

DEEP GENERATIVE MODEL LEARNING FOR BLIND SPECTRUM CARTOGRAPHY WITH NMF-BASED RADIO MAP DISAGGREGATION

Sagar Shrestha[†], Xiao Fu[†], and Mingyi Hong^{*}

[†]School of EECS, Oregon State University

^{*}Department of ECE, University of Minnesota

(shressag, xiao.fu)@oregonstate.edu, mhong@umn.edu

ABSTRACT

Spectrum cartography (SC) aims at estimating the multi-aspect (e.g., space, frequency, and time) interference level caused by multiple emitters from limited measurements. Early SC approaches rely on model assumptions about the radio map, e.g., sparsity and smoothness, which may be grossly violated under critical scenarios, e.g., in the presence of severe shadowing. More recent data-driven methods train deep generative networks to distill parsimonious representations of complex scenarios, in order to enhance performance of SC. The challenge is that the state space of this learning problem is extremely large—induced by different combinations of key problem constituents, e.g., the number of emitters, the emitters’ carrier frequencies, and the emitter locations. Learning over such a huge space can be costly in terms of sample complexity and training time; it also frequently leads to generalization problems. Our method integrates the favorable traits of model and data-driven approaches, which substantially ‘shrinks’ the state space. Specifically, the proposed learning paradigm only needs to learn a generative model for the radio map of a *single* emitter (as opposed to numerous combinations of multiple emitters), leveraging a nonnegative matrix factorization (NMF)-based emitter disaggregation process. Numerical evidence shows that the proposed method outperforms state-of-the-art purely model-driven and purely data-driven approaches.

Index Terms— Radio map, spectrum cartography, cognitive radio, deep generative model, nonnegative matrix factorization

1. INTRODUCTION

Spectrum cartography (SC) aims at estimating the radio frequency (RF) interference level caused by emitters over multiple domains, e.g., space, time, and frequency—a multi-aspect radio map [1–4]. SC is the first step towards full RF awareness, and is considered a cornerstone for high-efficiency wireless resource management, networking, and user/system co-existence that are core tasks in future wireless communications systems [1].

SC is an ill-posed high-dimensional data sampling and recovery problem, since the radio map can only be sparsely measured in the domains of interest, e.g., space and time. In the past decade, a plethora of approaches were proposed to tackle this problem, using dictionary learning, kernel interpolation, and tensor decomposition; see, e.g., [2–6]. In a nutshell, these works can all be categorized

into the *model-based* family, which rely on model assumptions—e.g., that the radio map exhibits sparsity, smoothness, or/and low-rank structures in certain domains. Model-based approaches work to a certain extent, but serious challenges remain. For example, the work in [2] assumes that the emitters’ *spatial loss fields* (SLFs)—which models the power propagation in space—are low-rank matrices. However, this assumption could be easily violated when the shadowing effect is strong (e.g., in urban or indoor environments).

Recently, [7, 8] proposed a data-driven approach. There, the radio map is treated as a 3D image, and a neural network is employed to learn its generative model. This way, the burden of modeling is ‘shifted’ to an off-line learning process. The new challenge is that the state space of interest is extremely large, which contains all possible scenarios of different numbers of emitters, their frequency usage, their locations, and all possible fading/shadowing effects. Having a representative training set would cost a large amount of resources, let alone the costly training process. Learning over such a huge space could also lead to high generalization errors, since the ‘test data’ could be far from any train sample (e.g., the number of emitters can change drastically from training set to test set).

In this work, we propose a mixed model and data-driven framework to address these challenges. Our idea is to keep the well-modeled part in classic model-based approaches intact (i.e., the fact that the overall radio map is an aggregation of the radio maps of individual emitters) [2, 4, 9], while let a deep neural network to represent the complex/uncertain part in the signal generating process—particularly, the SLFs of emitters. This way, that the ‘modeling burden’ of the neural network is substantially lower relative to that in [7, 8]. To enable the training process, we propose a carefully designed nonnegative matrix factorization (NMF)-based signal disaggregation method, which ensures that the individual radio maps of different emitters can be separated from the incomplete observations, under reasonable conditions. Simulations show that our method outperforms a number of state-of-the-art baselines. In particular, our framework successfully overcomes the problem of emitter miss detection that is observed in existing purely data-driven methods, indicating substantially improved generalization performance.

2. PROBLEM STATEMENT AND BACKGROUND

We are interested in a spatio-spectral radio map recovery problem as in [2, 4, 6, 9]. Consider a two-dimensional geographical region that is discretized into $I \times J$ grids. We hope to estimate the power spectrum density (PSD) over K frequency bins at every grid—which gives rise to an $I \times J \times K$ radio map tensor, i.e., $\underline{\mathbf{X}} \in \mathbb{R}^{I \times J \times K}$. Note that the tensor fiber $\underline{\mathbf{X}}(i, j, :) \in \mathbb{R}^K$ represents the power spectrum density (PSD) of the received signal measured at the location (i, j) .

[†] The work of S. Shrestha and X. Fu is supported in part by the NSF-Intel MLWiNS Program CNS-2003082, and in part by NSF ECCS-2024058.

^{*} The work of M. Hong is supported in part by the NSF-Intel MLWiNS Program CNS-2003033, and in part by ARO W911NF-19-1-0247.

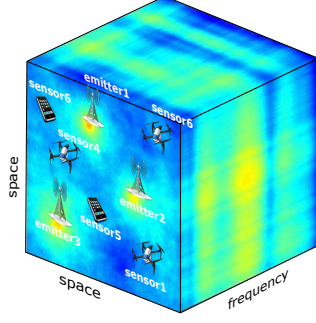


Fig. 1. The SC scenario of interest.

To acquire the spectral information over the region, a number of sensors are placed at certain locations (i_n, j_n) for $n = 1, \dots, N$ where $i_n \in [I]$ and $j_n \in [J]$ and we use the notation $[I] = \{1, \dots, I\}$ (see Fig. 1 for illustration). Assume that there is a set $\Omega \subseteq [I] \times [J]$ such that for every $(i, j) \in \Omega$, there is a spectrum sensing device (a sensor) that can measure the PSD over K frequency bands at the location (i, j) . This means that $\underline{X}(i, j, :) \in \mathbb{R}^K$, $\forall (i, j) \in \Omega$ is available, and we wish to use such information to recover the entire $\underline{X} \in \mathbb{R}^{I \times J \times K}$.

2.1. Classic Model and Methods

In [2, 4, 7, 10], a model for $\underline{X}$ is advocated. To be specific, consider R emitters that are transmitting signals in the region. The power propagation pattern of emitter r is a nonnegative matrix $\mathbf{S}_r \in \mathbb{R}^{I \times J}$, i.e., the SLF of emitter r . Let $\mathbf{c}_r \in \mathbb{R}^K$, column of $\mathbf{C} \in \mathbb{R}^{K \times R}$, be the PSD of emitter r . Then, $\mathbf{S}_r(i, j)c_{k,r}$ represents the interference power level at location (i, j) and frequency k contributed by emitter r ; i.e.,

$$\underline{X}(i, j, k) = \sum_{r=1}^R \mathbf{S}_r(i, j, k) \mathbf{C}(k, r). \quad (1)$$

This model is considered fairly accurate if the bandwidth of interest is not too wide relative to the central carrier frequency (e.g., 20MHz bandwidth at a carrier frequency within 2-5 GHz.) [11, 12]. Recently, [2] modeled $\mathbf{S}_r$ as a low-rank matrix and connected (1) to a block-term tensor decomposition (BTD) model [13], and provided a tensor recovery framework with provable guarantees for estimating $\underline{X}$.

The theoretical guarantees in [2] are appealing. However, the challenge is that the presumed models are often not accurate enough, and such modeling error oftentimes gives rise to (large) performance degradation in practice. For example, the low-rank assumption on $\mathbf{S}_r$ could be easily violated if shadowing is strong, which may be true for city centers or indoor environments (see Fig. 2).

2.2. Deep Generative Model-Based Approaches

Recently, [7, 8] proposed to utilize deep generative models for completing the radio maps from sparse measurements. This line of methods does not impose any signal model on $\underline{X}$. Instead, a mapping from the sampled measurements to the complete data is learned through a deep generative network, e.g., deep autoencoders (DAEs) [7] or generative adversarial networks (GANs) [8]. To implement this idea, L simulated radio maps are generated (denoted by $\underline{X}^{(\ell)}$ for $\ell = 1, \dots, L$), and (randomly) sampled (denoted by $\underline{X}^{(\ell)}(\Omega_\ell, :)$ where $\Omega_\ell \subset [I] \times [J]$ is a random sampling pattern for training example ℓ). Then, the pairs $\{(\underline{X}^{(\ell)}(\Omega_\ell, :), \underline{X}^{(\ell)})\}_{\ell=1}^L$

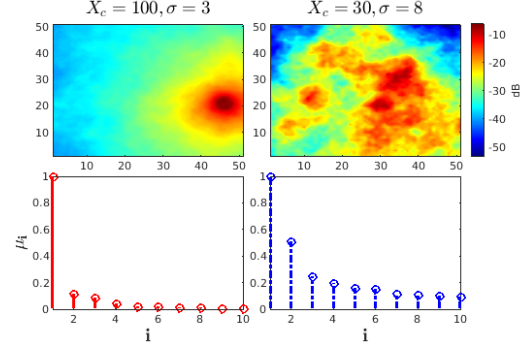


Fig. 2. Shadowing effect affects the soundness of the low-rank assumption on the SLFs. The SLFs under test are from a $50 \times 50\text{m}^2$ region. The maps are generated following (1) and (3)-(4); see [2, 14]. μ_i denotes the i th singular value normalized w.r.t. μ_1 . The singular value decays much slower when the shadowing effect is more severe.

serve as the training pairs for learning a ‘completion network’. The training problem can be summarized as follows

$$\min_{\theta} \frac{1}{L} \sum_{\ell=1}^L \left\| \mathbf{f}_{\theta}(\underline{X}^{(\ell)}(\Omega_\ell, :)) - \underline{X}^{(\ell)} \right\|_F^2 \quad (2)$$

where $\mathbf{f}_{\theta} : \mathbb{R}^{I \times J \times K} \rightarrow \mathbb{R}^{I \times J \times K}$ is the generative network.

The DAE/GAN based approaches can effectively circumvent the modeling challenges as in purely model-based approaches, e.g., [2, 4, 9]. However, other challenges arises. To be specific, since the domain of $\underline{X}$ consists of numerous scenarios with various numbers, locations, and frequencies of emitters, generating a training set that well covers all the possible scenarios is almost impossible. Consequently, no matter how many scenarios are considered in the training set, the test case can easily go beyond these scenarios, since the system designer typically has no control about some key problem parameters, e.g., the number of emitters in the region of interest.

3. PROPOSED APPROACH

In this work, we offer a model-aided deep learning approach, with the aim of reducing the ‘modeling burden’ of the network.

3.1. Main Framework

Our framework can be summarized in the following two steps:

- **Step 1)** From $\underline{X}(\Omega, :)$, estimate $\mathbf{C}$ and $\mathbf{Q}_r = \mathbf{W}_{\Omega} \circledast \mathbf{S}_r$ of emitter r for $r = 1, \dots, R$, where $\circledast$ is the Hadamard product, and $\mathbf{W}_{\Omega}(i, j) = 1$ if $(i, j) \in \Omega$ and $\mathbf{W}_{\Omega}(i, j) = 0$ otherwise.
- **Step 2)** Use a deep generative network to complete $\mathbf{S}_r$ from $\mathbf{Q}_r$ for all r .

In the above, $\mathbf{Q}_r$ represents the SLF of emitter r but only the entries indexed by $(i, j) \in \Omega$ are observed (see Fig. 3). Our motivation is as follows. The modeling challenge of the radio map largely lies in how to model the SLFs, while the aggregation model in (1) is relatively accurate—see [2, 4, 11, 12] for real data validations. Using the well-established aggregation model, one may be able to separate the contributions from different emitters. This way, the SLF completion stage only concerns a single emitter—the state space of this

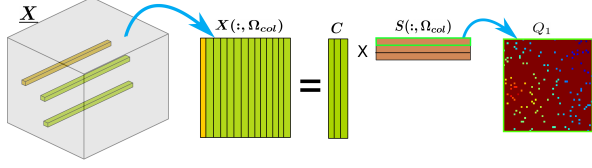


Fig. 3. Illustration for re-arranging the sensed PSDs at various locations (fibers in the left tensor). $S(r, \Omega_{\text{col}})$ is reshaped into Q_r .

learning problem is largely shrunk from that of the approach in [7, 8], which leads to a deep generative model that is much easier to train, compared to those in [7, 8].

3.2. Nonnegativity-Based Disaggregation

To implement our framework, we offer an NMF-based approach for separating the emitters' SLFs. Consider the complete radio map $\underline{\mathbf{X}} \in \mathbb{R}^{I \times J \times K}$. One can re-arrange all the spectrum measurements at all the locations $\underline{\mathbf{X}}(i, j, :) \in \mathbb{R}^K$ as a matrix, i.e.,

$$\mathbf{X} = [\underline{\mathbf{X}}(1, 1, :), \underline{\mathbf{X}}(2, 1, :), \dots, \underline{\mathbf{X}}(I, J, :)] \in \mathbb{R}^{K \times IJ}.$$

By (1), we have $\mathbf{X} = \mathbf{C}\mathbf{S}$, where $\mathbf{S}(r, :) = \text{vec}(\mathbf{S}_r)$. The column index of $\mathbf{X}$ and the location coordinates (i, j) have the following relation: $q = I(i - 1) + j$. Hence, there is a bijective mapping such that $\Omega \subseteq [I] \times [J]$ can be mapped to $\Omega_{\text{col}} \subseteq [IJ]$. Consequently, the observed data can be expressed as follows

$$\widetilde{\mathbf{X}} = \mathbf{X}(:, \Omega_{\text{col}}) = \mathbf{C}\mathbf{S}(:, \Omega_{\text{col}}) = \mathbf{C}\mathbf{H}.$$

The model $\widetilde{\mathbf{X}} = \mathbf{C}\mathbf{H}$ is an NMF model, since $\mathbf{C}$ consists of R PSDs of the emitters as its columns, and the rows of $\mathbf{H}$ are the sampled SLFs of the emitters. By their physical meaning, $\mathbf{C}$ and $\mathbf{H}$ are both nonnegative. Then, we have the following proposition:

Proposition 1 *Under (1), assume that $\mathbf{S}_r$'s are drawn from any joint absolutely continuous distribution, that $\text{rank}(\mathbf{C}) = R$, and that every emitter has a designated frequency f_r such that $\mathbf{C}(f_r, r) > 0$ and $\mathbf{C}(f_r, k) = 0$ for $k \neq r$. Then, with probability one, there exists a polynomial time algorithm that can estimate $\widehat{\mathbf{C}} = \mathbf{C}\Pi\Lambda$, $\widehat{\mathbf{H}} = \Pi^\top\Lambda^{-1}\mathbf{H}$, given that the number of samples satisfies $|\Omega| \geq R$.*

In a nutshell, the designated frequency condition translates to the separability condition in NMF [15, 16]—which is normally satisfied if the emitters sparsely take different subbands; see more discussions in [12]. Then, the proof relies on the fact that the so-called separable NMF admits polynomial time solvers. One can see that only R samples are needed for the disaggregation stage. In this work, we employ the NMF algorithm in [16]. The method is a greedy algorithm that is similar to the Gram-Schmidt procedure, and thus is fast and scalable—which fits the need of wireless systems.

3.3. Deep Generative Model-based SLF Completion

After $\widehat{\mathbf{H}}(r, :)$ is obtained, we obtain an estimate for the ‘masked version’ of $\mathbf{S}_r$, i.e., $\widehat{\mathbf{Q}}_r = \mathbf{W}_\Omega \otimes \widehat{\mathbf{S}}_r$ (see Fig. 3). The remaining task is to recover $\mathbf{S}_r$ from $\mathbf{W}_\Omega \otimes \widehat{\mathbf{S}}_r$, where $\otimes$ denotes the Hadamard product. This is a matrix completion problem—but low-rank completion techniques may not be applicable, as seen in Fig. 2.

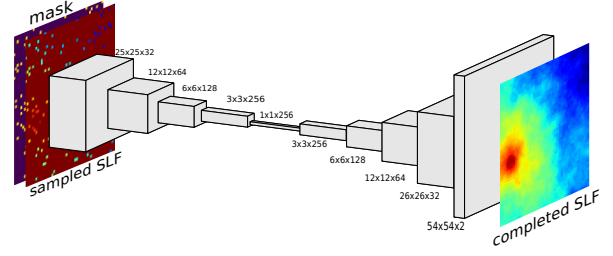


Fig. 4. The autoencoder-based SLF completion network.

In SC, it is reasonable to assume that one has some general knowledge about the region of interest—e.g., if this is an indoor, outdoor, urban, or suburban area. With such information, shadowing and fading simulators can be employed to generate a large number of synthetic SLFs for an individual emitter. For example, one can follow a joint path loss and log-normal shadowing model in [14]:

$$\mathbf{S}_r(\mathbf{y}) = \|\mathbf{y} - \mathbf{z}_r\|_2^{-\alpha_r} 10^{z_r(\mathbf{y})/10} \quad (3)$$

where $\mathbf{y} = (i, j)$ denotes the spatial coordinates, $\mathbf{z}_r$ is the location of r -th emitter, α_r is the path loss coefficient of r , and $z_r(\mathbf{y})$ is the correlated log-normal shadowing component sampled from zero-mean Gaussian with variance σ_r^2 with the autocorrelation between $\mathbf{y}$ and $\mathbf{y}'$ as follows:

$$\mathbb{E}(z_r(\mathbf{y}), z_r(\mathbf{y}')) = \sigma_r^2 \exp(-\|\mathbf{y} - \mathbf{y}'\|_2 / X_c), \quad (4)$$

in which X_c is the so-called decorrelation distance. For example, for an outdoor environment, the typical values of X_c and σ range from 50 to 100 and 4dB to 12dB, respectively [14]. Our simulator-generated training examples include as many as possible positions of the emitter, to cover more cases and scenarios. Let us denote this set of training samples as $\mathbf{M}^{(\ell)} \in \mathbb{R}^{I \times J}$ for $\ell = 1, \dots, L$. Our training is by finding a neural network that minimizes the following loss function:

$$\min_{\theta} \frac{1}{L} \sum_{\ell=1}^L \left\| \mathbf{f}_{\theta}(\mathbf{W}_{\Omega_{\ell}} \otimes \mathbf{M}^{(\ell)}) - \mathbf{M}^{(\ell)} \right\|_{\text{F}}^2, \quad (5)$$

where Ω_{ℓ} is a uniformly and random sampling mask for example ℓ and $\mathbf{f}_{\theta}(\cdot) : \mathbb{R}^{I \times J} \rightarrow \mathbb{R}^{I \times J}$ is a deep neural network. After the network is trained, the completion stage is implemented by the following: $\widehat{\mathbf{S}}_r = \mathbf{f}_{\theta}(\widehat{\mathbf{Q}}_r)$, where θ represents the trained network parameters by tackling the problem in (5). In this work, we set $\mathbf{f}_{\theta}$ to be an autoencoder, but other generative models such as the variational autoencoder and GAN can also be readily employed. We refer to the proposed approach as the NMF-DAE method.

4. NUMERICAL RESULTS

In this section, we use numerical examples to showcase the effectiveness of the proposed approach. We consider an urban outdoor region discretized to 51×51 grids over 64 frequency bins. We follow the shadowing model in [14] [see (3)-(4)] to generate SLFs for 3 to 6 emitters. Their PSDs are generated as in [2]. In our training examples, we randomly vary the standard deviation parameter in (4) (i.e., σ_r) from 2 to 7, decorrelation distance (X_c) from 50 to 100, which is a setting typical for outdoor scenarios [14]. The path-loss

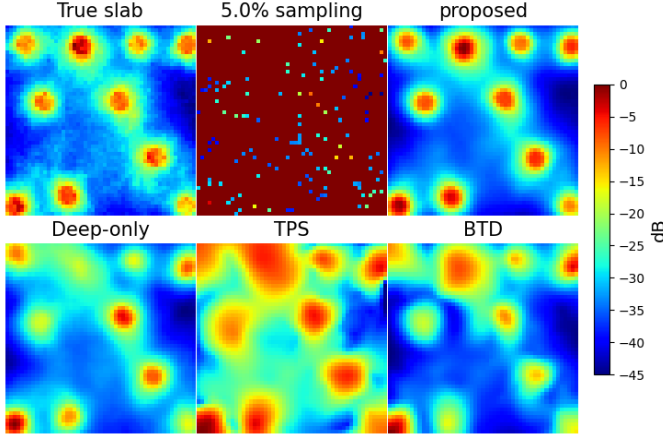


Fig. 5. Ground-truth and reconstructed radio maps by various methods at frequency bin 1; $\rho = 5\%$, $R = 5$, $X_c = 50$, $\sigma = 4$.

exponent α is sampled from the uniform distribution between 2 and 3. For sample ℓ , we generate a uniformly random mask Ω_ℓ such that $\rho = |\Omega_\ell|/(IJ) \times 100\%$ ranges from 1% to 20%. In the prediction stage, we use $\sigma_r = 4$, $X_c = 50$, unless specified otherwise. For NMF-DAE, we assume that R is previously estimated, which can be done by classic model order selection methods, e.g., those in [17]. Note that separable NMF algorithm can also be employed for estimating the model order; see [18, 19].

In terms of network structure, we use an 11-layer convolutional neural network (CNN)-based autoencoder, where layers 1-5 act as the encoder and layers 6-11 the decoder. Our filter size for all the layers is 4×4 , except that layers 5 and 6 use 3×3 filters. The stride size is set to be 2. The number of channels for all the layers are annotated in Fig. 4. This network structure is used for our approach and a baseline that directly completes $\mathbf{X}$ without disaggregation (denoted as ‘Deep-Only’). The baseline is reminiscent of the methods in [7, 8]. Both models are trained with the same hyperparameters and the Adam algorithm (where the learning rate is 0.001 for first 30 epochs and 0.0001 for the last 70 epochs; batch size of Adam is set to be 128) using 500,000 training examples. We also use the *thin plate spline* (TPS) method, which is a classic spatial kernel based interpolation method [20, 21], and the block-term tensor decomposition approach (BTB) in [2] as additional benchmarks.

Fig. 5 shows the ground-truth and recovered radio maps at the first frequency bin. Several observations are in order. First, the proposed NMF-DAE method outputs an estimated map that is visually most consistent with the ground-truth map among all the method under comparison. Second, one can see that the Deep-Only method misses some emitters—i.e., under estimates the power levels at some emitter locations. This can be detrimental to applications such as cognitive radio, where discovering unoccupied frequency bins and un-interfered locations are of great interest. This effect may attribute to the ‘out-of-distribution’ problem that the Deep-Only method encounters: Since the training samples at most contain 6 emitters and it not so often that 5 emitters appear in the same frequency in the training samples. Hence, the Deep-Only method experiences such generalization challenges. Since the proposed method first disaggregates the radio map to the individual emitter level, it does not have such issues. This effect will be quantified later in Fig. 6.

Fig. 7 [Left] and [Right] show the *normalized absolute error* (NAE) (see [2] for definition) under different ρ 's when $R =$

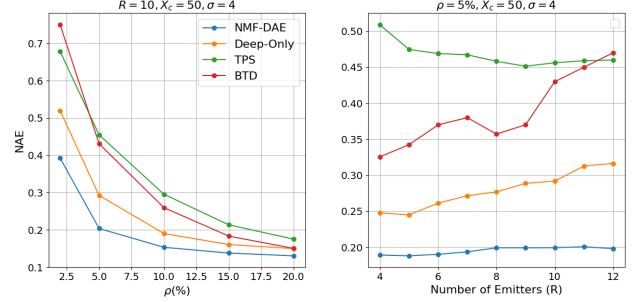


Fig. 6. [Left] Detection probability with respect to sampling size, Ω . [Right] Detection probability with respect to number of emitters.

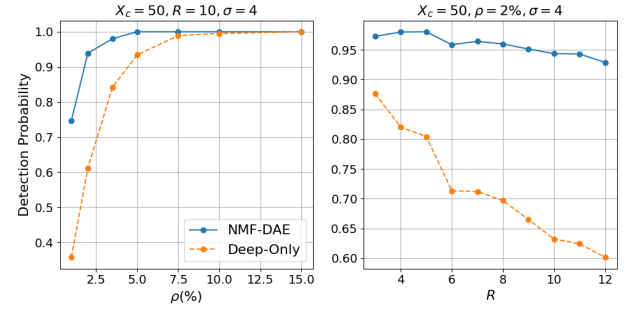


Fig. 7. [Left] Performance evaluation with respect to sampling size. [Right] Performance evaluation with respect to number of emitters.

10, $K = 64$, $X_c = 50$, $\sigma = 4$ and R 's when $\rho = 5\%$, $X_c = 50$, $\sigma = 4$, respectively. In Fig. 7, one can see that Deep-Only is comparable in terms of NAE relative to NMF-DAE when $\rho \geq 17.5\%$, but deteriorates much faster when ρ decreases. In Fig. 7 [Right], we explicitly test the performance of the auto-encoders beyond the range of parameters used in the training examples—by increasing the number of emitters beyond 6. One can see that Deep-Only is less competitive under such settings whereas the proposed method exhibits good robustness, again, since our NMF-disaggregation stage makes the method insensitive to such out-of-range scenarios.

In Fig. 6, we quantify the emitter miss detection problem by showing the detection probability of emitters at reconstruction. If an emitter at a frequency bin is estimated to admit a power level that is decreased by 75% from the original power value in the ground-truth map at the same emitter location, we treat it as miss-detected. Fig. 6 shows the detection probabilities of different methods under different ρ 's and R 's. It is clear from the figure that Deep-Only has a tendency to miss emitters, especially when $\rho < 5\%$. The proposed method keeps a good detection probability even if $\rho \approx 2\%$.

5. CONCLUSION

In this work, we proposed an NMF-aided learning framework for DAE-based blind spectrum cartography. With the assistance of a model-based signal disaggregation stage, the training and generalization of the deep learning-based SC stage exhibits robust and promising performance, outperforming purely model-based and purely data-driven approaches. Our results suggest that classic signal processing models may continue playing critical roles in deep learning-based system design, with proper design.

6. REFERENCES

- [1] S. Bi, J. Lyu, Z. Ding, and R. Zhang, "Engineering radio maps for wireless resource management," *IEEE Trans. Wireless Commun.*, vol. 26, no. 2, pp. 133–141, 2019.
- [2] G. Zhang, X. Fu, J. Wang, X.-L. Zhao, and M. Hong, "Spectrum cartography via coupled block-term tensor decomposition," *IEEE Trans. Signal Process.*, 2020.
- [3] J. A. Bazerque and G. B. Giannakis, "Distributed spectrum sensing for cognitive radio networks by exploiting sparsity," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1847–1862, 2010.
- [4] J. A. Bazerque, G. Mateos, and G. B. Giannakis, "Group-lasso on splines for spectrum cartography," *IEEE Trans. Signal Process.*, vol. 59, no. 10, pp. 4648–4663, 2011.
- [5] J. A. Bazerque and G. B. Giannakis, "Distributed spectrum sensing for cognitive radio networks by exploiting sparsity," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1847–1862, 2009.
- [6] D. Romero, S.-J. Kim, G. B. Giannakis, and R. López-Valcarce, "Learning power spectrum maps from quantized power measurements," *IEEE Trans. Signal Process.*, vol. 65, no. 10, pp. 2547–2560, 2017.
- [7] Y. Teganya and D. Romero, "Data-driven spectrum cartography via deep completion autoencoders," in *ICC 2020-2020 IEEE International Conference on Communications (ICC)*. IEEE, 2020, pp. 1–7.
- [8] X. Han, L. Xue, F. Shao, and Y. Xu, "A power spectrum maps estimation algorithm based on generative adversarial networks for underlay cognitive radio networks," *Sensors*, vol. 20, no. 1, p. 311, 2020.
- [9] D. Romero, S.-J. Kim, G. B. Giannakis, and R. López-Valcarce, "Learning power spectrum maps from quantized power measurements," *IEEE Trans. Signal Process.*, vol. 65, no. 10, pp. 2547–2560, 2017.
- [10] D. Lee, S.-J. Kim, and G. B. Giannakis, "Channel gain cartography for cognitive radios leveraging low rank and sparsity," *IEEE Trans. Wireless Commun.*, vol. 16, no. 9, pp. 5953–5966, 2017.
- [11] X. Fu, N. D. Sidiropoulos, J. H. Tranter, and W.-K. Ma, "A factor analysis framework for power spectra separation and multiple emitter localization," *IEEE Trans. Signal Process.*, vol. 63, no. 24, pp. 6581–6594, 2015.
- [12] X. Fu, N. D. Sidiropoulos, and W.-K. Ma, "Power spectra separation via structured matrix factorization," *IEEE Trans. Signal Process.*, vol. 64, no. 17, pp. 4592–4605, 2016.
- [13] L. De Lathauwer, "Decompositions of a higher-order tensor in block terms—Part II: Definitions and uniqueness," *SIAM J. Matrix Anal. Appl.*, vol. 30, no. 3, pp. 1033–1066, 2008.
- [14] A. Goldsmith, *Wireless communications*. Cambridge university press, 2005.
- [15] X. Fu, K. Huang, N. D. Sidiropoulos, and W.-K. Ma, "Nonnegative matrix factorization for signal and data analytics: Identifiability, algorithms, and applications," *IEEE Signal Process. Mag.*, vol. 36, no. 2, pp. 59–80, 2019.
- [16] N. Gillis and S. A. Vavasis, "Fast and robust recursive algorithms for separable nonnegative matrix factorization," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 4, pp. 698–714, 2013.
- [17] P. Stoica, R. L. Moses *et al.*, "Spectral analysis of signals," 2005.
- [18] X. Fu, W.-K. Ma, T.-H. Chan, and J. M. Bioucas-Dias, "Self-dictionary sparse regression for hyperspectral unmixing: Greedy pursuit and pure pixel search are related," *IEEE J. Sel. Topics Signal Process.*, vol. 9, no. 6, pp. 1128–1141, 2015.
- [19] X. Fu and W.-K. Ma, "Robustness analysis of structured matrix factorization via self-dictionary mixed-norm optimization," *IEEE Signal Process. Lett.*, vol. 23, no. 1, pp. 60–64, 2016.
- [20] S. Üreten, A. Yongacoglu, and E. Petriu, "A comparison of interference cartography generation techniques in cognitive radio networks," in *2012 IEEE International Conference on Communications (ICC)*, 2012, pp. 1879–1883.
- [21] F. L. Bookstein, "Principal warps: Thin-plate splines and the decomposition of deformations," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, no. 6, pp. 567–585, 1989.