CONTRASTIVE SELF-SUPERVISED LEARNING FOR WIRELESS POWER CONTROL

Navid Naderializadeh

HRL Laboratories, LLC
Malibu, California, USA

ABSTRACT

We propose a new approach for power control in wireless networks using self-supervised learning. We partition a multi-layer perceptron that takes as input the channel matrix and outputs the power control decisions into a backbone and a head, and we show how we can use contrastive learning to pre-train the backbone so that it produces similar embeddings at its output for similar channel matrices and vice versa, where similarity is defined in an information-theoretic sense by identifying the interference links that can be optimally treated as noise. The backbone and the head are then fine-tuned using a limited number of labeled samples. Simulation results show the effectiveness of the proposed approach, demonstrating significant gains over pure supervised learning methods in both sum-throughput and sample efficiency.

1. INTRODUCTION

Machine learning (ML), and especially deep learning, methods have shown very promising performance in many domains where the main challenge is to implicitly or explicitly model the underlying input distribution or the input/output relationship in the data. Among many areas that have benefited from the breakthroughs in ML, wireless communications is one of the most prominent ones, where ML has been used to tackle various problems, including network optimization and radio resource management (RRM) [1].

Under the ML-based RRM paradigm, earlier works focused on supervised approaches [2]. In these methods, ground-truth labels for RRM decisions were created using baseline algorithms, e.g., weighted minimum mean-squared error (WMMSE) [3], that had been shown to work well across a variety of scenarios, but exhibited much lower complexity than solving the original problems. Nevertheless, deriving optimal solutions for RRM problems is challenging as the corresponding optimization problems are typically very complex [4]. This has given rise to a plethora of work using other types of ML for RRM, such as reinforcement learning and unsupervised learning, which do not necessarily rely on manual labeling of the data for training the underlying neural networks [5, 6, 7, 8, 9].

Another paradigm in ML that has recently gained significant traction is self-supervised learning (SSL). The core idea of SSL is that the underlying neural network parameters can be pre-trained using labels that are generated manually from the data. For example, in case of visual data, an input image can be manually rotated, and the network can then be trained to predict the rotation angle [10]. SSL has recently shown striking success in representation learning in computer vision and natural language processing [11, 12].

In this paper, we use the notion of contrastive learning [13, 14, 15] for injecting SSL into learning optimal RRM decisions. We first break the underlying multi-layer perceptron (MLP) into a backbone, which maps the input channel matrix to an intermediate embedding, and a regression head, which maps the embedding to an output power control vector. We first pre-train the backbone using contrastive loss, which forces the embeddings of two similar channel matrices to be close, while the embeddings of two dissimilar channel matrices are pushed farther from each other. We use the information-theoretic condition for the optimality of treating interference as noise [16, 17] to identify and randomly remove weak interference links across the network to create similar channel matrices. Once pre-training is complete, the regression head can then be trained alongside the backbone using only a limited set of labeled training samples.

We show that the proposed self-supervised pre-training, coupled with the aforementioned information-theoretic channel augmentations, can effectively partition the embedding space into clusters within each of which the power control decisions are similar. Note that such a clustering happens without any supervision, i.e., the use of ground-truth power control decisions. We further show how the subsequent supervised training phase can train the regression head to produce power control decisions with a much fewer number of labeled samples compared to a pure supervised baseline.

2. PROBLEM FORMULATION

We consider a wireless network with $N$ transmitter-receiver pairs $\{(T_{x_i}, R_{x_i})\}_{i=1}^{N}$, where each transmitter $T_{x_i}$ intends to
send a message to its corresponding receiver $R_x_i$. The trans-
missions are assumed to occur using the same time and fre-
quency resources, implying that simultaneous transmissions of
two or more transmitters would cause interference on each
other. Let $x_i$ denote the signal transmitted by transmitter $T_x_i$, 
such that a transmit power constraint of $P_{\max}$. Then, the re-
served signal of receiver $R_x_i$ can be written as

$$y_i = h_{ii} x_i + \sum_{j \neq i} h_{ij} x_j + n_i,$$

(1)

where $h_{ij}$ denotes the channel gain between transmitter $T_x_j$
and receiver $R_x_i$, and $n_i \sim \mathcal{CN}(0, \sigma^2)$ denotes the additive
white Gaussian noise at receiver $R_x_i$. Letting $P_i \in [0, P_{\max}]$
denote the transmit power used by transmitter $T_x_i$, the Shan-
non capacity of the channel between transmitter $T_x_i$ and rec-
iever $R_x_i$ can be written as

$$R_i = \log_2 \left( 1 + \frac{|h_{ii}|^2 \gamma_i}{\sum_{j \neq i} |h_{ij}|^2 \gamma_j + \frac{\sigma^2}{P_{\max}}} \right),$$

(2)

where $\gamma_i = \frac{P_i}{P_{\max}} \in [0, 1]$ denotes the fraction of the max-
imum transmit power used by transmitter $T_x_i$.

In this paper, we focus on the problem of power control for sum-rate maximization across the network. In particular, we intend to solve the following optimization problem:

$$\max_{\gamma_1, \ldots, \gamma_N} \sum_{i=1}^{N} R_i$$

(3a)

s.t. $R_i = \log_2 \left( 1 + \frac{|h_{ii}|^2 \gamma_i}{\sum_{j \neq i} |h_{ij}|^2 \gamma_j + \frac{\sigma^2}{P_{\max}}} \right), \forall i$ (3b)

$$\gamma_i \in [0, 1], \forall i.$$ (3c)

3. PROPOSED APPROACH

We propose solving problem (3) using a learning-based ap-
proach. In particular, we consider a multi-layer perceptron
(MLP), represented as a mapping $f : \mathbb{R}^{N \times N} \rightarrow [0, 1]^N$, 
which takes as input the $N \times N$ channel matrix

$$H := \begin{bmatrix} |h_{11}|^2 & \cdots & |h_{1N}|^2 \\ \vdots & \ddots & \vdots \\ |h_{N1}|^2 & \cdots & |h_{NN}|^2 \end{bmatrix},$$

(4)

flattened into an $N^2$-dimensional vector, and generates the
power control vector $\Gamma = [\gamma_1 \ldots \gamma_N]^T$ at the output.

In order to train this MLP, we need a training dataset com-
prising multiple samples, each with an input channel matrix
$H$ and the corresponding output power control vector
$\Gamma$. However, the optimization problem (3) is non-convex and
challenging to solve exactly, especially as the number of transmitter-receiver pairs grows. To circumvent that is-
ue, similar to [2], we resort to deriving sub-optimal power control decisions using WMMSE [3]. Nevertheless, running
the WMMSE algorithm itself to create a massive dataset is
computationally expensive, which justifies using both labeled
samples, i.e., $(H, \Gamma)$ pairs, as well as unlabeled samples, only
including an input channel matrix $H$.

More formally, we assume that we have access to a set of
$M$ training samples, which consist of $M_L$ labeled samples
$\{ (H_i, \Gamma_i) \}_{i=1}^{M_L}$, alongside $M_U = M - M_L$ unlabeled samples
$\{ (H_i) \}_{j=M_L+1}^{M_U}$. Note that generating the input channel matrices is much more computationally than deriving the WMMSE
power control decisions, so it can be the case that $M_U \gg M_L$, 
which represents the so-called “low-label” regime.

As in most prior work in the area of wireless power control based on supervised learning, e.g., [2], we resort to deriving sub-optimal power
control decisions using stochastic gradient descent (SGD). In particular, using a batch of samples $\{ (H_i, \Gamma_i) \}_{i \in B_L}$, $B_L \subseteq \{ 1, \ldots, M_L \}$, the regression objective minimized via SGD can be written as

$$L_{\text{supervised}, B_L} = \frac{1}{|B_L|} \sum_{i \in B_L} \| \text{MLP}(H_i) - \Gamma_i \|_2^2$$

(5)

which denotes the mean-squared error (MSE) between the
MLP outputs and the ground-truth power control decisions.

In the following, we will demonstrate how we can use all
the labeled and unlabeled training samples to pre-train the
MLP parameters using self-supervised learning.

3.1. Self-Supervised Pre-Training

Let $l$ denote the number of neurons in the last hidden layer
of the MLP, right before the $N$-dimensional power control
decision outputs. The MLP can, therefore, be decomposed into a backbone $f : \mathbb{R}^{N \times N} \rightarrow \mathbb{R}^l$ and a head $g : \mathbb{R}^l \rightarrow [0, 1]^N$, where for any input channel matrix $H \in \mathbb{R}^{N \times N}$,

$$\text{MLP}(H) = g \circ f(H).$$

(6)

We call $f(H)$ the embedding of the channel matrix $H$. This
embedding is an intermediate representation of the channel
matrix, which is implicitly learned as part of the supervised
learning process. However, here we show that this embed-
ding can be pre-trained without the need for any labels, i.e.,
ground-truth WMMSE power control decisions.

We use the idea of contrastive learning [13][14][15], which
states that the embeddings of two similar channel matrices
should end up close to each other in the $l$-dimensional embed-
ding space, while the embeddings for two different channel
matrices should get as far as possible from each other. In par-
icular, consider a set of channel matrices tuples $\{ (H_i, \overline{H}_i) \}_{i \in B}$, where for each $i \in B$, $\overline{H}_i$ and $H_i$ are similar, while for any
$j \neq i$, $\overline{H}_i$ is different from $H_j$. Then, the backbone $f(\cdot)$ can be trained so as to minimize the following contrastive loss,

$$L_{\text{contrastive}, B} = \frac{1}{|B|} \sum_{i \in B} \left[ - \log \frac{e^{j(f(H_i), f(\overline{H}_i))}}{e^{j(f(H_i), f(\overline{H}_i))} + \sum_{j \neq i} e^{j(f(H_i), f(\overline{H}_j))}} \right],$$

(7)
where $\tau$ denotes a scalar temperature hyperparameter. In order to implement the contrastive loss in (7), we use mini-batch gradient descent on both labeled and unlabeled samples, i.e., \( \{H_i\}_{i=1}^M \). In particular, given a mini-batch \( \{H_i\}_{i \in B} \), where \( B \subseteq \{1, \ldots, M\} \), for each channel matrix \( H_i \), we create two augmentations \( \overline{H}_i \) and \( \widetilde{H}_i \), treating them as similar samples, while we use an augmentation of the rest of the channel matrices in the mini-batch for the dissimilar samples.

To create the channel matrix augmentations, we leverage the information-theoretic optimality condition for treating interference as noise [15]. Theorem 4 in [15] states that in a Gaussian interference channel as modeled in Section 2, treating interference as noise is information-theoretically optimal (to within a constant gap of the capacity region) if

\[
\frac{P_{\text{max}}|h_{ii}|^2}{\sigma^2} \geq \frac{P_{\text{max}}|h_{ij}|^2}{\sigma^2} \cdot \frac{P_{\text{max}}|h_{kl}|^2}{\sigma^2}, \forall i, j, k \neq i. \tag{8}
\]

As noted in [17], it is easy to verify that a sufficient condition for the condition in (8) would be that

\[
\frac{P_{\text{max}}|h_{ij}|^2}{\sigma^2} \leq \sqrt{\frac{P_{\text{max}}|h_{ik}|^2}{\sigma^2}}, \forall j \neq i. \tag{9}
\]

Indeed, if an interference link between transmitter \( T_{X_j} \) and a non-desired receiver \( T_{X_k} \) is so weak that it satisfies (9), then that interference link might be treated as noise without the loss of optimality (again, to within a constant gap).

In light of the aforementioned conditions, given a channel matrix \( H_1 \), we identify the weak interference links that satisfy (9). Then, we create an augmentation of the channel matrix by keeping each weak interference link with probability 0.5 and discarding it (setting it to zero) otherwise. Therefore, two different augmentations of a given channel matrix have the identical signal links as well as strong interference links, while their weak interfering link topologies might differ.

**Example 1.** Consider the case of \( N = 3 \) transmitter-receiver pairs, and assume that for a channel matrix \( H \) and some values of \( P_{\text{max}} \) and \( \sigma^2 \), we have

\[
\frac{P_{\text{max}}}{\sigma^2} \cdot H = \begin{bmatrix} 3 & 4 & 1 \\ 1 & 4 & 2 \\ 1 & 1 & 3 \end{bmatrix}. \tag{10}
\]

It can be verified that entries \( \{(1,3),(2,1),(3,1),(3,2)\} \) in (10) satisfy (9). Therefore, two possible augmentations for this channel matrix can be

\[
\overline{H} = \sigma^2 \frac{P_{\text{max}}}{\sigma^2} \begin{bmatrix} 3 & 4 & 0 \\ 0 & 4 & 2 \\ 1 & 0 & 3 \end{bmatrix}, \quad \widetilde{H} = \sigma^2 \frac{P_{\text{max}}}{\sigma^2} \begin{bmatrix} 3 & 4 & 0 \\ 1 & 4 & 2 \\ 0 & 1 & 3 \end{bmatrix}. \tag{11}
\]

To illustrate the effectiveness of such an augmentation method coupled with the contrastive self-supervised pre-training, we consider a network with \( N = 3 \) transmitter-receiver pairs, and a backbone \( f(\cdot) \) with a single 256-node hidden layer, LeakyReLU non-linearity, and an embedding size of \( l = 2 \). As suggested in prior work in the contrastive representation learning literature [14] [15] [18], we use \( \ell_2 \) normalization to ensure that each input channel matrix is mapped to an embedding on the unit sphere/circle. For simplicity, we set \( \frac{P_{\text{max}}}{\sigma^2} = 1 \) and generate \( M = 10^4 \) training channel matrices, where each channel gain in each matrix is a realization of a Rayleigh random variable with zero mean and unit variance, independent of all the other channel gains in the dataset. We use a batch size of 64, a learning rate of 0.05, and a temperature hyperparameter of \( \tau = 0.1 \).

Prior to the self-supervised pre-training procedure, we derive the 2-D embeddings of all the \( M = 10^4 \) channel matrices in the training set. We then pre-train the backbone for 10 epochs using the aforementioned augmentations and the contrastive loss (7), without using any of the ground-truth power control decisions. At the end of the pre-training phase, we derive the 2-D embeddings of the channel matrices again. Figure 1 demonstrates the embeddings before/after the pre-training phase. Each column represents the embeddings of the channel matrices with similar binary power control decisions.

**Fig. 1:** Embeddings of channel matrices before/after contrastive self-supervised pre-training for a network with \( N = 3 \) transmitter-receiver pairs. Each column illustrates the embeddings of channel matrices with similar optimal power control decisions. Each binary vector represents the on/off patterns of the transmitters, which is closest to the output of the WMMSE algorithm [3]. Note that these clusters emerge solely based on the channel matrices and the corresponding augmentations satisfying (9), without the use of any ground-truth power control decisions.
Figure 2 demonstrates the achieved sum-rates of our proposed approach (SSL) as a function of the fraction of labeled samples, i.e., \( \frac{N_u}{N} \). For comparison, we have included the WMMSE and full reuse (FR) baselines, where in FR, all transmitters always use their full transmit power \( P_{\text{max}} \). Moreover, the performance of a conventional approach which only uses the labeled samples for supervised learning is also demonstrated. The solid lines and shadowed regions respectively represent the mean and standard deviation over five training runs with different random seeds. As the figure shows, our proposed approach exhibits a significant performance boost compared to a pure supervised approach, especially in the low-label regime and for larger network sizes. In particular, for a given fraction of labeled samples, the proposed approach provides a higher sum-rate performance, and for a given target sum-rate, much fewer labeling effort is required to achieve the same performance level.

4. CONCLUSION

We considered the problem of downlink power control to maximize sum-throughput in wireless networks, where a multi-layer perceptron (MLP) is trained to produce power control decisions given an input channel matrix. We proposed using a self-supervised pre-training approach, in which the backbone (the MLP excluding the final layer) is trained using contrastive loss, so that similar channel matrices are represented using similar embeddings and vice versa. We used an information-theoretic condition for the optimality of treating interference as noise to create augmentations for the input channel matrices as a basis for computing the contrastive loss. We showed how a subsequent supervised training phase can boost the resulting sum-rate of the proposed method, using many fewer labels than a pure supervised approach.
5. REFERENCES

[1] Mingzhe Chen, Ursula Challita, Walid Saad, Changchuan Yin, and Mérouane Debbah, “Artificial neural networks-based machine learning for wireless networks: A tutorial,” IEEE Communications Surveys & Tutorials, vol. 21, no. 4, pp. 3039–3071, 2019.

[2] Haoran Sun, Xiangyi Chen, Qingjiang Shi, Mingyi Hong, Xiao Fu, and Nikos D Sidiropoulos, “Learning to optimize: Training deep neural networks for wireless resource management,” in 2017 IEEE 18th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC). IEEE, 2017, pp. 1–6.

[3] Qingjiang Shi, Meisam Razaviyayn, Zhi-Quan Luo, and Chen He, “An iteratively weighted mmse approach to distributed sum-utility maximization for a mimo interfering broadcast channel,” IEEE Transactions on Signal Processing, vol. 59, no. 9, pp. 4331–4340, 2011.

[4] Karl-Ludwig Besser, Bho Matthiesen, Alessio Zappone, and Eduard A Jorswieck, “Deep learning based resource allocation: How much training data is needed?,” in 2020 IEEE 21st International Workshop on Signal Processing Advances in Wireless Communications (SPAWC). IEEE, 2020, pp. 1–5.

[5] Yasar Sinan Nasir and Dongning Guo, “Multi-agent deep reinforcement learning for dynamic power allocation in wireless networks,” IEEE Journal on Selected Areas in Communications, vol. 37, no. 10, pp. 2239–2250, 2019.

[6] N. Naderializadeh, J. Sydir, M. Simsek, and H. Nikopour, “Resource management in wireless networks via multi-agent deep reinforcement learning,” in 2020 IEEE 21st International Workshop on Signal Processing Advances in Wireless Communications (SPAWC). IEEE, 2020, pp. 1–5.

[7] Mark Eisen, Clark Zhang, Luiz FO Chamon, Daniel D Lee, and Alejandro Ribeiro, “Learning optimal resource allocations in wireless systems,” IEEE Transactions on Signal Processing, vol. 67, no. 10, pp. 2775–2790, 2019.

[8] Mark Eisen and Alejandro Ribeiro, “Optimal wireless resource allocation with random edge graph neural networks,” IEEE Transactions on Signal Processing, 2020.

[9] N. Naderializadeh, M. Eisen, and A. Ribeiro, “Wireless power control via counterfactual optimization of graph neural networks,” in 2020 IEEE 21st International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), 2020, pp. 1–5.

[10] Spyros Gidaris, Praveer Singh, and Nikos Komodakis, “Unsupervised representation learning by predicting image rotations,” in International Conference on Learning Representations, 2018.

[11] Alexander Kolesnikov, Xiaohua Zhai, and Lucas Beyer, “Revisiting self-supervised visual representation learning,” in Proceedings of the IEEE conference on Computer Vision and Pattern Recognition, 2019, pp. 1920–1929.

[12] Zhenzhong Lan, Mingda Chen, Sebastian Goodman, Kevin Gimpel, Piyush Sharma, and Radu Soricut, “Albert: A lite bert for self-supervised learning of language representations,” in International Conference on Learning Representations, 2020.

[13] Aaron van den Oord, Yazhe Li, and Oriol Vinyals, “Representation learning with contrastive predictive coding,” arXiv preprint arXiv:1807.03748, 2018.

[14] Ting Chen, Simon Kornblith, Mohammad Norouzi, and Geoffrey Hinton, “A simple framework for contrastive learning of visual representations,” arXiv preprint arXiv:2002.05709, 2020.

[15] Tongzhou Wang and Phillip Isola, “Understanding contrastive representation learning through alignment and uniformity on the hypersphere,” arXiv preprint arXiv:2005.10242, 2020.

[16] Chunhua Geng, Navid Naderializadeh, Amir Salman Avestimehr, and Syed A Jafar, “On the optimality of treating interference as noise,” IEEE Transactions on Information Theory, vol. 61, no. 4, pp. 1753–1767, 2015.

[17] Navid Naderializadeh and Amir Salman Avestimehr, “ITLinQ: A new approach for spectrum sharing in device-to-device communication systems,” IEEE Journal on Selected Areas in Communications, vol. 32, no. 6, pp. 1139–1151, 2014.

[18] Feng Wang, Xiang Xiang, Jian Cheng, and Alan Loddon Yuille, “Normface: L2 hypersphere embedding for face verification,” in Proceedings of the 25th ACM international conference on Multimedia, 2017, pp. 1041–1049.

[19] Anders Gjendemsjø, David Gesbert, Geir E Øien, and Saad G Kiani, “Binary power control for sum rate maximization over multiple interfering links,” IEEE Transactions on Wireless Communications, vol. 7, no. 8, pp. 3164–3173, 2008.