IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received October 8, 2016, accepted October 25, 2016, date of publication November 14, 2016,
date of current version November 28, 2016.

Digital Object Identifier 10.1109/ACCESS.2016.2627243

A Joint Tensor Completion and Prediction Scheme for Multi-Dimensional Spectrum Map Construction

MENGYUN TANG[1], GUORU DING[1,2], (Senior Member, IEEE), QIHUI WU[3], (Senior Member, IEEE), ZHEN XUE[1], AND THEODOROS A. TSIFTSIS[4], (Senior Member, IEEE)

[1]College of Communications Engineering, PLA University of Science and Technology, Nanjing 210007, China
[2]National Mobile Communications Laboratory, Southeast University, Nanjing 210018, China
[3]College of Electronic and Information Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China
[4]School of Engineering, Nazarbayev University, Astana 010000, Kazakhstan

Corresponding author: G. Ding (dr.guoru.ding@ieee.org) and Q. Wu (wuqihui2014@sina.com)

This work was supported in part by the National Natural Science Foundation of China under Grant61631020, Grant 61501510, and Grant 61301160, in part by the Natural Science Foundation of Jiangsu Province under Grant BK20150717, in part by the China Postdoctoral Science Foundation Funded Project, and in part by the Jiangsu Planned Projects for Postdoctoral Research Funds.

ABSTRACT Spectrum data, which are usually characterized by many dimensions, such as location, frequency, time, and signal strength, present formidable challenges in terms of acquisition, processing, and visualization. In practice, a portion of spectrum data entries may be unavailable due to the interference during the acquisition process or compression during the sensing process. Nevertheless, the completion work in multi-dimensional spectrum data has drawn few attention to the researchers working in the field. In this paper, we first put forward the concept of spectrum tensor to depict the multi-dimensional spectrum data. Then, we develop a joint tensor completion and prediction scheme, which combines an improved tensor completion algorithm with prediction models to retrieve the incomplete measurements. Moreover, we build an experimental platform using Universal Software Radio Peripheral to collect real-world spectrum tensor data. Experimental results demonstrate that the effectiveness of the proposed joint tensor processing scheme is superior than relying on the completion or prediction scheme only.

INDEX TERMS Spectrum tensor, tensor completion, tensor prediction, cognitive radio.

I. INTRODUCTION

A. BACKGROUND AND MOTIVATION

With the skyrocketing development of the mobile internet, the shortage of spectrum resource is critical to be tackled [1], [2]. Several approaches have been suggested to meet the traffic growth [3], e.g., spectrum refarming, spectrum sharing and network densification, etc. Among these approaches, cognitive radio (CR) is considered as a promising way to improve the utilization of radio spectrum [4]. Specifically, the CR technology [5] allows the unlicensed secondary users to access the spectrum holes unoccupied by the licensed primary users opportunistically [6]. However, it's doubtful that unlicensed users would not interfere with the incumbents without the knowledge of temporal and spatial information of these bands. Therefore, it is well recognized that it's necessary to construct and manage a spectrum database to obtain the temporal and spatial spectrum availability information [7].

B. RELATED WORK

Relying on the sensors collecting the spectrum measurements, spectrum map can obtain the distribution of signal strength and estimate the spectrum utilization in a particular region, it can be seen as a visible cartography according to the geo-location database. For example, the TV white space coverage map, which combines terrain data and propagation models together to determine the max transmission power of each channel [8], enables opportunistic transmission across the available channels. The author in [9] presents sensing approaches to learn the power spectrum density (PSD) maps of spatial fields, and minimized communication overhead via measurement compression and quantization. In addition, many schemes are focused on the application of the spectrum map. In [10], the author constructs spectrum map from real-world data and estimates different radio and network performance metrics. The author in [11] proposes a joint indoor localization and spectrum map construction scheme that can be directly deployed with limited calibration load in indoor environments. DARPA has also launched its Advanced RF Mapping program, known as Radio Map, which aims to provide real-time awareness of radio spectrum use across frequency, geography and time [12]. Radio Map is a developing technology that visually overlays spectrum information on a map enabling rapid frequency deconfliction and maximizing use of available spectrum. However, most of the existing studies on the spectrum map construction solely describe the distribution of the signal strength in space, while the frequency and time information are overlooked. Moreover, if we take the indirect spectrum data into consideration, more factors such as the user data, terrain data, and meteorological and hydrographic data [13] would lead to different result.

In our previous work, we consider two dimensional spectrum map as the first step, which is presented in [14]. The goal of that work is to construct a TV white space database which collects the spectrum signal strengths from white space devices. In this paper, we extend our work to model the multi-dimensional spectrum data from the perspective of tensor [15]. We design and conduct an experiment to collect the real-world signal strength measurements, and compare the effectiveness of the introduced spectrum tensor processing scheme to the matrix completion algorithm.

C. CONTRIBUTIONS

In order to depict the multi-dimensional characteristics of the spectrum data, in this paper tensor pattern is introduced to represent the extending work from a 2D spectrum map to 3D or higher dimensional spectrum map. A tensor can be represented as an organized multi-dimensional array of numerical values, which has been widely applied in computer vision and graphics field [16], [17], [18]. Different from the prior studies, we propose the concept of the "spectrum tensor" and model the spectral manifold information in a novel way. The integration of the tensor into spectrum data is a promising paradigm to extend the study of spectrum map in depth. Specifically, it makes sense to estimate the spectrum tensor with low-cost sensors distributed across a particular domain since only signal strength is required. In order to facilitate practical implementations with sensor networks, where the number of communication devices is limited, two methods are proposed to reduce the overhead. One is to decrease the number of sensing sensors, and completion scheme is required to reconstruct the unknown measurements. The other is to incorporate prediction models to capture the relevance among the frequency or time domains via mining available prior information.

Briefly, in this paper, we conduct an experiment using Universal Software Radio Peripheral (USRP) to collect real-world spectrum tensor data, and propose a joint tensor completion and prediction scheme to retrieve the incomplete measurements. The main contributions throughout this paper are summarized as follows:

- Model the multi-dimensional spectrum data from the perspective of a spectrum tensor. To our best knowledge, this is the first work to put forward the concept of the spectrum tensor.
- Formulate the problem of spectrum tensor completion and present data analysis of real-world spectrum measurements to reveal the low rank characteristic of the spectrum data.
- Improve the low rank tensor completion algorithm, and evaluate it by comparing the improved spectrum tensor completion, the original one, and the spectrum matrix completion scheme.
- Propose a joint tensor completion and prediction scheme, and obtain an advanced performance than the state-of-the-art schemes.

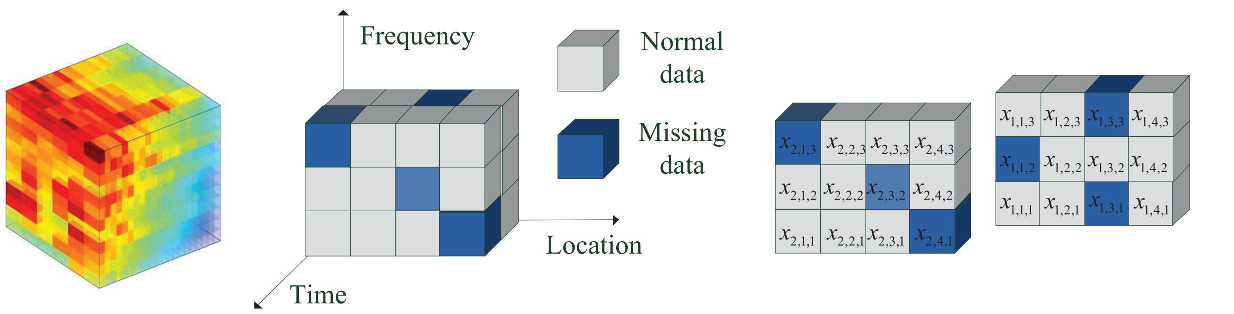
D. ORIGINATION AND NOTATION

The remainder of the paper is organized as follows. Section II presents the tensor model and the algebraic formulation. The joint tensor completion and prediction scheme is proposed in Section III. The designed experiment and the experimental performance are provided in section IV. The conclusions are drawn in Section V.

We use upper case letters for matrices, e.g., M, and lower case letters for the entries, e.g., $x_{t,l,f}$. The Frobenius norm of the matrix M is defined as $\| M\|_F = (\sum_{i,j} |x_{ij}|^2)^{\frac{1}{2}}$. For a tensor, let $\chi \in R^{I_1 \times I_2 \times \ldots \times I_N}$ denote a n-order spectrum tensor with its $(i_1, i_2, \ldots, i_N)$-th element denoted by $x_{i_1,i_2,\ldots,i_n}$. Here, the mode N of a tensor is the number of dimensions. Unfolding is an operation that transforms a spectrum tensor into a series of matrices. Specifically, a tensor χ can be unfolded along $i - th$ mode, denoted as $\chi(i)$.

II. MODEL FOR SPECTRUM TENSOR

An n-order spectrum tensor can be defined as $\chi \in R^{I_1 \times I_2 \times \ldots \times I_N}$, with $x_{i_1,i_2,\ldots,i_n}$ denoting its $(i_1, i_2, \ldots, i_N)$-th element. Generally speaking, the location is a two- dimensional parameter in the plane view, and the location is spatially viewed as a three-dimensional parameter [19]. However, in order to simplify the dimensions of multi-dimensional spectrum map, we assume only one-dimensional parameter to represent the information of the location, and take other factors like frequency and time into account. In Fig. 1(a), we show the real-world spectrum tensor collected in our experiment. It can be modeled as a 3-order spectrum tensor $\chi \in R^{I_1 \times I_2 \times I_3}$ in Fig. 1(b), whose x-axis is the indices of time slots, y-axis is the indices of continuous locations, and z-axis is the indices of frequency bands. Each tensor element $x_{t,l,f}$ denotes the received signal strength in the t-th time slot, the l-th location and the f-th frequency band. As a generalization of a two-dimensional matrix to higher dimensions, also known as ways or modes, a tensor can be unfolded as matrices along each mode. As an example, the 3-order spectrum tensor χ can be unfolded as follows (1), (2), and (3) are shown at the bottom of the next page, where $\chi_{(f)}$, $\chi_{(t)}$, $\chi_{(l)}$ denote the tensor unfolded by frequency-mode, time-mode,



(a) Real-world spectrum tensor (b) Spectrum tensor formulation (c) Spectrum tensor unfolded to spectrum matrix

FIGURE 1. Modeling for spectrum tensor.

and location-mode, respectively. The unfolded matrices contain all the entries inside the original tensor.

In practice, a portion of spectrum data entries may be unavailable due to the interference in the acquisition process, or the compression strategies in the sensing process would make the data incomplete [20]. Moreover, when mapping a spectrum tensor requires a large amount of measurements with sensor networks, the overhead of time or energy resources can be narrowed down by means of sampling. In this context, an effective tensor reconstruction method would be essential for recovering incomplete spectrum tensor from a few known samples.

As shown in Fig. 1(c), the white entries denote the normal entries, while the entries in dark color denote the incomplete or unavailable entries. Given a subset Ω, data in the subset Ω indicate the observed entries in χ, the other "unknown" entries would set to be "0." The subset tensor χ_Ω can be defined as

$$\chi_\Omega = \begin{cases} \chi, & (i_1, i_2, \ldots, i_n) \in \Omega \\ 0, & otherwise. \end{cases} \tag{4}$$

Then, the spectrum tensor completion for the "unknown" value can be modeled as an optimization problem which completes the tensor data using the known subset, implemented as the following problem

$$\min_{\chi} \| \chi \|_* \quad s.t. \ \chi_\Omega = \Gamma_\Omega, \tag{5}$$

where $\| . \|_*$ originally denotes the tightest convex envelop for the rank of matrices, χ_Ω, Γ_Ω are n-mode spectrum tensors with the same size in each mode, entries of Γ from the set Ω are given while the remaining entries are missing and χ is the incomplete spectrum tensor to be completed. The tensor can be seen as the expansion from the matrix perspective. Generally, low rank matrix completion issue can be solved by defining of the norm to approximate the rank of the matrix, but computing the rank of a tensor (mode number > 2) is an NP hard problem [21]. Thus, we introduce the following definition for the tensor trace norm [24]

$$\| \chi \|_* := \sum_{i=1}^{n} \alpha_i \| \chi_{(i)} \|_*, \tag{6}$$

where $\alpha_i \geq 0$ and $\sum_{i=1}^{n} \alpha_i = 1$, $||\chi_{(i)}||_*$ denotes the norm of unfolded matrix $\chi_{(i)}$. The trace norm of a tensor is consistent with all the matrices unfolded along each mode. Under this definition, the optimization problem can be rewritten as

$$\min_{\chi} \sum_{i=1}^{n} \alpha_i \| \chi_{(i)} \|_* \quad s.t. \ \chi_\Omega = \Gamma_\Omega. \tag{7}$$

III. JOINT TENSOR COMPLETION AND PREDICTION

In this section, we propose a novel joint scheme which combines a prediction scheme with the tensor completion algorithm together. As a spectrum tensor can be seen as a

$$\chi_{(f)} = \left[\begin{array}{c|c} x_{1,1,3}, x_{1,2,3}, x_{1,3,3}, x_{1,4,3} & x_{2,1,3}, x_{2,2,3}, x_{2,3,3}, x_{2,4,3} \\ x_{1,1,2}, x_{1,2,2}, x_{1,3,2}, x_{1,4,2} & x_{2,1,2}, x_{2,2,2}, x_{2,3,2}, x_{2,4,2} \\ x_{1,1,1}, x_{1,2,1}, x_{1,3,1}, x_{1,4,1} & x_{2,1,1}, x_{2,2,1}, x_{2,3,1}, x_{2,4,1} \end{array} \right], \tag{1}$$

$$\chi_{(t)} = \left[\begin{array}{c|c|c} x_{1,1,1}, x_{1,2,1}, x_{1,3,1}, x_{1,4,1} & x_{1,1,2}, x_{1,2,2}, x_{1,3,2}, x_{1,4,2} & x_{1,1,3}, x_{1,2,3}, x_{1,3,3}, x_{1,4,3} \\ x_{2,1,1}, x_{2,2,1}, x_{2,3,1}, x_{2,4,1} & x_{2,1,2}, x_{2,2,2}, x_{2,3,2}, x_{2,4,2} & x_{2,1,3}, x_{2,2,2}, x_{2,3,3}, x_{2,4,3} \end{array} \right], \tag{2}$$

$$\chi_{(l)} = \left[\begin{array}{c|c} x_{1,1,1}, x_{1,1,2}, x_{1,1,3} & x_{2,1,1}, x_{2,1,2}, x_{2,1,3} \\ x_{1,2,1}, x_{1,2,2}, x_{1,2,3} & x_{2,2,1}, x_{2,2,2}, x_{2,2,3} \\ x_{1,3,1}, x_{1,3,2}, x_{1,3,3} & x_{2,3,1}, x_{2,3,2}, x_{2,3,3} \\ x_{1,4,1}, x_{1,4,2}, x_{1,4,3} & x_{2,4,1}, x_{2,4,2}, x_{2,4,3} \end{array} \right], \tag{3}$$

composition of many spectrum maps, this trial can take the advantage of the correlation between the inherent regularity of these spectrum maps and the completed tensor measurements.

A. IMPROVED TENSOR COMPLETION ALGORITHM

So far, a lot of algorithms have been proposed to complete the tensor data. The authors in [22] provide a new compressive sensing reconstruction formula for multi-dimensional measurements. A fully Bayesian CP factorization which can naturally handle incomplete and noisy tensor data is presented in [23]. In this paper, we propose an improved tensor completion algorithm of an efficient algorithm presented in [24]. The algorithm can work even with a small amount of samples and estimate larger missing regions of the spectrum data.

Because of the interdependence among the matrix trace norm terms, the problem in (7) is difficult to be tackled. In order to simplify the original problem, additional matrices $M_1, \ldots, M_n$ are introduced to split these interdependent terms, so that they can be solved independently. Hence, we can obtain the following equivalent formulation

$$\begin{aligned}&\min_{\chi, M_i} \sum_{i=1}^{n} \alpha_i \parallel M_i \|_* \\ &s.t.\ \ \chi_{(i)} = M_i \quad \text{for } i = 1, \ldots, n \\ &\qquad \chi_\Omega = \Gamma_\Omega. \end{aligned} \tag{8}$$

The algorithm HaLRTC (high accuracy low rank tensor completion) applies the alternating direction method of multipliers (ADMM) to solve the convex optimization problem by breaking them into smaller pieces [25], and each of them becomes easier to handle. Firstly, let the matrices M_is be replaced with their tensor versions φ_i

$$\begin{aligned}&\min_{\chi, \varphi_1, \ldots, \varphi_n} \sum_{i=1}^{n} \alpha_i \parallel \varphi_{i(i)} \|_* \\ &s.t.\ \ \chi = \varphi_i, \quad \text{for } i = 1, \ldots, n \\ &\qquad \chi_\Omega = \Gamma_\Omega. \end{aligned} \tag{9}$$

Based on this formulation, we can observe that α is a scale vector. In this sense, in order to tackle this problem, it is equivalent to figure out which $i-th$ dimensional unfolded matrix would have the minimum $\parallel \varphi_i \|_*$. Thus, we introduce the following liner convex optimization problem for α independently

$$\begin{aligned}&\min \sum_{i=1}^{n} \alpha_i \parallel \varphi_i \|_* \\ &s.t.\ \ \chi = \varphi_i, \quad i = 1, \ldots, n \\ &\sum_{i=1}^{n} \alpha_i = 1. \end{aligned} \tag{10}$$

Then, the augmented Lagrangian function is defined as follows

$$\begin{aligned}&L_p(\chi, \varphi_i, \gamma_i, \alpha_i) \\ &\quad = \sum_{i=1}^{n} \alpha_i \|\varphi_{i(i)}\|_* + \langle \chi - \varphi_i, \gamma_i \rangle + \frac{\rho_1}{2} \parallel \chi - \varphi_i \|_F^2 \\ &\qquad + \mu(\sum_{i=1}^{n} \alpha_i - 1) + \frac{\rho_2}{2} \parallel \sum_{i=1}^{n} \alpha_i - 1 \|_F^2, \end{aligned} \tag{11}$$

where ρ_1, ρ_2 are the penalty parameters of the two penalty functions, respectively. According to the framework of ADMM, the augmented Lagrangian function is applied to update φ_i,χ ,γ_i and α_i iteratively

$$\begin{aligned}&\left\{\varphi_1^{k+1}, \ldots, \varphi_n^{k+1}\right\} \\ &\quad = \arg\min_{M_1, \ldots, M_n} L_\rho(\chi^k, \varphi_1, \ldots, \varphi_n, y_1^{k+1}, \ldots, y_n^{k+1}), \end{aligned} \tag{12}$$

$$\chi^{k+1} = \arg\min_{\chi \in Q} L_\rho \left(\chi, \varphi_1^{k+1}, \ldots \varphi_n^{k+1}, y_1^k, \ldots, y_n^{k+1}\right), \tag{13}$$

$$y_i^{k+1} = y_i^k - \rho_1(\varphi_i^{k+1} - \chi^{k+1}), \tag{14}$$

$$\mu_i^{k+1} = \mu_i^k - \rho_2(\sum_{i=1}^{n} \alpha_i^{k+1} - 1). \tag{15}$$

This algorithm can also be accelerated with an increasing value of ρ_1, and ρ_2. Although the convergence rate is not the fastest, the convergence of the general ADMM algorithm is guaranteed [25]. When the optimization is finished, we can obtain the optimized α vector that satisfies the constraints from the default value, which means the weight of the best completion performance of the unfolded matrix is 1, while the others are 0. The optimized algorithm is summarized in the Algorithm I. The shrinkage operator $D_\tau(X)$ indicates $D_\tau(X) = U \sum_\tau V^T$, and lines 5, 7, 8, 9 mentioned herein correspond to the optimized augmented Lagrangian function.

B. A JOINT TENSOR COMPLETION AND PREDICTION SCHEME

Spectrum prediction can forecast future unknown spectrum maps from previous collected spectrum data by exploiting the correlation among them, and provide forecast views of the spectrum utilization of various wireless applications. For a linear dynamic systems with Gaussian noise, the Kalman filter is the minimum-variance state estimator. Even if the noise is non-Gaussian, the Kalman filter can also derive the optimal state estimator in closed form. Since the signal strength measurements change with the radiation source, we can assume that in a short time, the evolution of spectrum maps can be seen as a linear dynamic system.

1) DESIGN RATIONALE

Reviewing the existing studies, prediction schemes have numerous applications for technologies as it can infer the

Algorithm 1 α Optimized HaLRTC

1: **Initialize:** Set K as the number of iterations, N as the dimensions of the spectrum tensor, and θ as a scale vector parameter. The parameter ρ is set to be 1×10^{-6}, and the temporally used variables are set as $\varphi = \mathbf{0}$, $\gamma = \mathbf{0}$, $\mu = \mathbf{0}$.
2: **Input:** χ with $\chi_\Omega = \Gamma_\Omega$, K and α.
3: **for** $k = 1, 2, \ldots K$ **do**
4:   **for** $i = 1, \ldots, N$ **do**
5:     $\varphi_i = \mathrm{fold}_i \left[D_{\frac{\alpha_i}{\rho}} \left(\chi_i + \frac{1}{\rho}\gamma_{i(i)} \right) \right]$.
6:   **end for**
7:   $\chi_\Omega = \frac{1}{n} \left(\sum_{i=1}^{n} \varphi_i - \frac{1}{\rho}\gamma_i \right)_{\bar{\Omega}}$;
8:   $\gamma_i = \gamma_i - \rho_1(\varphi_i - \chi)$;
9:   $\mu_i = \mu_i - \rho_2(\sum_{i=1}^{n} \alpha_i - 1)$.
10: **end for**
11: **Output:** The completed spectrum data χčand the scale vector α .

state evolution. Kalman filter is one of these common prediction schemes, which is an iterative algorithm that obtain the results from the consecutive cycles of prediction and filtering [26]. It is a widely applied concept in time series analysis used in fields such as signal processing and econometrics, and it is also one of the main topics in robotic motion planning and control [27] and positioning systems [28].

Prediction and iteration are the two key steps of Kalman filtering. The spectrum tensor calculated by the improved tensor completion algorithm can be seen as observation value, while a linear prediction model also produces a prediction tensor. The iteration of Kalman gain would adapt the weighted average more and more accurate from the observation and prediction results. This joint prediction and completion scheme does not take any assumption errors into account, due to the inaccuracy of the tensor completion algorithm.

2) SCHEME DESIGN

In this paper, we apply the core idea of Kalman filtering to establish a connection between the estimate measurements from the tensor completion algorithm and the prediction results. This connection endows the completion scheme with the universality of Kalman prediction and provides a new solution to figure the optimized problem.

Firstly, a linear prediction model should be provided for making estimates of the incomplete current spectrum state, since the spectrum tensor can be decomposed into many maps. We retain a complete spectrum map as the prior knowledge of state, and then divide the incomplete spectrum tensor into two categories. One is a training dataset for building the prediction model, while the other is the input dataset. In the training step, the remaining entries of the training dataset are compared with the reference spectrum map, and the statistical learning tool is adopted to obtain a linear fitting curve as a prediction model

$$\eta_{\bar{\Omega}}(k/k) = A\eta_{\bar{\Omega}}(k/k-1) + B, \tag{16}$$

where η is the predicted spectrum tensor. The incomplete entries, which are not included in Ω would be retrieved via this prediction model. The current $\eta_{\bar{\Omega}}(k/k)$ can be forecast by the previous spectrum maps $\eta_{\bar{\Omega}}(k/k-1)$. A and B are the state vector parameters, which is acquired by training the input dataset with a curve fitting tool. In this experiment, A is set to be 1.027, and B is 2.224. Then in the prediction step, we retain the existing entries in Ω, and calculate the estimations of the incomplete current spectrum map by a curve fitting tool based on the input dataset.

With the finished complete predicted tensor, the spectrum tensor would be updated using the Kalman gain to return a balance between the completed spectrum tensor and the predicted spectrum tensor.

$$\eta(k/k) = \eta(k/k-1) + kg(k)(\chi(k) - \eta(k/k-1)). \tag{17}$$

Once the outcome of the measurements from the tensor completion scheme $\chi(k)$ is observed, a dynamic weight kgčwhich is also referred to Kalman gain, would be assigned to achieve the final integrated spectrum tensor $\eta(k/k)$. The gain is determined by comparing the observation error of the completion scheme with the prediction scheme.

$$kg(k) = \frac{\Delta(\eta(k/k))}{\Delta(\eta(k/k)) + \Delta(\chi(k))}, \tag{18}$$

where $0 \leq kg \leq 1$, Δ refers to the error of observation and prediction value. Due to the collected original complete spectrum tensor, we can calculate the error of the completed and predicted tensor, respectively. However, in actual fact, the observation error is usually unknown in the real-world experiment, and it can be set at random with Gaussian distribution.

3) IMPLEMENTATION

By employing the core idea of the Kalman prediction algorithm, the joint completion and prediction scheme include two flow paths, which we can observe from Fig. 2. one is the advanced tensor completion algorithm, and the other is Kalman filtering prediction process. Both of the two paths can produce a reconstructed spectrum tensor, respectively. Since A 3-D spectrum tensor can be seen as a composer of many 2-D spectrum maps, a prior knowledge of state is necessary for building a prediction model, which means we should collect a complete spectrum map, before monitoring environmental signal strength measurements with several sensors. When the incomplete dataset is captured, some of the spectrum maps would be compared with the prior spectrum map, to train a fitting model for the next input dataset. Meanwhile, the incomplete spectrum tensor would be processed by the advanced tensor completion algorithm, and produce a complete spectrum tensor. It can run in an online mode, when we collect the signal strength measurements, the joint

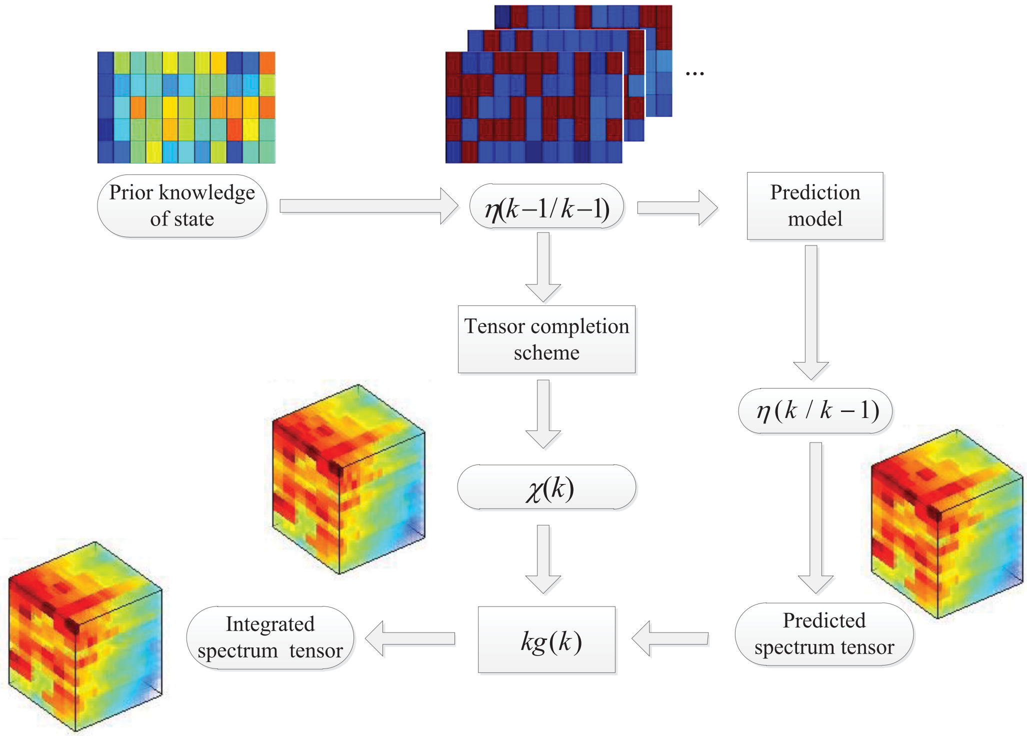


FIGURE 2. The flow chart of the joint tensor completion and prediction scheme.

scheme can take these current entries as new input datasets, and process these maps in real-time.

The dynamic weight is an important parameter that winds these two states together. It reveals that which scheme the final integrated result tends to believe more. For example, when the amount of remaining entries of a spectrum tensor is small, tensor completion scheme can hardly play an excellent role, but the prediction model learning from the previous tensor measurements may lead to a more accurate result. On the contrary, when the weight is large, it means the final state estimates trust the prediction results more than tensor completion results.

IV. PERFORMANCE EVALUATION

In the section, we design and conduct an experiment to collect the real-world signal strength measurements, and experimental results demonstrate the effectiveness of the introduced spectrum tensor processing scheme over the state-of-the-art spectrum matrix processing scheme.

A. EXPERIMENT SETUP

In this experiment, we used five USRPs (Universal Software Radio Peripheral), equipped with an antenna for each, to collect the real-world signal strength measurements. The experiment utilizes one USRP as the signal transmitter and the other four USRPs as the receivers, and the table lists the parameters of the transmitter and the receivers. Firstly, we sweep the frequency and find out that the band between 580 MHz-690 MHz is relatively idle. This band is included in the TV white space [7]. Then, we choose 660 MHz to be the central frequency to transmit the sine wave patten. In order to acquire the real-world multi-dimensional spectrum tensor, we divide the lab room into 6 × 12 grids and measure the received signal strength of these grids.

TABLE 1. Transmitting and receiving parameters

Parameter	Value
Transmitting IQ rate	1MHz
Carrier frequency	660MHz
Tune frequency	10KHz
Receiving IQ rate	10MHz
Samples	1000
Resolution	10KHz
Transmitting gain	0:1:20

The experiment performed considers 100 channels of bandwidth 10 kHz each in the TV white space band, and we change the transmitting power by adjusting the gain parameter. The signal strength at different locations received from the transmitter is ranged from -80dB to -30dB, and the size of the collected spectrum tensor in our experiment is $6 \times 12 \times 21$.

B. RANK ANALYSIS

In order to characterize the rank of the spectrum tensor data, we first make a preprocessing by unfolding the tensor into matrices along each mode. Then, we apply singular value decomposition (SVD) [29] to analyze the rank distribution of all unfolded spectrum matrices.

In Fig. 4, we plot the normalized singular values in a descending order for the unfolded matrices along various

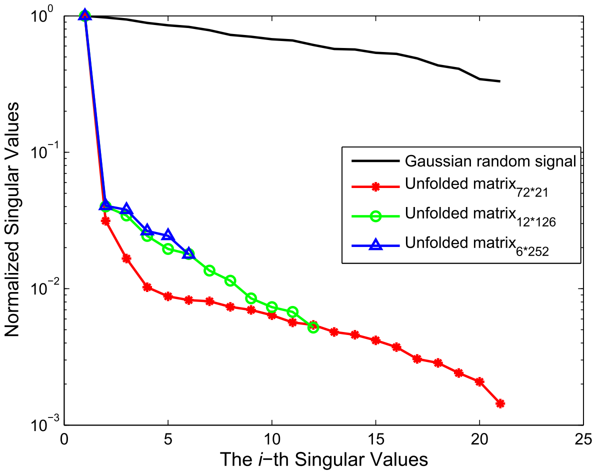


FIGURE 3. Visualization for the completion process of the real-world measurements. (a) A three-dimensional tensor collected by USRPs. (b) We randomly remove 50% entries from the original spectrum tensor measurements. (c) The reconstruction result by the tensor completion algorithm.

locations (x-axis and y-axis) and transmitting powers, respectively. For comparison, an independent and identically distributed (i.i.d.) Gaussian random signal dataset of the same size is also analyzed. It is suggested that, the distribution of the signal strength is always concentrated on the top several singular values in the measured real-world spectrum tensor, which is quite different from the Gaussian random signal dataset. This phenomenon reveals that practical spectrum tensor exhibits approximate low-rank structure.

C. EXPERIMENTAL RESULTS

We perform simulations to evaluate the tensor completion algorithm we introduced above. Suppose that we sample the spectrum tensor data at random with the sampling rate sr, where sr indicates the percentage of the observed elements in the tensor.

Under the representative simulation, Fig. 3 presents the completion process of the tensor completion scheme. The real-world tensor data is shown in Fig. 3(a), where the three-dimensional axis denote the different mode of the tensor, and different colors of the grids denote different signal strength. Then, we do the sampling randomly that only 50% entries are retained in the spectrum tensor measurements (see Fig. 3(b). The reconstruction performance of the tensor completion algorithm is shown in Fig. 3(c)), which verifies the effectiveness of the incomplete data estimation.

In order to measure the recovery performance of the tensor completion algorithm, we define the root square error (RSE) in dB as follows

$$RSE\ [dB] = 10\log_{10}\frac{||\tilde{\chi} - \chi||_2}{||\chi||_2}, \tag{19}$$

where $\tilde{\chi}$ is the reconstructed tensor and χ is the original real-world spectrum tensor. Fig. 5 compares the RSE of the tensor completion scheme with the matrix completion scheme against various sampling rates. Here we select a fast and robust matrix completion algorithm in widely use [30], which can solve large matrix rank minimization problems. Three dashed lines indicate the RSEs of the matrix completion scheme for the unfolded matrix by different mode, while the two lines indicate the performance of the two tensor completion scheme. We can observe that the performance of the *alpha* optimized scheme has about 2 dB better than the average weight for the original tensor completion algorithm. It is also observed that, when sampling rate is less than 0.3 or larger than 0.7, the RSE of the tensor completion scheme has better performance than the matrix completion scheme, while on the contrary, the RSE is not always better than all the unfolded matrix completion results.

It has to be mentioned that in Fig. 5, the matrix completion scheme cannot reconstruct the incomplete data because the sampling rate is too low, according to [30]. Given the rank of the matrix $m \times n$ is r, we sample a subset of p entries and list $FR = r(m+n-r)/p$, and if $FR > 1$, then there is always an infinite number of matrices with rank r with the given entries,

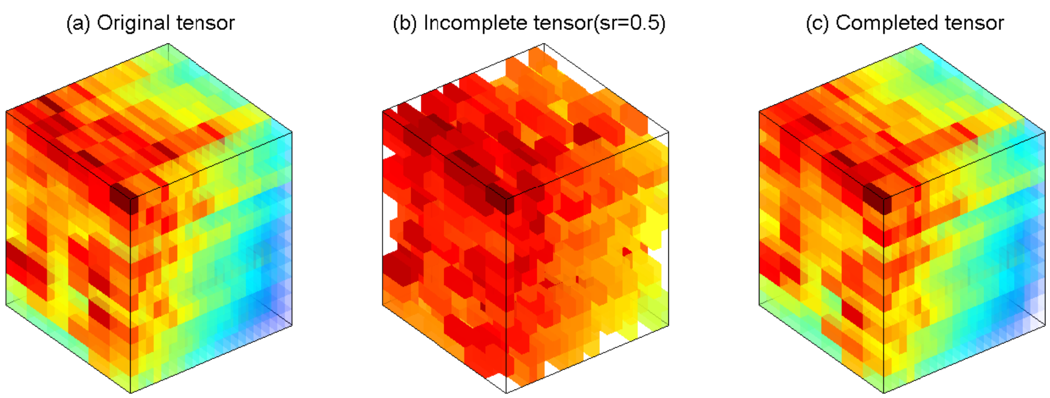


FIGURE 4. Normalized singular values of unfolded spectrum tensor datasets.

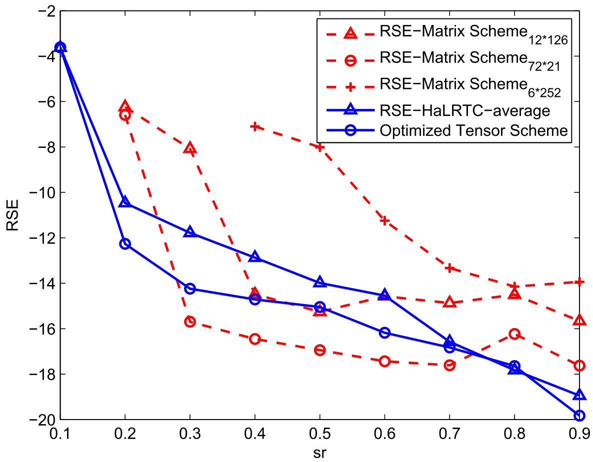


FIGURE 5. RSE performance of the two tensor completion scheme and the unfolded matrix scheme completion.

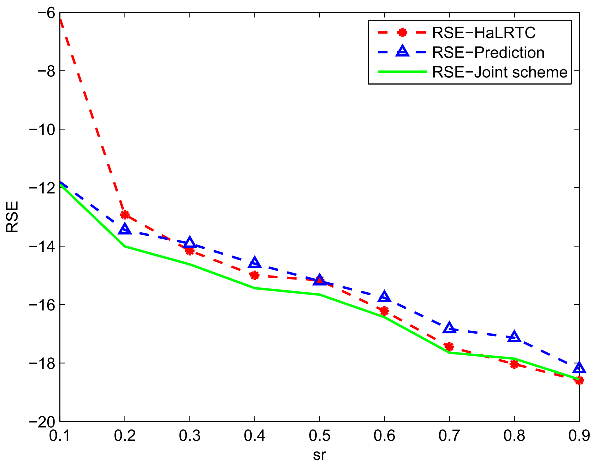


FIGURE 6. RSE performance of the online joint tensor completion and prediction scheme.

and we can hardly recover the incomplete data successfully in this situation.

Fig. 6 shows the comparison of the joint tensor completion and prediction scheme. The two dashed lines indicate the RSE of the tensor completion, and relying on the prediction scheme only. It is obvious that when *sr* is low, which means the amount of valid observed measurements is small, and the prediction model is relative accurate and stable. When *sr* is high, tensor completion scheme has a better RSE performance than the prediction model. The joint scheme is represented as the line in Fig. 6, which combines the advantage of the reliability of the prediction model when *sr* is low, and the accuracy of the tensor completion when *sr* is relatively high. The performance of the joint scheme is superior when *sr* is below 0.5, and is close to the tensor completion line when over 0.5.

V. CONCLUSION

In this paper, a novel spectrum tensor concept is coined to depict the multi-dimensional spectrum data and a joint tensor completion and prediction scheme is proposed to interpolate the incomplete or unknown spectrum measurements. The experimental results with real-world spectrum measurements show that the superiority of the joint scheme is over the state-of-the-art studies. For the future work, we would take more factors such as the user data, terrain data, and meteorological and hydrographic data into account, and improve the accuracy of the spectrum tensor.

ACKNOWLEDGMENT

The authors would like to thank the Editor and Reviewers for their precious time and efforts in reviewing our paper. They also thank Prof. Xiaofei Zhang for his helpful discussions.

REFERENCES

[1] R. Liu, G. Yu, F. Qu, and Z. Zhang, "Device-to-device communications in unlicensed spectrum: Mode selection and resource allocation," *IEEE Access*, vol. 4, pp. 4720–4729, 2016.
[2] Y. Wu, R. Schober, D. W. K. Ng, C. Xiao, and G. Caire, "Secure massive MIMO transmission with an active eavesdropper," *IEEE Trans. Inf. Theory*, vol. 62, no. 7, pp. 3880–3900, Jul. 2016.
[3] Q. C. Li, H. Niu, A. T. Papathanassiou, and G. Wu, "5G network capacity: Key elements and technologies," *IEEE Veh. Technol. Mag.*, vol. 9, no. 1, pp. 71–78, Mar. 2014.
[4] X. L. Huang, F. Hu, J. Wu, H. H. Chen, G. Wang, and T. Jiang, "Intelligent cooperative spectrum sensing via hierarchical Dirichlet process in cognitive radio networks," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 5, pp. 771–787, May 2015.
[5] J. Mitola and G. Q. Maguire, Jr., "Cognitive radio: Making software radios more personal," *IEEE Pers. Commun.*, vol. 6, no. 4, pp. 13–18, Apr. 1999.
[6] G. Ding, Q. Wu, Y.-D. Yao, J. Wang, and Y. Chen, "Kernel-based learning for statistical signal processing in cognitive radio networks: Theoretical foundations, example applications, and future directions," *IEEE Signal Process. Mag.*, vol. 30, no. 4, pp. 126–136, Jul. 2013.
[7] G. Ding, J. Wang, Q. Wu, Y.-D. Yao, F. Song, and T. A. Tsiftsis, "Cellular-base-station-assisted device-to-device communications in TV white space," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 1, pp. 107–121, Jan. 2016.
[8] R. Murty, R. Chandra, T. Moscibroda, and P. Bahl, "SenseLess: A database-driven white spaces network," *IEEE Trans. Mobile Comput.*, vol. 11, no. 2, pp. 189–203, Feb. 2012.
[9] D. Romero, S.-J. Kim, G. B. Giannakis, and R. Lopez-Valcarce. (2016). "Learning power spectrum maps from quantized power measurements." [Online]. Available: http://arxiv.org/abs/1606.02679
[10] S. Debroy, S. Bhattacharjee, and M. Chatterjee, "Spectrum map and its application in resource management in cognitive radio networks," *IEEE Trans. Cognit. Commun. Netw.*, vol. 1, no. 4, pp. 406–419, Dec. 2015.
[11] S. Sorour, Y. Lostanlen, S. Valaee, and K. Majeed, "Joint indoor localization and radio map construction with limited deployment load," *IEEE Trans. Mobile Comput.*, vol. 14, no. 5, pp. 1031–1043, May 2015.
[12] *Advanced RF Mapping (Radio Map)*, accessed on Dec. 11, 2016. [Online]. Available: http://www.darpa.mil/program/advance-rf-mapping
[13] Q, Wu, G, Ding, Z, Du, Y, Sun, M, Jo, and A. V. Vasilakos, "A cloud-based architecture for the Internet of spectrum devices over future wireless networks," *IEEE Access*, vol. 4, pp. 2854–2862, 2016.
[14] M. Tang, Z. Zheng, G. Ding, and Z. Xue, "Efficient TV white space database construction via spectrum sensing and spatial inference," in *Proc. IEEE Int. Perform. Comput. Commun. Conf.*, Dec. 2015, pp. 1–5.
[15] M. Tang, G. Ding, Z. Xue, J. Zhang, and H. Zhou, "Multi-dimensional spectrum map construction: A tensor perspective," in *Proc. IEEE Int. Conf. Wireless Commun. Signal Process.*, Oct. 2016, pp. 1–5.
[16] P. J. Basser, J. Mattiello, and D. LeBihan, "MR diffusion tensor spectroscopy and imaging," *Biophys. J.*, vol. 66, no. 1, pp. 259–267, 1994.
[17] D. Le Bihan *et al.*, "Diffusion tensor imaging: Concepts and applications," *J. Magn. Reson. Imag.*, vol. 13, no. 4, pp. 534–546, 2001.
[18] E. Acar, D. M. Dunlavy, T. G. Kolda, and M. Mørup, "Scalable tensor factorizations for incomplete data," *Chemometrics Intell. Lab. Syst.*, vol. 106, no. 1, pp. 41–56, 2011.

[19] N. Wu, Y. Xiong, H. Wang, and J. Kuang, "A performance limit of TOA-based location-aware wireless networks with ranging outliers," *IEEE Commun. Lett.*, vol. 19, no. 8, pp. 1414–1417, Aug. 2015.
[20] G. Tsagkatakis, B. Beferull-Lozano, and P. Tsakalides, "Singular spectrum-based matrix completion for time series recovery and prediction," *EURASIP J. Adv. Signal Process.*, vol. 2016, no. 1, p. 1, 2016.
[21] C. J. Hillar and L.-H. Lim, "Most tensor problems are NP hard," *Comput. Res. Repository*, 2009. [Online]. Available: http://arxiv.org/abs/0911.1393
[22] C. F. Caiafa and A. Cichocki, "Stable, robust, and super fast reconstruction of tensors using multi-way projections," *IEEE Trans. Signal Process.*, vol. 63, no. 3, pp. 780–793, Feb. 2015.
[23] Q. Zhao, L. Zhang, and A. Cichocki, "Bayesian CP factorization of incomplete tensors with automatic rank determination," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 9, pp. 1751–1763, Sep. 2015.
[24] J. Liu, P. Musialski, P. Wonka, and J. Ye, "Tensor completion for estimating missing values in visual data," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 1, pp. 208–220, Jan. 2013.
[25] M. Fazel, H. Hindi, and S. Boyd, "A rank minimization heuristic with application to minimum order system approximation," in *Proc. Amer. Control Conf.*, vol. 6. 2001, pp. 4734–4739.
[26] M. I. Ribeiro, "Kalman and extended Kalman filters: Concept, derivation and properties," *Inst. Syst. Robot.*, 2004.
[27] S. Y. Chen, "Kalman filter for robot vision: A survey," *IEEE Trans. Ind. Electron.*, vol. 59, no. 11, pp. 4409–4420, Nov. 2012.
[28] H. S. Ahn and K. H. Ko, "Simple pedestrian localization algorithms based on distributed wireless sensor networks," *IEEE Trans. Ind. Electron.*, vol. 56, no. 10, pp. 4296–4302, Oct. 2009.
[29] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An Algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4311–4322, Nov. 2006.
[30] S. Ma, D. Goldfarb, and L. Chen, "Fixed point and Bregman iterative methods for matrix rank minimization," *Math. Program.*, vol. 128, nos. 1–2, pp. 321–353, 2011.

MENGYUN TANG received the B.S. degree in electronic and information engineering from the Nanjing University of Post and Telecommunications, Nanjing, China, in 2014. She is currently pursuing the M.S. degree in information and communication engineering with the College of Communications Engineering, PLA University of Science and Technology. Her research interests focus on cognitive radio networks, machine learning, and big spectrum data analytics for future wireless networks.

GUORU DING (S'10–M'14–SM'16) received the B.S. degree (Hons.) in electrical engineering from Xidian University, Xi'an, China, in 2008, and the Ph.D. degree (Hons.) in communications and information systems from the College of Communications Engineering, Nanjing, China, in 2014. Since 2014, he has been an Assistant Professor with the College of Communications Engineering and a Research Fellow with the National High Frequency Communications Research Center, China. Since 2015, he has been a Post-Doctoral Research Associate with the National Mobile Communications Research Laboratory, Southeast University, Nanjing, China. His research interests include cognitive radio networks, massive MIMO, machine learning, and big data analytics over wireless networks.

Dr. Ding has acted as a Technical Program Committees member for a number of international conferences, including the IEEE Global Communications Conference, the IEEE International Conference on Communications, and the IEEE Vehicular Technology Conference (VTC). He is a Voting Member of the IEEE 1900.6 Standard Association Working Group. He was a recipient of the Best Paper Awards from the EAI MLICOM 2016, the IEEE VTC 2014-Fall, and the IEEE WCSP 2009. He currently serves as a Guest Editor of the IEEE Journal on Selected Areas in Communications and an Associate Editor of the *KSII Transactions on Internet and Information Systems*.

QIHUI WU (SM'13) received the B.S. degree in communications engineering and the M.S. and Ph.D. degrees in communications and information systems from the Institute of Communications Engineering, Nanjing, China, in 1994, 1997, and 2000, respectively. From 2003 to 2005, he was a Post-doctoral Research Associate with Southeast University, Nanjing. From 2005 to 2007, he was an Associate Professor with CCE, PLA UST, Nanjing, where he served as a Full Professor from 2008 to 2016. Since 2016, he has been a Full Professor with the College of Electronic and Information Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing. In 2011, he was an Advanced Visiting Scholar with the Stevens Institute of Technology, Hoboken, USA. His current research interests span the areas of wireless communications and statistical signal processing, with emphasis on system design of software defined radio, cognitive radio, and smart radio.

ZHEN XUE received the B.S. degree in electrical engineering from Xidian University, Xi'an, China, in 2012, and the M.S. degree in information and communication engineering from the College of Communications Engineering, PLA University of Science and Technology (CCE, PLA UST), in 2015. He is currently pursuing the Ph.D. degree with CCE, PLA UST. His research interests focus on wireless resource optimization, machine learning, and big spectrum data analytics for future wireless networks.

THEODOROS A. TSIFTSIS (S'02–M'04–SM'10) was born in Lamia, Greece, in 1970. He received the B.Sc. degree in physics from the Aristotle University of Thessaloniki, Greece, in 1993, the M.Sc. degree in digital systems engineering from Heriot-Watt University, Edinburgh, U.K., in 1995, the M.Sc. degree in decision sciences from the Athens University of Economics and Business, Greece, in 2000, and the Ph.D. degree in electrical engineering from the University of Patras, Greece, in 2006. He joined the Department of Electrical Engineering, Technological Educational Institute of Central Greece, in 2010. He is currently an Associate Professor of Communication Technologies with the Department of Electrical and Electronic Engineering, School of Engineering, Nazarbayev University, Astana, Kazakhstan. He has authored or co-authored over 100 technical papers in scientific journals and international conferences. His research interests include the broad areas of cooperative communications, communication theory, wireless communications, and optical wireless communication systems.

Dr. Tsiftsis acts as a Reviewer for several international journals. He was a member of the Editorial Board of the IEEE Transactions on Vehicular Technology and the IEEE Communications Letters. He is currently an Area Editor of the Wireless Communications II of the IEEE Transactions on Communications.