Deep Learning based Channel Extrapolation for Large-Scale Antenna Systems: Opportunities, Challenges and Solutions

Shun Zhang, Yushan Liu, Feifei Gao, Chengwen Xing, Jianping An, and Octavia A. Dobre

Abstract

With the depletion of spectrum, wireless communication systems turn to exploit large antenna arrays to achieve the degree of freedom in the space domain, such as millimeter wave massive multi-input multi-output (MIMO), reconfigurable intelligent surface assisted communications and cell-free massive MIMO. In these systems, how to acquire accurate channel state information is difficult and becomes a bottleneck of the communication links. In this article, we introduce the concept of channel extrapolation that relies on a small portion of channel parameters to infer the remaining channel parameters. Since the substance of channel extrapolation is a mapping from one parameter subspace to another, we can resort to deep learning (DL), a powerful learning architecture, to approximate such a mapping function. Specifically, we first analyze the requirements, conditions and challenges for channel extrapolation. Then, we present three typical extrapolations over the antenna dimension, the frequency dimension, and the physical terminal, respectively. We also illustrate their respective principles, design challenges and DL strategies. It will be seen that channel extrapolation could greatly reduce the transmission overhead and subsequently enhance the performance gains compared with the traditional strategies. In the end, we provide several potential research directions on channel extrapolation for future intelligent communication systems.

Introduction

With the rapid development of wireless technologies, we are entering the era of the fifth generation (5G) wireless communications and are heading toward the sixth generation (6G) [1, 2]. Massive multi-input multi-output (MIMO) will continue to serve as one of the key technologies for 6G, as it can provide much more degrees of freedom than conventional MIMO. In particular, massive MIMO combined with millimeter wave (mmWave) communications can achieve orders of magnitude enhancement in system throughput. For cell-free massive MIMO, distributed systems can potentially reduce the inter-cell interference through coherent cooperation between base stations (BSs) and provide higher coverage than co-located massive MIMO. The reconfigurable intelligent surface (RIS) with a massive number of passive antennas can effectively perform the desired beamforming and reconstruct the radio scattering environment into an intelligent environment. Unfortunately, as the number of antennas increases, all these promising wireless technologies require a large amount of training overhead in order to achieve accurate channel state information (CSI).

Traditionally, channel reciprocity is utilized in time division duplex (TDD) massive MIMO systems to reduce the cost of downlink channel estimation as long as the uplink CSI is obtained. For frequency division duplex (FDD) massive MIMO, the compressive algorithms that exploit the sparsity of channel in angle domain have been used to estimate CSI. However, the performance of compressive sensing is restricted with the linear sparsity assumption that does not accurately match the complex nonlinear sparse characteristics of the environment. Recently, Alkhateeb *et al*. revealed the deterministic relationship among different channels at different antennas and frequency bands, and introduced a new concept, namely channel mapping in the space and frequency domains [3]. With a similar principle, Esswie *et al*. proposed a spatial channel estimation scheme to reconstruct the downlink channel from the uplink channel measurements for FDD MIMO systems [4]. In [5], Ali *et al.* constructed the channel covariance of the mmWave link from the spatial characteristics of the sub-6 GHz band.

In fact, the essential principle of deterministic mapping among different channels is that users at different spaces or different frequencies experience the same electromagnetic environment. Utilizing such deterministic mapping, the CSI at one space/frequency point can be used to predict the CSI at another space/frequency point, which is referred to as *channel extrapolation*. Clearly, channel extrapolation can be excitingly useful to reduce the training cost over massive MIMO related systems. Inspired by the universal approximation capability of deep learning (DL) [6], it is possible to use DL to characterize the mapping among channels at different space/frequency locations.

In this article, we introduce the mechanism for channel extrapolation and analyze its major challenges over three scenarios: antenna domain extrapolation, frequency domain extrapolation

Digital Object Identifier: 10.1109/MWC.001.2000534

Shun Zhang and Yushan Liu are with Xidian University; Feifei Gao (corresponding author) is with Tsinghua University; Chengwen Xing and Jianping An are with Beijing Institute of Technology; Octavia A. Dobre is with Memorial University.

and physical terminal extrapolation. Specifically, for antenna extrapolation, we propose to use the channel of a few uplink antennas to reconstruct the full downlink channel for massive MIMO and RIS-aided communication, and present the corresponding DL-based solution. For frequency extrapolation, we consider two cases: channel mapping between two subcarrier sets within one frequency band, and channel mapping across different frequency bands. For terminal extrapolation, we consider channel mapping among different distributed antennas/users, where the sensors are applied to provide additional information about the environment and DL is used to achieve the transfer between neural networks (NNs) of distinct antennas/users. Finally, we discuss several potential research directions on channel extrapolation for future intelligent communications.

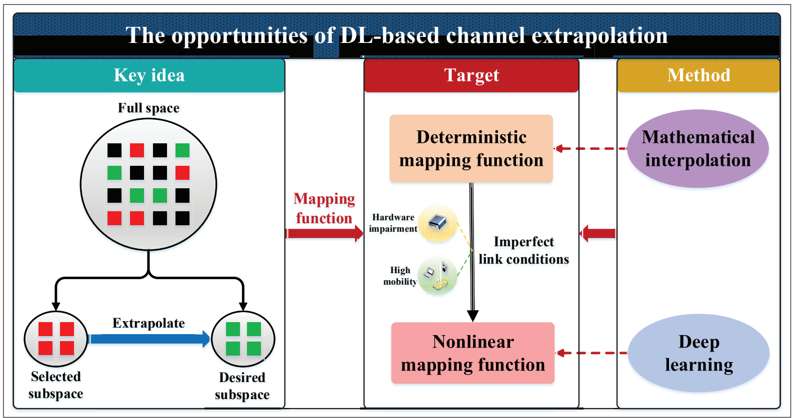

FIGURE 1. The opportunities of DL-based channel extrapolation.

Problem Formulation

Concept of Channel Extrapolation

From the propagation characteristics of the electromagnetic wave, we know that when the user switches the frequency or moves in a short time, its surrounding environment would not change much, which makes the channels at different spaces or different frequencies to have a certain mapping relationship. Many efforts have been made to explore such mapping, and some effective mapping models, for example, the one between the uplink and downlink channels over massive MIMO systems, have been established. Nevertheless, channel mappings, like the one between the sub-6 and mmWave bands or the one among different distributed antennas, are difficult to describe mathematically. In [3], the authors proved that large-scale antenna arrays are able to assure the *bijectiveness* condition, which makes the channel mapping along the frequency and space dimensions unique. Then, the authors utilized the DL approach to successively approximate such channel mapping. The key idea of DL-based channel extrapolation is illustrated in Fig. 1.

Challenges of DL-Based Channel Extrapolation

In fact, channel extrapolation can be treated as a mapping from one parameter subspace to another, which depends on three aspects: the acquisition of the original subspace information, the choice of the original subspace, and the mapping scheme from the original subspace to the targeted one. Notice that the task of "the original subspace acquisition" is to estimate the sub-sampled channel of small size and can be easily implemented through the conventional Bayesian linear estimator, message passing algorithms, or DL-based algorithms. Therefore, in the following, we will focus on the other two challenges.

Subspace Selection: For the given full space, the raw input information is closely related with the selection of subspace. Different selections would correspond to distinguished information compression rates and *selection patterns*, where the selected space is marked as "1" while the others are marked as "0." With the power, hardware, or performance constraints, it is possible to optimize the selection pattern and to provide a good starting point for a specific extrapolation task. The sparser the selection pattern is, the less the training costs and hardware power consumption are, but the poorer the performance of the channel extrapolation will be. Thus, with DL-based channel extrapolation, it is important to determine the selection pattern before starting the transmission. In fact, the core operation of DL is the gradient descent algorithm that requires the mapping function to be continuous and differentiable. However, it can be checked that the subspace selection is a discrete combinatorial optimization problem, for which the gradient descent algorithm is not easy to implement. Hence, the optimization of the subspace selection is challenging for DL-based channel extrapolation.

Mapping Scheme: DL-based channel extrapolation is similar to super-resolution in the field of image processing, in which it is important to properly exploit the correlation between data elements for information completion. In order to improve the performance of the NN, we can increase the number of data layers or modify the NN structure. However, more layers result in heavier calculation, and when the number of layers reaches a certain number, the degree of improvement becomes smaller. Sometimes, excessively deepening the NN will even cause gradient explosion and disappearance. Thus, we aim to construct a robust but simple NN structure to achieve better extrapolation performance and faster convergence, which is another challenge of the DL-based channel extrapolation.

In the following, we present channel extrapolation over the antenna dimension, the frequency dimension and the physical terminal separately, and show the effectiveness of each type of extrapolation.

Channel Extrapolation over Antenna Dimension

A practical way to implement large antenna arrays, for example, massive MIMO, is to use the hybrid analog and digital architecture, where a small number of radio frequency (RF) chains are connected to massive antennas through the full connection or subarray connection structure. In this case, the limited number of RF chains have to connect the active antennas in turn for the acquisition of downlink CSI, which consumes significant time resources, especially when physically switching the antenna costs non-ignorable time. Over the RIS-aided communication network, the

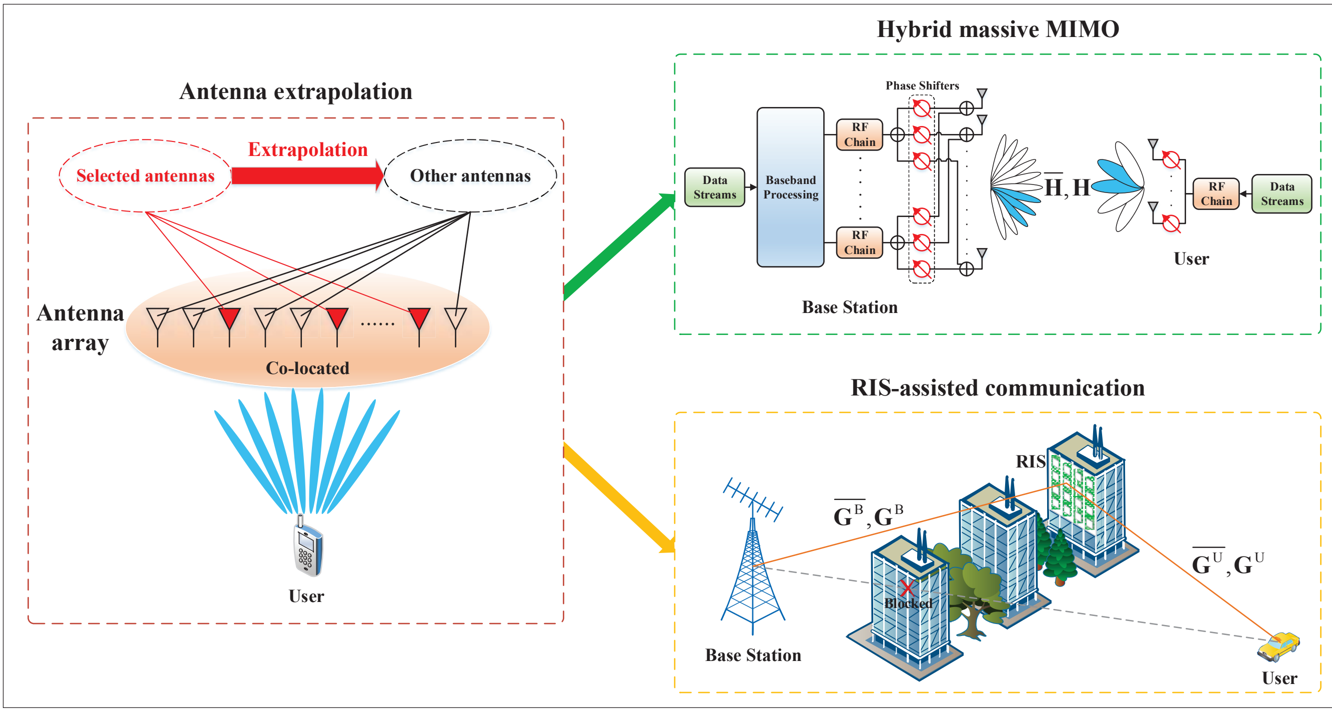

FIGURE 2. The channel extrapolation over antenna dimension for large-scale antenna systems. Note that $\overline{\mathbf{H}}$ and $\mathbf{H}$ denote the partial and full channels between the BS and the RIS for the hybrid massive MIMO respectively, while the same goes for $\overline{\mathbf{G}}^{B}$, $\mathbf{G}^{B}$, $\overline{\mathbf{G}}^{U}$ and $\mathbf{G}^{U}$.

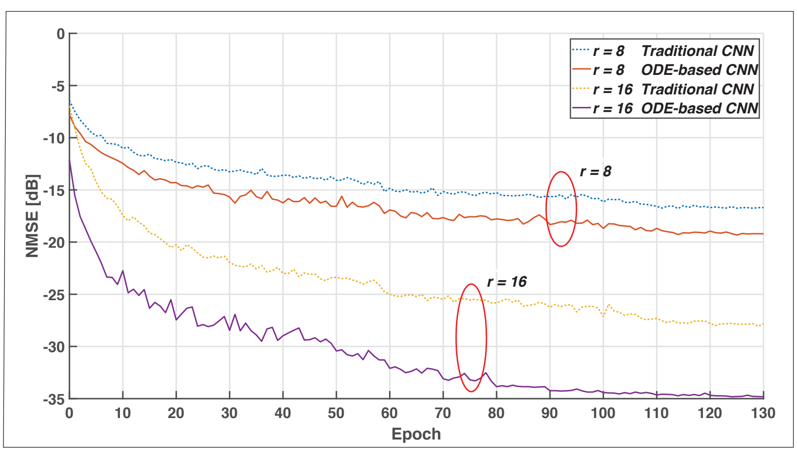

FIGURE 3. The NMSE comparison of traditional CNN and ODE-based CNN against epochs for antenna extrapolation in a RIS-assisted communication system. Note that *r* denotes the number of active antennas at RIS, the "Epoch" denotes the number of times that all training sets have been traversed for neural network training and the definition of NMSE can refer to [9].

channel size is in scale with the number of RIS elements, which is usually very large to achieve good electromagnetic reconfiguring performance. Thus, the channel estimation in a RIS-aided network also requires significant time resources. With the new idea of channel extrapolation, we can explore the mapping relation between the partial antennas and the full one in an offline manner, and accurately infer the CSI of full antennas from the partial antennas during online transmission. This type of channel extrapolation architecture is specifically referred to as antenna extrapolation and is illustrated in Fig. 2.

Recently, there are some preliminary results on implementing the antenna extrapolation. In [7], the authors examined antenna extrapolation for a massive MIMO system based on the deep NN (DNN), which can capture the inherent relationship between the uplink and downlink channel data sets and extrapolate the downlink channels from a subset of the uplink CSI. In [8], Zhang *et al*. realized antenna extrapolation through a convolutional NN (CNN) in a RIS-aided communication system. Moreover, the authors proposed an antenna selection network that utilizes the probabilistic sampling theory to select the optimal locations of those active antennas. In [9], we modify the NN structure by ordinary differential equation (ODE) which can describe the latent relations among different data layers and improve the performance gains of antenna extrapolation.

In Fig. 3, we provide evaluation results of antenna extrapolation for the traditional CNN structure and the ODE-based CNN structure in a RIS-aided communication system. The parameter settings are: a 4×4 uniform planar array (UPA) at BS, a 8×8 UPA at RIS and single antenna at the user, where the antenna spacing is half wavelength. The carrier frequency is 2.4 GHz, the system bandwidth is 20 MHz, and the number of subcarriers is 64. Moreover, the number of active antennas at RIS is separately taken as 8 and 16, and the uniform selection pattern is adopted. Notice that the "active antennas" denotes the selected antennas for achieving the initial information of extrapolation. The distribution of training and test users in the DeepMIMO dataset is the same as in [7]. It can be seen from Fig. 3 that with the increase of iteration time, the normalized mean square error (NMSE) curves decay fast first, and then slow down after 110 epochs. As the number of active antennas increases, the NMSEs of the two structures decrease and the performance gap between the two structures becomes larger. Hence, for a given number of active antennas, the ODE-based CNN outperforms the traditional CNN and trains faster than the traditional CNN.

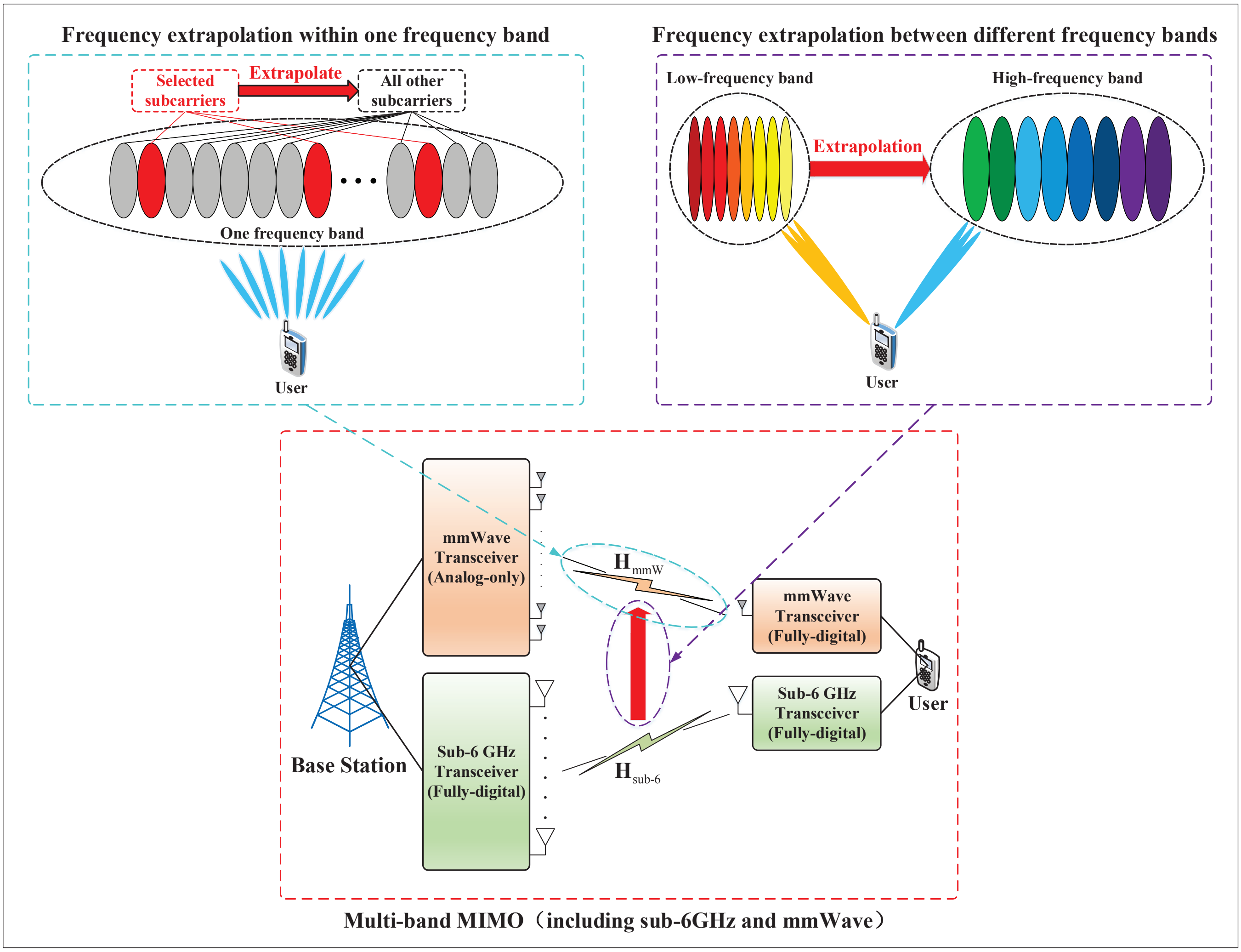

FIGURE 4. The channel extrapolation over frequency for multi-band MIMO systems, including sub-6 GHz and mmWave. Note that H_{mmW} and H_{sub-6} respectively denote the mmWave and sub-6 GHz channels between the BS and the user.

In fact, the DL-based algorithms should adapt to the environmental changes and customize the antenna extrapolation schemes according to the environmental information. Hence, once the environment changes, say when the users are moving, the corresponding NNs should be re-trained. Thus, how to achieve a good balance between the performance gains and the cost caused from retraining the NNs is another challenge for antenna extrapolation. A possible solution could be the transfer and meta learning that has the ability to adapt to the new environment quickly with a small amount of new training samples [10].

Channel Extrapolation over Frequency Dimension

Along the frequency dimension, channel extrapolation can use one set of subcarriers to infer another set of subcarriers, which is named *frequency extrapolation*. We present two typical applications for frequency extrapolation in Fig. 4. One is to implement channel extrapolation between two subcarrier sets within a given frequency band. The other is to extrapolate channels among different frequency bands, being suitable for the multi-band systems, say FDD massive MIMO systems. As seen later, the frequency extrapolation principle may even be feasible when the gap between different frequency bands is large.

Frequency Extrapolation within a Frequency Band

Orthogonal frequency division multiplexing (OFDM) technology with a large amount of subcarriers is usually adopted to capture the potential efficiency of large bandwidth and deal with the frequency selective fading. Since mathematical channel modeling among different subcarriers is available, one can sparsely place pilots at partial subcarriers, sound the incomplete channels and reconstruct the entire CSI over the full band via a signal processing approach. Although the performance of the signal processing approach is very good in a stable environment, its effectiveness degrades in more complicated scenarios. For example, in high mobility scenarios, OFDM may encounter significant inter-carrier interference due to the Doppler spread. Moreover, when the length of the predefined cyclic prefix (CP) is smaller than that of the channel's delay spread, inter-symbol interference appears and destroys the signal modeling in the frequency domain. In addition, the hardware impairments, such as the I/Q imbalance, phase noise and nonlinearity of the power amplifier, would cause nonlinear distortion in received signals.

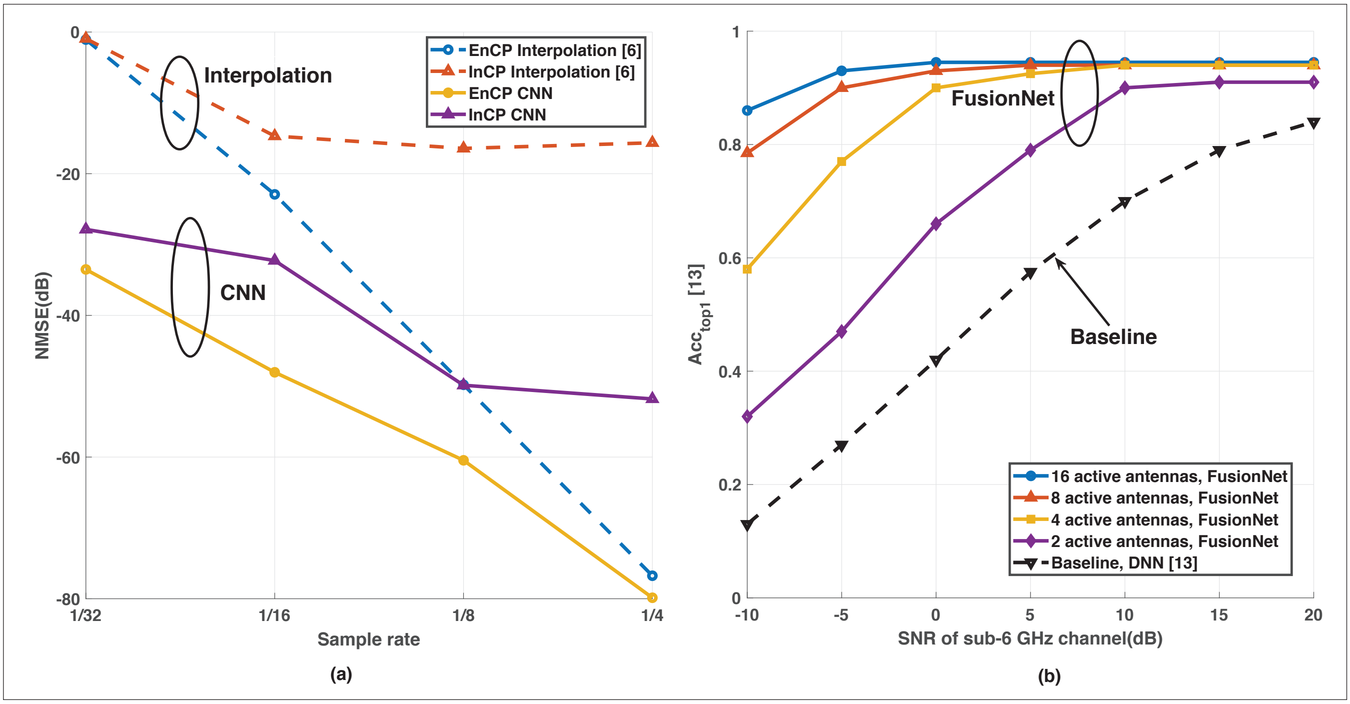

FIGURE 5. The performance analysis of two frequency extrapolations in a multi-band system (including sub-6 GHz and mmWave bands): a) frequency extrapolation within one frequency band; b) frequency extrapolation between sub-6 GHz and mmWave bands.

A possible solution is to utilize the universal approximation capability of DL to characterize the channel modeling in these complicated scenarios and continue to reconstruct the channels of all subcarriers from a small subset of subcarriers. Another key problem is that the traditionally adopted uniform pilot pattern cannot guarantee the optimal performance anymore and the design of the pilot pattern should be carefully addressed.

In Fig. 5a, we offer the evaluation results of frequency extrapolation in the mmWave band based on the CNN and conventional interpolation scheme [6]. The antenna configuration is the same as that in Fig. 3. The carrier frequency is 60 GHz, and the subcarrier spacing is 60 kHz. The number of subcarriers is 1024, and the number of selected subcarriers for partial channel estimation is 256. Notice that the uniform selection pattern is adopted. The curves labeled by 'InCP' correspond to the case of insufficient CP, while the ones marked by 'EnCP' represent the case of enough CP. It can be seen from Fig. 5a that the NMSE curves decay with the increase of the sample rate on the subcarrier. Meanwhile, the NMSEs of CNN are always better than those of conventional interpolation, especially in the case of insufficient CP. Moreover, in the case of enough CP, the gap between CNN and conventional interpolation becomes larger as the sample rate decreases.

Frequency Extrapolation Between Different Frequency Bands

DL-based frequency extrapolation can also be used to infer the downlink channel from the uplink channel in FDD massive MIMO, which solves the problem resulting from the fact that the channel is not reciprocal in the two different frequency bands. In fact, the technology of multiple bands transmission is a new trend, especially that recent communication systems are designed to work at both sub-6 GHz and mmWave bands simultaneously [5, 11, 12]. In the dual band systems, the mmWave spectrum would provide a high speed link and offer gigabit-per-second data rates, but it would also face huge training overhead and high sensitivity to blockages. Since the mmWave and sub-6 GHz links experience the same scattering environment, it is possible to build a deterministic relationship between their channels and extrapolate the CSI of the mmWave band from that of the sub-6 GHz band.

In [11], Li *et al*. used the sub-6 GHz spatial information to help beam selection in a mmWave system by exploiting the feature that the support set of the mmWave channels is a subset of that for the sub-6 GHz channels with the same grid quantization. In [12], Alrabeiah *et al.* performed beam prediction at the mmWave band from the CSI at the sub-6 GHz band with the aid of a DNN. Moreover, the authors incorporated the materials' dielectric coefficients at the sub-6 GHz and mmWave bands, which is critical information for taccurate mapping between the sub-6 GHz and mmWave signals.

Although the effectiveness of the above work has been verified, the frequency extrapolation from sub-6 GHz to mmWave is practically inaccurate. This is mainly because the mmWave signal propagation is highly sensitive to blockages, and the electromagnetic microwave impinging on the arrays of two links have different angular spread. Therefore, the mmWave channel is a sub-category of the sub-6 GHz channel, and the CSI of mmWave band could seriously deviate from that in sub-6 GHz. In order to calibrate the CSI deviation, we design a simple but tricky dual-input NN, referred to as FusionNet, to merge the features of sub-6 GHz channels and a few inherent pilots at the mmWave band to improve the beam prediction [13]. It is worth mentioning that the number of mmWave pilots is not enough to meet the

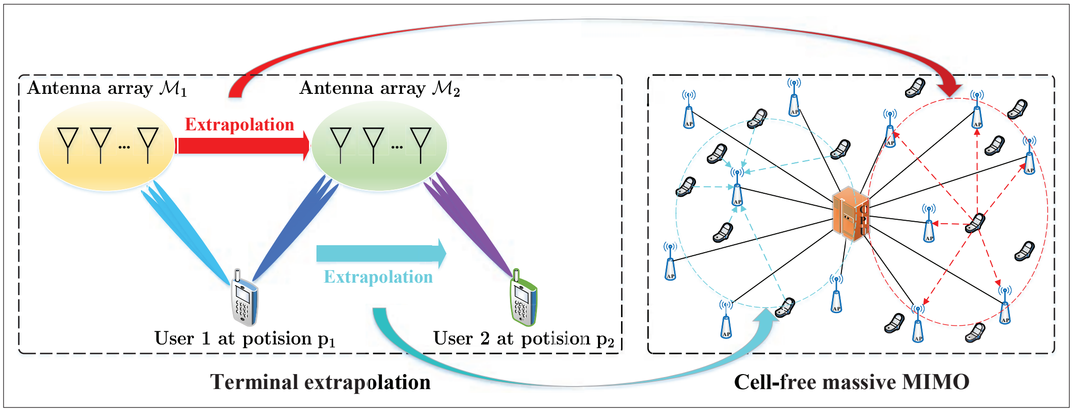

FIGURE 6. The channel extrapolation over the physical terminal for a cell-free massive MIMO system. Note that the terminal can be antenna array and user.

channel estimation requirements; however, it can help refine the mmWave beam direction on top of the CSI from the sub-6 GHz band. Moreover, the balance between the auxiliary pilot overhead and the extrapolation performance can be further optimized, and the pilot pattern along the frequency dimension can be improved through the probabilistic sampling theory.

In Fig. 5b, we display the evaluation results of mmWave beam prediction based on DNN and FusionNet, respectively. The parameter configuration and the definition of top-1 accuracy Acc_{top1} are the same as in [13]. The "Baseline" curve is the performance of the DNN, and the "active antennas" represent the mmWave antennas that participate in the mmWave channel estimation. The number of active antennas in the mmWave system is taken as 2, 4, 8, 16 respectively, and the mmWave pilot signal-to-noise ratio (SNR) is 20 dB. Notice that the uniform pilot pattern is adopted. It can be seen from the Fig. 5b that the prediction accuracy of the FusionNet with any number of mmWave active antennas is always better than that of the baseline method, especially at low SNR for the sub-6 GHz channel estimation. Moreover, as the number of active mmWave antennas increases, the beam prediction accuracy of FusionNet improves significantly but slows down after the number of active mmWave antennas exceeds 8.

Channel Extrapolation over Physical Terminal

Inspired by channel-to-channel mapping, the authors of [3] verified that the channel extrapolation among different terminals at any position is possible, and then applied this mapping concept to the distributed (cell-free) massive MIMO. With the key idea of [3], one can employ a subset of terminals to infer the information for the other terminals, which is named *terminal extrapolation*, illustrated in Fig. 6. If the terminals are close or in a similar environment, then terminal extrapolation is feasible and provides enough accuracy. However, with the increase of the distance between terminals, the number of common scattering objects becomes smaller and the latent relation becomes weaker, which deteriorates the performance of terminal extrapolation. In addition, this mapping would be closely related with the link conditions, such as the physical position of terminals and hardware states, which causes difficulties for terminal extrapolation. Hence, unlike the antenna and frequency extrapolations, terminal extrapolation occurs among different parameter subspaces with large differences, and it is difficult to realize satisfactory extrapolation performance only through NNs.

Meanwhile, with the application of various types of sensors, the communication systems can avail of diverse information, such as users' positions and mobility states. Hence, one can incorporate the information from sensors to enhance the performance of terminal extrapolation [14]. Explicitly, the terminals are equipped with sensors to detect the map of the static environment. Then, they can be divided into groups of terminals that share similar electromagnetic scattering environments [15]. Similar to frequency extrapolation, we can resort to FusionNet to merge a few pilots and sensing information to achieve a good extrapolation among terminals in one group. Nevertheless, the terminals' channels in different groups are less correlated, and thus, these groups cannot share the common extrapolation NN. Separately designing and training the extrapolation NNs for different groups would yield a large computation burden.

Inspired by humans' ability to transfer knowledge from previous experience, transfer learning has become a promising technology in the field of machine learning to solve similar tasks with limited labeled data. This aims to improve the performance of target tasks by exploiting the knowledge from source tasks. Recently, Yang *et al.* utilized a meta-learning scheme and effectively adapted the CSI estimation NN to the new environment [10], which provides a feasible way to apply transfer learning or meta-learning to the terminal extrapolation. Given the constraints on the performance of the terminal extrapolation, one should also determine which group should be selected to infer the NNs of the other groups. Moreover, the size of the group should be carefully designed to balance the performance of terminal extrapolation and training cost.

Future Research Directions

Although channel extrapolation has been established in theory and some preliminary results have been presented, there are still many open problems that need to be investigated over large-scale antenna systems. We highlight several potential research directions as follows.

Joint Channel Extrapolation over Antenna, Frequency and Physical Terminal: The previously

Similar to frequency extrapolation, we can resort to FusionNet to merge a few pilots and sensing information to achieve a good extrapolation among terminals in one group. Nevertheless, the terminals' channels in different groups are less correlated, and thus, these groups cannot share the common extrapolation NN. Separately designing and training the extrapolation NNs for different groups would yield a large computation burden.

Channel extrapolation is a very promising and powerful tool to handle complicated transmission scenarios when the mathematical modeling of the channel is inaccurate or even unavailable. This is especially applicable to mmWave massive MIMO, RIS-assisted communications, and cell-free massive MIMO, and deserves a full investigation and exploitation for such communication systems, which will incorporate intelligence in the future.

introduced three different channel extrapolations can be designed in a joint manner. For example, one can apply the OFDM system with large bandwidth into RIS-assisted communications, that is, the combination of antenna and frequency extrapolation. In such a case, one needs to implement 2D sampling along the antenna-frequency domain, and then use the small 2D subspace to construct the full 2D space over the antenna and frequency dimensions.

Optimization of NN and Resource Deployment: The existing works mainly demonstrated the effectiveness of DL-based channel extrapolations, while further effort is needed to optimize the parameters of the NN, for example, the number of layers, as well as to consider optimal resource deployment, for example, the location of auxiliary mmWave pilots over frequency extrapolation, and distributed antennas/users scheduling over terminal extrapolation.

Model-Driven Channel Extrapolation: The previously discussed channel extrapolation mainly works in a data-driven manner. If the link conditions are good, the mapping function among different subspaces may be modeled mathematically, and a model-driven approach may be designed to enhance the performance of channel extrapolation. We may also design schemes that could intelligently switch between the data-driven and model-driven manners according to the complexity requirement and the dynamic environment.

Vision-Based Channel Extrapolation: In general, the communication environment information, such as building location, obstacle shape and material, dynamic pedestrian and vehicle distribution, would determine the propagation of the electromagnetic waves at any location and any frequency band. Hence, one may resort to sensors such as camera or radar to capture the 3D scene image of the environment and then design the corresponding DL algorithm to assist all three types of channel extrapolation.

Conclusion

In this article, we have presented the channel extrapolation concept and analyzed its three major challenges, that is, the acquisition of the original subspace information, the selection of the original subspace, and the mapping scheme from the original subspace to the targeted one. We divided channel extrapolation into three typical types: antenna extrapolation, frequency extrapolation and terminal extrapolation. For antenna extrapolation, we found that the ODE-based NN outperforms the traditional NN model. For frequency extrapolation, when the gap between different frequency bands is large, we need a small number of pilots to refine the extrapolation. For terminal extrapolation, due to subspaces with large differences, it is difficult to achieve a good performance only through NNs. Hence, the utilization of sensory data and transfer learning can improve extrapolation performance. Finally, we have introduced several potential research directions on channel extrapolation.

In summary, channel extrapolation is a very promising and powerful tool to handle complicated transmission scenarios when the mathematical modeling of the channel is inaccurate or even unavailable. This is especially applicable to mmWave massive MIMO, RIS-assisted communications, and cell-free massive MIMO, and deserves a full investigation and exploitation for such communication systems, which will incorporate intelligence in the future.

Acknowledgment

The work of Feifei Gao was supported by the National Key Research and Development Program of China (2018AAA0102401). The work of Shun Zhang and Yushan Liu was supported by the National Natural Science Foundation of China under Grant 61871455, Grant 61931017, and Grant 61901329. The work of Octavia A. Dobre was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) through its Discovery program.

References

[1] Y. Zhou *et al.*, "Fog Computing Enabled Future Mobile Communication Networks: A Convergence of Communication and Computing," *IEEE Commun. Mag.*, vol. 57, no. 5, May 2019, pp. 20–27.

[2] Y. Zhou *et al.*, "Service-Aware 6G: An Intelligent and Open Network Based on the Convergence of Communication, Computing and Caching," *Digit. Commun. Netw.*, vol. 6, no. 3, Aug. 2020, pp. 253–60.

[3] M. Alrabeiah and A. Alkhateeb, "Deep Learning for TDD and FDD Massive MIMO: Mapping Channels in Space and Frequency," *Proc. IEEE Asilomar Conf. Signals, Systems, and Computers*, 2019, pp. 1465–70.

[4] A. A. Esswie *et al.*, "Spatial Channel Estimation-Based FDD-MIMO Interference Alignment Systems," *IEEE Wireless Commun. Lett.*, vol. 6, no. 2, Apr. 2017, pp. 254–57.

[5] A. Ali *et al.*, "Spatial Covariance Estimation for Millimeter Wave Hybrid Systems Using Out-of-Band Information," *IEEE Trans. Wireless Commun.*, vol. 18, no. 12, Dec. 2019, pp. 5471–85.

[6] X. Yi and C. Zhong, "Deep Learning for Joint Channel Estimation and Signal Detection in OFDM Systems," *IEEE Commun. Lett.*, vol. 24, no. 12, Dec. 2020, pp. 2780–84.

[7] Y. Yang *et al.*, "Deep Learning Based Antenna Selection for Channel Extrapolation in FDD Massive MIMO," *Proc. Int'l. Conf. Wireless Commun. Signal Process.*, Oct. 2020, pp. 182–87.

[8] S. Zhang *et al.*, "Deep Learning Optimized Sparse Antenna Activation for Reconfigurable Intelligent Surface Assisted Communication," arXiv:2009.01607, 2020; available: https://arxiv.org/abs/2009.01607.

[9] M. Xu *et al.*, "Ordinary Differential Equation-Based CNN for Channel Extrapolation over RIS-Assisted Communication," *IEEE Commun. Lett.*, 2021, pp. 1–1.

[10] Y. Yang *et al.*, "Deep Transfer Learning-Based Downlink Channel Prediction for FDD Massive MIMO Systems," *IEEE Trans. Commun.*, vol. 68, no. 12, Dec. 2020, pp. 7485–97.

[11] B. Li, Z. Zhang, and Y. Chen, "Beam Selection in Multi-User Millimeter Wave System Based on Out-of-Band Spatial Information," *Proc. IEEE Int'l. Conf. Comput. Commun. (ICCC)*, 2019, pp. 871–76.

[12] M. Alrabeiah and A. Alkhateeb, "Deep Learning for mmWave Beam and Blockage Prediction Using Sub-6 GHz Channels," *IEEE Trans. Commun.*, vol. 68, no. 9, Sept. 2020, pp. 5504–18.

[13] C. Bian *et al.*, "FusionNet: Enhanced Beam Prediction for mmWave Communications Using Sub-6GHz Channel and a Few Pilots," arXiv:2009.02655, 2020; available: https://arxiv.org/abs/2009.02655.

[14] Y. Yang *et al.*, "Deep Multimodal Learning: Merging Sensory Data for Massive MIMO Channel Prediction," *IEEE JSAC*, 2020, pp. 1–1.

[15] Y. Xu *et al.*, "User Grouping for Massive MIMO in FDD Systems: New Design Methods and Analysis," *IEEE Access*, vol. 2, Sept. 2014, pp. 947–59.

Biographies

Shun Zhang [M'13] received the B.S. degree in communication engineering from Shandong University, Jinan, China, in 2007, and the Ph.D. degree in communications and signal processing from Xidian University, Xi'an, China, in 2013. He is currently with the State Key Laboratory of Integrated Services Networks, Xidian University, where he is an associate professor.

His research interests include communication theory, signal processing for communications, artificial intelligence for communications, array signal processing, and convex optimizations.

Yushan Liu received the B.S. degree in communication engineering from Xidian University, Xi'an, China, in 2019, where he is currently pursuing the M.S. degree with the State Key Laboratory of Integrated Services Networks. His research interests include MIMO-OFDM systems, parameter estimation theory, and data detection.

Feifei Gao [F'20] is currently an associate professor in the Department of Automation, Tsinghua University, Beijing, China. His research interests include signal processing for communications, array signal processing, convex optimizations, and artificial intelligence assisted communications. He has authored/coauthored more than 150 refereed IEEE journal articles and more than 150 IEEE conference proceedings papers that are cited more than 8800 times in Google Scholar. He has served as a technical committee member for more than 50 IEEE conferences. He has also served as the Symposium Co-Chair of the 2019 IEEE ICC, the 2018 IEEE VTC Spring, the 2015 IEEE ICC, the 2014 IEEE Globecom, and the 2014 IEEE VTC Fall. He has served as an editor for IEEE TWC, IEEE TCCN, IEEE WCL, and *China Communications*, a lead guest editor for IEEE J-STSP, and a senior editor for IEEE SPL and IEEE COMML.

Chengwen Xing [S'08, M'10] received the B.Eng. degree from Xidian University, Xi'an, China, in 2005, and the Ph.D. degree from the University of Hong Kong, Hong Kong, China, in 2010. Since September 2010, he has been with the School of Information and Electronics, Beijing Institute of Technology, Beijing, China, where he is currently a full professor. From September 2012 to December 2012, he was a visiting scholar at the University of Macau. His current research interests include machine learning, statistical signal processing, convex optimization, multivariate statistics, array signal processing, high frequency band communication systems and so on.

Jianping An [M'08] received the Ph.D. degree from the Beijing Institute of Technology, China, in 1996. He joined the School of Information and Electronics, Beijing Institute of Technology, in 1995, where he is currently a full professor and the Dean. His research interests are in the field of digital signal processing, cognitive radio, wireless networks, and high-dynamic broadband wireless transmission technology. He has received two national awards for technological inventions and science and technology progress.

Octavia A. Dobre [M'05, SM'07, F'20] is a professor and Research Chair at Memorial University, Canada. She was a Royal Society and a Fulbright Scholar, and a visiting professor at Massachusetts Institute of Technology. Her research interests include technologies for beyond 5G, as well as optical and underwater communications. She has published over 300 referred papers in these areas. She serves as the Editor-in-Chief (EiC) of the *IEEE Open Journal of the Communications Society*. She was the EiC of *IEEE Communications Letters*, a senior editor and an editor with prestigious journals, as well as General Chair and Technical Co-Chair of flagship conferences in her area of expertise. She was the recipient of diverse awards, such as Best Paper Awards at IEEE ICC, IEEE Globecom, and IEEE WCNC conferences, as well as a Distinguished Lecturer of the IEEE Communications Society. She is a fellow of the the Engineering Institute of Canada.