



Towards an AI-native, user-centric air interface for 6G networks

D2.5 Digital Twin Training Sandbox

Contractual Delivery Date:	30-06-2025
Actual Delivery Date:	24-06-2025
Editor: <i>(name, organization)</i>	Osvaldo Simeone, King's College London
Deliverable nature:	R
Dissemination level:	PU
Version:	1.0

Keywords: Digital Twin, Prediction-Powered Inference, Doubly-Robust Learning, Context-aware Semi-Supervised Learning

ABSTRACT

One of the primary objectives of WP2 is the development of robust digital twin (DT) platforms for training and monitoring AI-driven methods. These platforms employ virtual twins to emulate their physical counterparts, facilitating iterative cycles of simulation, analysis, and optimization to enhance system performance. To ensure the reliability and generalizability of DT systems, we propose a semi-supervised learning (SSL) framework incorporating two novel training algorithms. The first method improves AI model training by



leveraging synthetic labels and reducing biases through a tuned cross-prediction-powered inference scheme. This approach is compatible with state-of-the-art SSL techniques, allowing seamless integration for enhanced performance. Additionally, we introduce context-aware doubly-robust (CDR) learning, an SSL strategy that dynamically adjusts its reliance on pseudo-labeled data based on the varying fidelity levels of the DT across different operational contexts. Evaluated on downlink beamforming tasks, CDR demonstrates superior performance over existing methods, achieving a 24% reduction in loss compared to conventional context-agnostic learning in low labeled-data regimes. These contributions improve the robustness and reliability of AI models training within DT platforms, ensuring their scalability and real-world applicability.

Disclaimer

This document contains material, which is the copyright of certain CENTRIC consortium parties, and may not be reproduced or copied without permission.

All CENTRIC consortium parties have agreed to full publication of this document.

Neither the CENTRIC consortium as a whole, nor a certain part of the CENTRIC consortium, warrant that the information contained in this document is capable of use, nor that use of the information is free from risk, accepting no liability for loss or damage suffered by any person using this information.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101096379. This publication reflects only the author's view and the European Commission is not responsible for any use that may be made of the information it contains.



Impressum

Full project title: Towards an AI-native, user-centric air interface for 6G networks

Short project title: CENTRIC

Number and title of the work package: WP2: AI-AI Hardware Platforms and Enablers

Number and title of task: T2.3: Training and Monitoring Environments

Document title: Digital Twin Training Sandbox

Editor: Osvaldo Simeone, King's College London

Work-package leader: Osvaldo Simeone, King's College London

© 2025 KCL and members of the CENTRIC consortium

Executive summary

This document presents the latest algorithmic advances from Task 2.3 of WP 2, all aimed at strengthening the reliability of digital-twin (DT) platforms for training AI models in 6G networks. We introduce two complementary semi-supervised learning frameworks that leverage DT-generated pseudo-data while explicitly correcting for its biases and context-dependent fidelity.

Reliable Semi-Supervised Learning via DTs: Semi-supervised learning (SSL) enhances model training by augmenting scarce labeled data with synthetic labels generated from a machine learning (ML) model but naively treating these pseudo-labels as ground truth can degrade performance. Building on the prediction-powered inference (PPI) framework, we explore methods to harness pseudo-labels while correcting their inherent bias. We introduce two novel PPI variants: Tuned cross-prediction-powered inference (CPPI), which adds a flexibility parameter to better adapt to pseudo-label quality, and Meta-CPPI (MCPPI), which jointly optimizes both the ML labeling models and the target model parameters. We validate these approaches on the problem of beam alignment using channel knowledge maps in mmWave networks and demonstrate that PPI-based methods consistently outperform standard SSL and purely supervised baselines. Notably, Meta-CPPI achieves the best results across all benchmarks, especially when labeled data are extremely limited.

Context-aware Doubly Robust Learning: The integration of AI into next-generation communication networks is hindered by the variability of traffic patterns and site-specific conditions, which demand highly contextual data. Leveraging both real-world measurements and synthetic pseudo-data generated by a DT offers a promising path forward, but its success depends critically on the DT's context-dependent fidelity. To overcome this challenge, we introduce context-aware doubly-robust (CDR) learning, a semi-supervised framework that dynamically adjusts its trust in pseudo-data according to the DT's accuracy in each context. We demonstrate CDR's effectiveness on a downlink beamforming task, where it consistently outperforms existing state-of-the-art semi-supervised methods.

List of authors

Company	Author	Contribution
KCL	Oswaldo Simeone	Editor
KCL	Housseem Sifaou	Contributor
KCL	Clement Ruah	Contributor

Table of Contents

Executive summary	4
List of authors	5
List of figures	7
Abbreviations	8
1 Introduction	9
1.1 Reliable SSL with Limited Labeled Data	9
1.2 Context-aware Doubly Robust Learning	9
2 Reliable Semi-Supervised Learning via Digital Twins	11
2.1 Background	11
2.2 Tuned CPPI with Neural Networks	11
2.3 Integration of Tuned CPPI with Meta Pseudo-Labels	12
2.4 Application: Beam Alignment	13
3 Context-Aware Doubly Robust Learning	15
3.1 Problem Setup	15
3.2 Background	15
3.3 Context-Aware Doubly Robust Self-Training	16
3.4 Experiments	16
5 Conclusion	19
References	20

List of figures

<i>Figure 1: Illustration of Meta Pseudo labels.</i>	<i>12</i>
<i>Figure 2: Illustration of Meta CPPI.....</i>	<i>12</i>
<i>Figure 3: Channel capacity as a function of the size of labeled dataset n when the number of BS antennas is fixed to 200.</i>	<i>13</i>
<i>Figure 4: Channel capacity as a function of the number of BS antennas when the size of labeled data is fixed to 500.....</i>	<i>14</i>
<i>Figure 5: Horizontal slice of the map of the urban region of interest in the beamforming experiments. The layouts of the buildings are depicted as empty white regions, and the BS position is marked by a black cross. The colored regions in (a) depict the ground-truth azimuth angles of the strongest path at each device location, while (b) reports the teacher model's azimuth estimates, which ignore the presence of LoS blockages due to buildings.</i>	<i>16</i>
<i>Figure 6: : Test loss as a function of the labeled data ratio ρ for ERM, P-ERM (2), DR [5], TDR, and the proposed CDR. The solid lines represent the median of 10 independent runs, and shaded areas indicate the range between the first and the third quartiles.</i>	<i>17</i>
<i>Figure 7: Test loss maps for the (a) ERM, (b) P-ERM (2), (c) DR [5], (d) TDR, and (e) CDR (4) with $n = 150$ labeled data points. (f) Test loss per context. (g) Evolution of the computed tuning parameters λ through training for TDR (purple) and CDR in NLoS (green dashed) and LoS (green solid) conditions, as well as the effective tuning parameter $1/(1+n/N)$ assumed by DR (black dashed). All levels correspond to the median of 10 independent runs, and error bars in (f) and shaded areas in (g) indicate the range between the first and the third quartiles.</i>	<i>18</i>

Abbreviations

AI	Artificial Intelligence
AoD	Angle of Arrival
BS	Base Station
CDR	Context-Aware Doubly Robust
CKM	Channel Knowledge Map
CPPI	Cross-Prediction-Powered Inference
CSI	Channel State Information
DR	Doubly Robust
DT	Digital Twin
ERM	Empirical Risk Minimization
LoS	Line of Sight
MCPPI	Meta Cross-Prediction-Powered Inference
MIMO	Multiple-Input Multiple-Output
ML	Machine Learning
MPL	Meta Pseudo Labels
MSE	Mean Squared Error
NLoS	Non-Line of Sight
P-ERM	Pseudo-empirical Risk Minimization
PPI	Prediction-Powered Inference
SS	Semi-Supervised
SSL	Semi-supervised Learning
TDR	Tuned Doubly Robust
UE	User Equipment

1 Introduction

Emerging 6G wireless networks are set to leverage increasingly advanced machine learning (ML) techniques to meet the demands of ultra-reliable, low-latency communications and massive connectivity. However, the practical deployment of these models is often constrained by the scarcity of high-quality labeled data in real-world environments. Semi-supervised learning (SSL) methods that combine limited labels with abundant unlabeled data offer a promising remedy, yet their success hinges on both the accuracy of generated pseudo-labels and the ability to adapt to changing contexts. In this report, we introduce robust SSL frameworks, Tuned cross-prediction-powered inference (CPPI) and context-aware doubly-robust (CDR) learning, that address bias in pseudo-labels, support modern neural architectures, and dynamically adjust to operational conditions for enhanced performance in 6G scenarios.

1.1 Reliable SSL with Limited Labeled Data

Next-generation wireless networks will increasingly rely on ML and data-driven decision-making [1] yet their success hinges on the availability of high-quality labeled data, a requirement that is often infeasible in practical deployments. SSL via pseudo-labeling offers a compelling solution by generating synthetic labels for unlabeled samples [2]. However, this approach can introduce bias when model predictions deviate from ground truth. To address this challenge, we built on the prediction-powered inference (PPI) framework [3] in *Deliverable D-2.3* to develop Tuned CPPI [4], a semi-supervised scheme that judiciously combines labeled and pseudo-labeled data for robust model training. In the present report, we extend Tuned CPPI to accommodate advanced architectures such as deep neural networks and demonstrate its broad applicability by integrating it with the state-of-the-art SSL techniques, thereby unlocking further performance gains.

1.2 Context-aware Doubly Robust Learning

We consider a similar SSL setting where, in addition to labeled data, the model has access to a pool of unlabeled data augmented with pseudo-labels generated by a teacher model. The conventional approach, pseudo-empirical risk minimization (P-ERM), trains a target model by minimizing the combined empirical loss over both labeled and pseudo-labeled data [5]. However, P-ERM's generalization performance is highly sensitive to the teacher model's accuracy, as errors in pseudo-labeling propagate into the final model.

To mitigate this limitation, doubly-robust (DR) self-training [6], inspired by [3], was introduced, correcting for bias arising from inaccurate pseudo-labels and demonstrating superior performance over existing SSL methods. Despite its advantages, DR exhibits a critical drawback in wireless network applications as it fails to incorporate contextual information, leading to suboptimal performance in dynamic environments. To bridge this gap, we propose context-aware doubly-robust (CDR) learning, an SSL framework that adaptively reweights pseudo-

labeled samples based on their contextual reliability. By dynamically adjusting confidence in pseudo-labels depending on operational conditions, CDR enhances robustness and generalization, particularly in scenarios with limited labeled data.

2 Reliable Semi-Supervised Learning via Digital Twins

2.1 Background

In many practical scenarios, obtaining high-quality labeled data is expensive and resource-intensive, whereas unlabeled data is often readily available in large quantities. Consider the case where we have access to a labeled dataset $D = \{X_i, Y_i\}_{i=1}^n$, consisting of independent and identically distributed (i.i.d.) samples drawn from an unknown joint distribution $P_{XY} = P_X \cdot P_{Y|X}$. In addition to this, we are provided with a substantially larger unlabeled dataset $\tilde{D} = \{\tilde{X}_i\}_{i=1}^N$, comprising N i.i.d. samples from the marginal distribution P_X . Typically, the size of the unlabeled dataset greatly exceeds that of the labeled dataset, i.e., $N \gg n$. Given a convex loss function $\ell_\theta(X, Y)$, the objective is to reliably estimate the minimizer θ^* of the population loss $L(\theta)$, i.e.,

$$\theta^* = \arg \min_{\theta} L(\theta), \text{ where } L(\theta) = \mathbb{E}[\ell_\theta(X, Y)].$$

Inspired by *Prediction-powered inference* (PPI) [3] [7] [8], we proposed in *Deliverable D-2.2*, Tuned CPPI, an SSL scheme that allows efficient use of labeled data for training teacher models to annotate unlabeled data while accounting for the potential bias of these models. Specifically, the labeled dataset is divided into K folds, D^1, \dots, D^K . For each k , a model $f^k(X)$ is trained on all labeled data except fold D^k . The latter will be used to account for the bias of model $f^k(X)$ as follows

$$L_\lambda^{CP}(\theta) = \frac{1}{n} \sum_{i=1}^n \ell_\theta(X_i, f(X_i)) + \lambda \left[\frac{1}{NK} \sum_{k=1}^K \sum_{i=1}^n \ell_\theta(\tilde{X}_i, f^k(\tilde{X}_i)) - \sum_{k=1}^K \sum_{i \in D^k} \ell_\theta(X_i, f^k(X_i)) \right].$$

Tuned CPPI offers the flexibility to tune the parameter $\lambda \in [0,1]$ as a function of the quality of the trained models, as detailed in our paper [4].

2.2 Tuned CPPI with Neural Networks

Although the PPI framework, including the proposed Tuned CPPI, relies on the assumption of convexity of the loss function $\ell_\theta(X, Y)$ for theoretical reasons, its practical benefits are not strictly limited to this setting. Empirical results suggest that the approach remains effective in real-world applications involving the training of neural networks, where loss functions are typically non-convex [6]. The approach consists of optimizing a slightly modified version of the PPI losses adapted to SGD algorithms. Specifically, at every gradient step t , the following loss is considered for Tuned CPPI

$$L_t^{PP}(\theta_t) = \lambda \sum_{k=1}^K \mathbb{E}_{\tilde{X} \sim \tilde{D}} [\ell_{\theta_t}(\tilde{X}, f^k(\tilde{X}))] - \kappa_t \frac{1}{K} \sum_{k=1}^K \mathbb{E}_{(X,Y) \sim D} [\lambda \ell_{\theta_t}(X, f^k(X)) - \lambda \ell_{\theta_t}(X, Y)]$$

where expectations in the above equation are approximated using minibatches from the corresponding dataset, and κ_t is a hyperparameter that allows for the gradual introduction of the bias term in the training process in order to avoid numerical instabilities.

2.3 Integration of Tuned CPPI with Meta Pseudo-Labels

Meta Pseudo Labels (MPL) [9] represent a strong benchmark in SSL, drawing inspiration from meta-learning techniques [10]. In contrast to traditional pseudo-labeling methods [1], where a fixed, pre-trained teacher model $f(\cdot)$ generates pseudo-labels for a separate student model $S(\cdot)$, MPL adopts a joint training strategy. Both the teacher and student models are trained simultaneously as shown in *Figure 1*. To enhance the quality of the pseudo-labels, a feedback mechanism is introduced: the teacher’s training loss incorporates a term that reflects the student’s performance on labeled data, thereby guiding the teacher to produce more effective pseudo-labels.

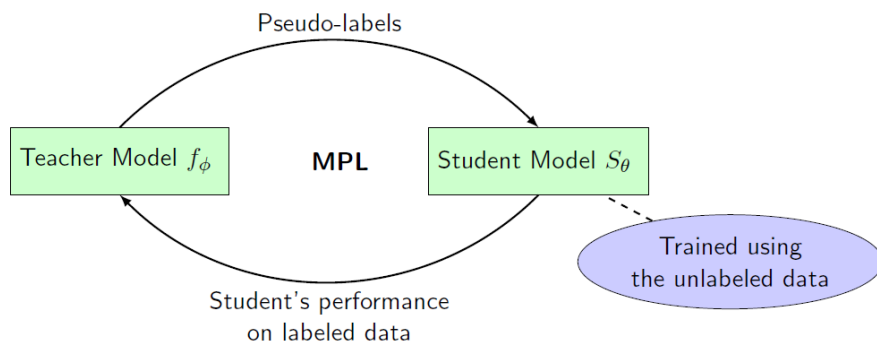


Figure 1: Illustration of Meta Pseudo labels.

We propose a novel framework that integrates MPL with the Tuned CPPI scheme. While Tuned CPPI follows a paradigm similar to classical pseudo-labeling, where a set of teacher models $\{f^k(\cdot)\}$ is trained once and then held fixed during the optimization of the student model parameters, our proposed approach here, termed Meta-CPPI (MCPPI), departs from this static setup. In MCPPI, the teacher models are trained jointly with the student model as illustrated in *Figure 2*. The student is optimized using the tuned CPPI loss, allowing dynamic feedback between the teacher and student to enhance learning performance.

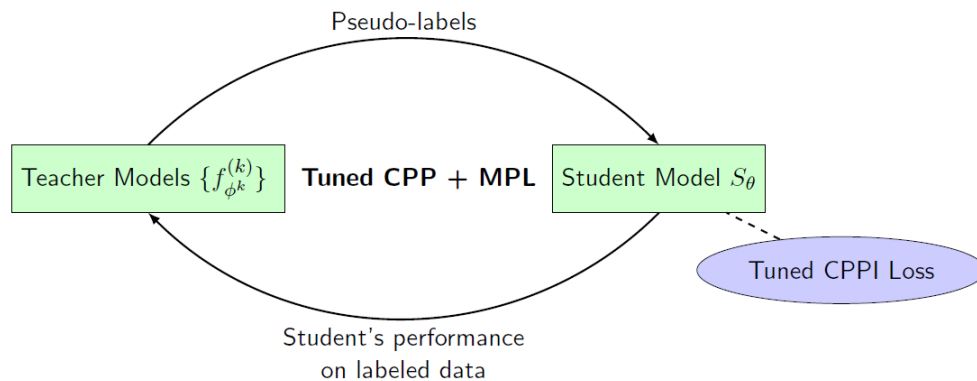


Figure 2: Illustration of Meta CPPI

2.4 Application: Beam Alignment

Beamforming plays a vital role in millimeter-wave (mmWave) massive MIMO systems [2], as it is essential for mitigating the significant path loss associated with high-frequency signal propagation. Traditional codebook-based beam alignment involves selecting the optimal beam from a predefined set of beams through beam sweeping, which relies on the transmission of pilot signals and can incur considerable training overhead. To address this limitation, a more recent approach introduces the concept of a Channel Knowledge Map (CKM), a site-specific database that captures channel characteristics associated with transmitter and receiver locations [11, 12]. We consider the same physical environment and dataset as in [11]. The dataset contains ground-truth multi-path channel information generated by using the *ray tracing* software Remcom Wireless Insite. The base station (BS) is equipped with a uniform planar array, and the user equipment (UE) is equipped with a single antenna. Furthermore, Kronecker product-based beamforming codebooks are employed. The total number of samples available in the dataset is 38038, from which n samples are reserved as labeled data, and the remaining $N = 38038 - n$ samples are considered as unlabeled data.

Applying the proposed framework, we train a mapping between a user's location and the optimal pair of beams by leveraging both labeled and unlabeled data. Labeled data consists of a user's location and the corresponding optimal beam pair, while unlabeled data only includes a user's location. We present here experiments of Tuned CPPI with neural networks, providing also comparison with MPL and with the proposed MCPPI.

We begin by evaluating the performance of Tuned CPPI against several benchmark methods, including empirical risk minimization (ERM), classical semi-supervised (SS) learning, PPI, Tuned PPI, and CPPI, as illustrated in *Figure 3*. Tuned CPPI demonstrates superior performance over CPPI and achieves a substantial improvement relative to ERM, which relies solely on labeled data. Both PPI and tuned PPI outperform ERM but their performance remains below that of CPPI and Tuned CPPI.

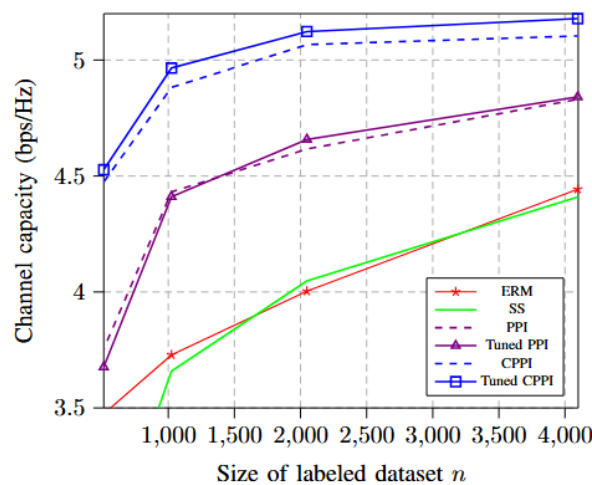


Figure 3: Channel capacity as a function of the size of labeled dataset n when the number of BS antennas is fixed to 200.

To further validate the effectiveness of the proposed methods, we compare their performance with the strong SSL benchmark MPL [9]. *Figure 4* presents the channel capacity as a function of the number of BS antennas for MPL, CPPI, Tuned CPPI, and the proposed MCPPI. The results indicate that while CPPI and Tuned CPPI achieve slightly lower performance than MPL, they remain competitive SSL approaches. Notably, MCPPI surpasses MPL, demonstrating improved performance and highlighting the benefits of the proposed joint training framework.

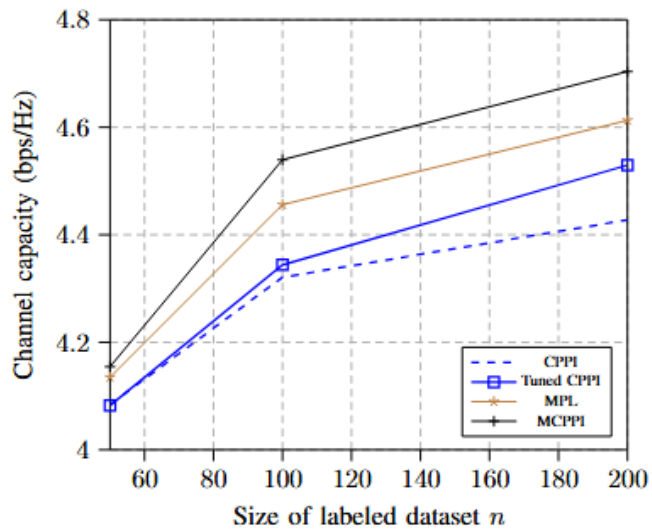


Figure 4: Channel capacity as a function of the number of BS antennas when the size of labeled data is fixed to 500.

3 Context-Aware Doubly Robust Learning

In this section, we introduce context-aware doubly-robust learning, an extension of classical doubly-robust training that adaptively adjusts the weighting of pseudo-labeled samples based on contextual information.

3.1 Problem Setup

We consider a learning scenario in which each sample is not only described by its input $X \in \mathcal{X}$ and label Y , but also by a discrete context variable $C \in \{1, \dots, K\}$ that stratifies the (X, Y) pairs into K distinct categories, representing properties shared by multiple data points such as collection-specific factors. The problem is cast in an SSL framework with a labeled dataset $\mathcal{D} = \{(x_i, c_i, y_i)\}_{i=1}^n \sim P(X, C, Y)$, and an unlabeled dataset $\tilde{\mathcal{D}} = \{(x_i, c_i)\}_{i=1}^N \sim P(X, C)$. Both datasets are assumed to be drawn i.i.d. from the joint distribution

$$P(X, C, Y) = P(C)P(X | C)P(Y | X, C).$$

In addition, we have access to a pre-trained *teacher model* $f: \mathcal{X} \rightarrow \mathcal{Y}$, which provides pseudo-label estimates $f(X)$ for each X . Our objective is to train a parametric probabilistic model $p(y|x; \theta)$ by optimizing over $\theta \in \Theta$. For any loss function $\ell_\theta(x, y)$, the optimal parameter $\hat{\theta}$ minimizes the population loss

$$L(\theta) = \mathbb{E}_{(X,Y) \sim P(X,Y)} [\ell_\theta(X, Y)],$$

approximated using the empirical loss on a given dataset \mathcal{D}

$$L_{\ell, \mathcal{D}}(\theta) = \frac{1}{|\mathcal{D}|} \sum_{(x,y) \in \mathcal{D}} \ell_\theta(x, y).$$

3.2 Background

Existing state-of-the-art methods typically neglect the role of contextual information. The conventional approach to the SSL problem, as described in the previous section, involves augmenting the labeled dataset with the samples from the unlabeled data with labels imputed using the teacher model $f(\cdot)$, *i. e.*, $\tilde{\mathcal{D}}^f = \{(\tilde{x}_i, f(\tilde{x}_i))\}_{i=1}^N$ [3, 7, 8, 4]. A direct approach to combine both datasets is to train an objective of the form

$$L^{P-ERM}(\theta) = L_{\ell, \mathcal{D} \cup \tilde{\mathcal{D}}^f}(\theta),$$

whose optimization is referred as pseudo-empirical risk minimization (P-ERM). The P-ERM training objective is a biased estimate of the population loss, leading to an inconsistent estimate. In order to circumvent this drawback, the DR scheme introduced in [6] adopts the training objective

$$L^{DR}(\theta) = L_{\ell, \mathcal{D}^f \cup \tilde{\mathcal{D}}^f}(\theta) - \left(L_{\ell, \mathcal{D}^f}(\theta) - L_{\ell, \mathcal{D}}(\theta) \right),$$

where the dataset $\mathcal{D}^f = \{(x_i, f(x_i))\}_{i=1}^n$ is constructed by replacing the true labels in \mathcal{D} with their corresponding pseudo-labels generated by the teacher model $f(\cdot)$. In the last equation, the first term represents the empirical loss evaluated on the union of pseudo-labeled data,

while the second term acts as a bias correction to account for the systematic error introduced by using the teacher model's predictions. This correction term ensures an unbiased estimate of the population loss.

3.3 Context-Aware Doubly Robust Self-Training

CDR starts by partitioning the labeled and unlabeled datasets according to the context variables c_i , that is, $\mathcal{D}_c = \{(x_i^c, y_i^c)\}_{i=1}^{n_c}$ and $\tilde{\mathcal{D}}_c = \{(\tilde{x}_i^c)\}_{i=1}^{N_c}$ for each context $c \in \{1, \dots, K\}$. Building on the stratified datasets, the contextual doubly-robust (CDR) objective extends the doubly-robust training loss. Specifically, it decomposes the contributions of both labeled and unlabeled data according to the different values of the context variable. For each context category c , a tuning parameter $\lambda_c \in [0,1]$ is introduced to control the balance between labeled and pseudo-labeled data contributions. This leads to the following training loss

$$L_\ell^{CDR(\lambda)}(\theta) = \sum_{c=1}^K \frac{\lambda_c N_c}{N} L_{\ell, \tilde{\mathcal{D}}_c^f}(\theta) - \left(\frac{n_c}{n} L_{\ell, \mathcal{D}_c}(\theta) - \frac{\lambda_c n_c}{n} L_{\ell, \mathcal{D}_c^f}(\theta) \right),$$

where $\lambda = [\lambda_1, \dots, \lambda_K] \in [0,1]^K$ is a tuning parameter vector. The inclusion of the parameters λ_c allows CDR to account for the extent to which the trained model relies on the pseudo-labeled data at context c . The tuning parameters are chosen at each iteration t to minimize the variance of the next parameter estimate at $t + 1$ as detailed in our paper [13].

3.4 Experiments

In this section, we study self-training for position-based downlink beamforming. The code for the experiments is available at [14]. As seen in *Figure 5*, we consider an outdoor urban setting with a single BS, where propagation paths between the BS and the devices are generated using ray-tracing as in [11].

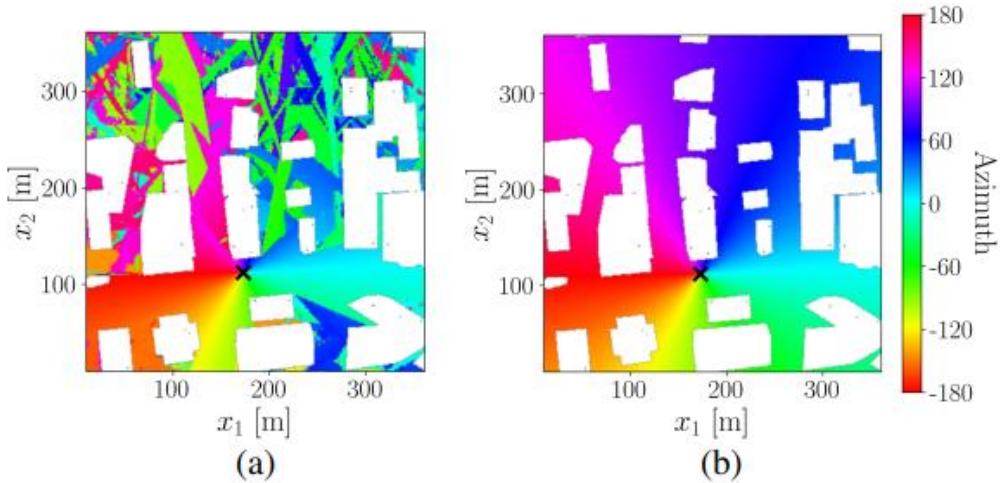


Figure 5: Horizontal slice of the map of the urban region of interest in the beamforming experiments. The layouts of the buildings are depicted as empty white regions, and the BS position is marked by a black cross. The colored regions in (a) depict the ground-truth azimuth angles of the strongest path at each device location, while (b) reports the teacher model's azimuth estimates, which ignore the presence of LoS blockages due to buildings.

As depicted in *Figure 5* for the azimuth, each device position X is associated with a groundtruth optimal azimuth and elevation angle of departure (AoD), obtained by considering the propagation path with the highest received power. The objective is to train a mapping g_θ that outputs an estimate of the optimal AoD at each device location. We augment each data point with a binary context variable $c_i \in [0,1]$, where $c_i = 1$ indicates that the device at location x_i is within line of sight (LoS) of the BS, and we set $c_i = 0$ otherwise, i.e., for non-line of sight (NLoS) coordinate.

Apart from the P-ERM and DR self-training procedures introduced earlier, we also consider the following benchmarks: 1) empirical risk minimization (ERM), which minimizes the standard loss limited to the labeled data; and 2) tuned doubly-robust (TDR) learning, which addresses the CDR objective with $\lambda_c = \lambda_0 \in [0,1]$ for all contexts $c \in \{1, \dots, K\}$, thus ignoring contextual information.

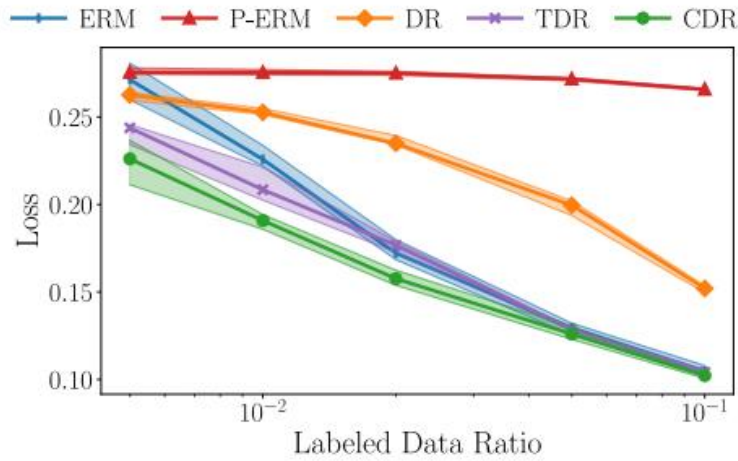


Figure 6: Test loss as a function of the labeled data ratio ρ for ERM, P-ERM (2), DR [6], TDR, and the proposed CDR. The solid lines represent the median of 10 independent runs, and shaded areas indicate the range between the first and the third quartiles.

Figure 6 shows the test loss for each estimator as we vary the proportion of labeled data. Both P-ERM and DR, lacking any mechanism to measure the teacher model’s reliability, perform worse than plain ERM, which ignores pseudo-labels altogether. By contrast, CDR dynamically adjusts its reliance on the teacher based on LoS and NLoS context, yielding lower test loss than all baselines, particularly when labeled samples are scarce.

Figure 7 illustrates how the test loss varies across spatial locations when only 150 labeled samples are used. Panels a–e display the loss map, while panel f highlights that both P-ERM and DR achieve near-optimal performance in LoS regions, where the teacher model is reliable, but fall behind standard ERM in NLo areas, where the teacher’s predictions degrade. By contrast, the TDR method effectively ignores the teacher in NLoS zones, as shown in panel g, its learned tuning parameter for NLoS converges to zero, causing its behavior to coincide with ERM. The proposed CDR approach goes further by adapting its reliance on the teacher separately for LoS and NLoS contexts. As a result, CDR matches DR’s performance in LoS

regions and ERM's robustness in NLoS regions, demonstrating its ability to seamlessly leverage the teacher model where it is accurate and disregard it where it is not.

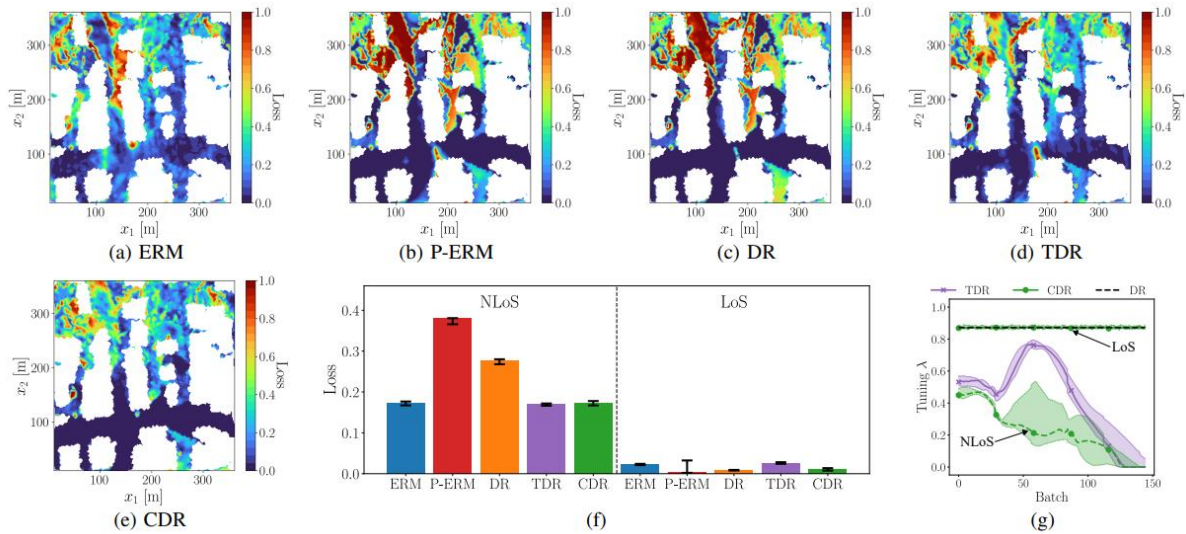


Figure 7: Test loss maps for the (a) ERM, (b) P-ERM (2), (c) DR [6], (d) TDR, and (e) CDR (4) with $n = 150$ labeled data points. (f) Test loss per context. (g) Evolution of the computed tuning parameters λ through training for TDR (purple) and CDR in NLoS (green dashed) and LoS (green solid) conditions, as well as the effective tuning parameter $1/(1+n/N)$ assumed by DR (black dashed). All levels correspond to the median of 10 independent runs, and error bars in (f) and shaded areas in (g) indicate the range between the first and the third quartiles.

5 Conclusion

This deliverable reports on Task 2.3 of WP 2, presenting several key enhancements to strengthen semi-supervised AI training for digital-twin platforms in 6G networks. We have extended the Tuned CPPI algorithm from *Deliverable D-2.2* to support high-capacity deep neural architectures and demonstrated its seamless integration with state-of-the-art SSL methods such as meta pseudo labels. Furthermore, we introduced a novel contextual doubly-robust scheme that dynamically adjusts the weight of pseudo-labels according to the teacher model's context-specific accuracy via an optimally derived tuning-parameter vector. Comprehensive downlink beamforming experiments confirm that CDR consistently outperforms context-agnostic self-training baselines, thereby enhancing the robustness, reliability, and efficiency of digital-twin platforms for 6G deployment.

References

- [1] O. Chapelle, B. Schölkopf and Z. Alexander, *Semi-supervised learning*, The MIT Press, 2006.
- [2] M. Giordani, M. Polese, A. Roy, D. Castor and M. Zorzi, "A tutorial on beam management for 3GPP NR at mmWave frequencies," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 1, pp. 173--196, 2018.
- [3] A. N. Angelopoulos, S. Bates, C. Fannjiang, M. I. Jordan and T. Zrnic, "Prediction-powered inference," *Science*, vol. 382, no. 6671, pp. 669--674, 2023.
- [4] H. Sifaou and O. Simeone, "Semi-Supervised Learning via Cross-Prediction-Powered Inference for Wireless Systems," *arXiv preprint arXiv:2405.15415*, 2024.
- [5] X. Yang, Z. Song, I. King and Z. Xu, "A survey on deep semi-supervised learning," *IEEE Transactions on Knowledge and Data Engineering*, vol. 35, no. 9, pp. 8934--8954, 2022.
- [6] B. Zhu, M. Ding, P. Jacobson, M. Wu, W. Zhan, M. Jordan and J. Jiao, "Doubly-Robust Self-Training," in *Advances in Neural Information Processing Systems*, 2023.
- [7] A. N. Angelopoulos, J. C. Duchi and T. Zrnic, "PPI++: Efficient prediction-powered inference," *arXiv preprint arXiv:2311.01453*, 2023.
- [8] T. Zrnic and E. J. Candès, "Cross-prediction-powered inference," *Proceedings of the National Academy of Sciences*, vol. 121, no. 15, p. e2322083121, 2024.
- [9] H. Pham, Q. Xie, D. Zihang and Q. V. Le, "Meta Pseudo Labels," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2021.
- [10] C. Finn, P. Abbeel and S. Levine, "Model-agnostic meta-learning for fast adaptation of deep networks," in *International conference on machine learning*, 2017.
- [11] Y. Zeng and X. Xu, "Toward environment-aware 6G communications via channel knowledge map," *IEEE Wireless Communications*, vol. 28, no. 3, pp. 84--91, 2021.
- [12] D. Wu, Y. Zeng, S. Jin and R. Zhang, "Environment-aware and training-free beam alignment for mmWave massive MIMO via channel knowledge map," in *IEEE International Conference on Communications Workshops*, 2021.
- [13] C. Ruah, H. Sifaou, O. Simeone and B. Al-Hashimi, "Context-Aware Doubly-Robust Semi-Supervised Learning," *IEEE Signal Processing Letters*, 2025.
- [14] C. Ruah, "Code Repository for Context-Aware Doubly-Robust SemiSupervised Learning," 2025. [Online]. Available: <https://github.com/kclip/cdr-ssl>.