Hypernetwork-aided Channel Estimation for Integrated Data and Energy Transfer

Yushi Lei, Yusha Liu, *Member, IEEE*, Jie Hu, *Senior Member, IEEE*, and Kun Yang, *Fellow, IEEE*

***Abstract*—Integrated data and energy transfer (IDET) is an important component of future communication networks due to its characteristic to provide continuous and reliable wireless power to battery-limited devices. In order to get full advantages of IDET technology in future communication networks and realize efficient beamforming design, the base station (BS) must obtain accurate downlink channel state information (CSI). In this work, we propose an energy harvesting (EH) aided channel estimation scheme using a novel deep learning (DL) architecture in an end-to-end training mode. This architecture simultaneously obtains CSI from both an energy harvester and a data decoder at an IDET receiver. By designing a hypernetwork-aided deep neural network (DNN), we achieve more accurate channel estimation, while effectively reducing the pilot overhead in channel estimation. Simulation results show that compared to the state of the art, our EH aided channel estimation scheme attains a lower normalized mean square error (NMSE).**

***Index Terms*—Integrated data and energy transfer (IDET), energy harvesting (EH) aided channel estimation, hypernetwork.**

I. Introduction

Radio frequency (RF) signals have been used both in wireless data transfer (WDT) and wireless energy transfer (WET). In recent years, with the explosive growth of low-power devices, such as Internet-of-Things (IoT) [1] and wireless sensor networks (WSN) [2], IDET has attracted much attention because of its potential in battery-limited wireless networks. Researchers have explored IDET in conjunction with various advanced technologies, including multiple-input multiple-output (MIMO) [3] and cognitive radio networks [4]. Among these, the integration of MIMO with IDET, commonly referred to as MIMO-IDET, has demonstrated significant improvements in transmission efficiency [3]. This integration allows for enhanced data and energy transfer capabilities, making it a promising solution for modern wireless communication systems. However, the effectiveness of such systems is heavily dependent on the accurate estimation of CSI, as CSI directly impacts the efficiency of both data transmission and EH processes. Given the importance of precise CSI estimation, ongoing research in relevant area focuses on developing advanced methodologies to enhance the accuracy and reliability of channel estimation. These efforts are vital for ensuring that MIMO-IDET systems can achieve their potential in terms of performance and efficiency, particularly in environments characterized by the presence of numerous low-power devices and the need for efficient energy management.

In existing literature, channel estimation methods typically include traditional approaches such as least squares (LS) and minimum mean-square error (MMSE), as well as advanced techniques like compressed sensing (CS) and DL-based methods. For WDT CSI acquisition, methods primarily employ the transmitted data signal for channel estimation, with uplink CSI usually playing a supporting role in downlink channel estimation, and increasingly, these methods are based on DL with different neural network architectures [5]–[7]. For example, Dong *et al.* [5] developed a deep convolutional neural network (CNN)-based channel estimation approach for MIMO-orthogonal frequency division multiplexing (OFDM) systems, which estimates the channel from the received signal at the receiver, enhancing robustness across various channel conditions. Guo *et al.* [6] proposed a high-precision DL-based downlink CSI acquisition framework that leverages uplink CSI. Recognizing the lack of full radio channel reciprocity between uplink and downlink in frequency division duplex (FDD) systems, Banerjee *et al.* [7] developed a conditional generative adversarial network (CGAN) approach for uplink-to-downlink CSI mapping. Different from WDT, WET channel estimation methods generally use the amplitude information present in energy signals to estimate partial downlink CSI. L.A.López *et al.* [8] leveraged the received energy signal for LS estimation to design the energy beamforming for powering multiple users with stringent EH demands in massive MIMO system. Other studies [9], [10] have employed DL methods for channel estimation by having the energy transmitter (ET) estimate the channel based on harvested energy feedback from the energy receiver (ER). On the basis of WDT and WET, the IDET system, which receives both data and energy signals at the receiver, offers potential for a more adaptable channel estimation scheme. However, existing WDT and WET channel estimation methods cannot be directly applied to IDET, as IDET aims to balance data rate and harvested energy, rather than prioritizing one over the other. Xu *et al.* [11] explored cascaded channel estimation in an intelligent reflecting surface-assisted IDET system, transforming the problem into a sparse

The authors are with School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu, Sichuan 611731, China (emails: 202321010929@std.uestc.edu.cn; yusha.liu@uestc.edu.cn; hujie@uestc.edu.cn; kunyang@uestc.edu.cn). Kun Yang is also with the School of Computer Science and Electronic Engineering, University of Essex, Colchester, CO4 3SQ, U.K.

This work was supported in part by MOST Major Research and Development Project under Grant 2021YFB2900204; in part by the Natural Science Foundation of China under Grant 62132004 and Grant 62301123; in part by the Key Research and Development Program of Zhejiang Province under Grant 2022C01093.

Corresponding author: Jie Hu.

Simulation codes of this article can be found at https://github.com/yushi-lei/hypernetwork-aided-channel-estimation-for-IDET.

signal reconstruction task solvable via CS. By contrast, with channel estimation implemented via an MMSE estimator, Amudala *et al.* [12] modeled a MIMO relaying system considering spatially-correlated relay-user channels and used a non-linear EH model to optimize the weighted sum energy efficiency metric. These diverse approaches highlight the need for innovative solutions tailored to the unique challenges of IDET systems.

However, despite employing various methods to enhance channel estimation performance, existing IDET literature has primarily concentrated on using data signals for channel estimation, neglecting the potential benefits of energy signals. This oversight limits the effectiveness of current approaches, since they fail to fully exploit the dual nature of IDET systems, which inspires our design. In this letter, we introduce the HyperDNN architecture, a high-performance DL-based joint scheme for pilot training and channel estimation in IDET systems by employing the hypernetwork framework. Our proposed HyperDNN leverages both data and energy signals, aiming to improve the channel estimation efficiency on an existing basis. The main contributions of this work are summarized as follows.

- We develop an end-to-end hypernetwork-aided DNN structure for channel estimation in an IDET system, referred to as HyperDNN, where pilot generation and CSI estimation are achieved by DNNs, and the whole process can be jointly optimized.
- We introduce a hypernetwork [13] for extracting channel amplitude information from energy harvested by the IDET receiver to help with channel estimation. Instead of employing a big DNN with all data and energy signals as inputs, the hypernetwork is a newly introduced relatively small DNN. It takes energy signal as input and outputs weight parameter for the main DNN to assist in channel estimation, where the main DNN is also a relatively small DNN with data signal as input. This hypernetwork guarantees a certain degree of the channel estimation accuracy while reducing network size.
- Simulation results show that by jointly utilizing the data and energy signals, the proposed HyperDNN architecture can perform channel estimation more accurately than traditional method, and its performance is better than that only uses data signal. Moreover, a shorter pilot sequence is designed by DNN in an end-to-end training manner, reducing the system overhead.

The rest of the paper is organized as follows. Section II describes our system model. Section III introduces the proposed HyperDNN architecture. Simulation results are provided in Section IV, while Section V provides the main conclusions.

Notations: In this paper, the uppercase and lowercase boldface letters, $\mathbf{X}$ and $\mathbf{x}$, denote matrices and vectors, respectively. $\mathbb{C}^{m\times n}$ represents the dimension of a matrix/vector. Furthermore, $(\cdot)^{-1}$, $(\cdot)^{T}$, and $(\cdot)^{H}$ represent matrix inversion, transpose, and Hermitian transpose operations, respectively. $|\cdot|$ and $\|\cdot\|$ denote the absolute value and the Euclidean norm. $\Re(\cdot)$ and $\Im(\cdot)$ denote the real and imaginary parts of a complex vector or matrix. $\mathrm{E}[\cdot]$ denotes the statistical expectation, $\mathrm{vec}(\cdot)$ denotes the vectorization of a matrix by stacking columns/rows. $\mathcal{CN}(m,n)$ represents the distribution of a circularly symmetric complex Gaussian (CSCG) random variables with a mean m and a variance n. $\mathcal{U}(a,b)$ represents a uniform distribution between a and b.

II. System Model

In this section, we introduce a multiple-input single-output (MISO) IDET system model, as shown in Fig. 1, where an M-antenna IDET transmitter at the BS transmits RF signal to a single-antenna IDET receiver.

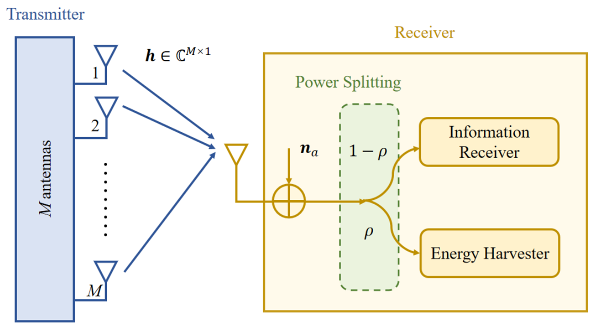

Fig. 1. System Model.

In the t-th $(t=1,2,\cdots,T)$ time slot, a pilot signal is first transmitted, which is followed by data transmission. At the beginning of each time slot, an M-antenna IDET transmitter first transmits L pilot symbols $\mathbf{X}\in\mathbb{C}^{M\times L}$ to the single-antenna IDET receiver. The power constraint of the pilot symbols is expressed as

$$\|\mathbf{x}_l\|^2 \le P,\ \forall l=1,\cdots,L, \tag{1}$$

where $\mathbf{x}_l$ represents the l-th column of the pilot signal $\mathbf{X}$ and P is the maximum transmit power of the IDET transmitter.

We consider a quasi-static flat-fading channel, where the channel remains invariant within each single time slot and changes independently across different time slots. Specifically, we assume a multi-path channel with P paths [14]. Each path is characterized by the angle of departure (AoD) θ_p, the path gain $\alpha_p \sim \mathcal{CN}\left(0,\sigma_\alpha^2\right)$, and fading amplitudes $\beta_p \sim \mathcal{CN}\left(0,\sigma_\beta^2\right)$. Based on the above assumptions, the channel $\mathbf{h}\in\mathbb{C}^{M\times 1}$ can be modelled as

$$\mathbf{h}=\sum_{p=1}^{P}\alpha_p\mathbf{a}\left(\theta_p\right)\beta_p, \tag{2}$$

where $\mathbf{a}(\theta_p)\in\mathbb{C}^{M\times 1}$ is the steering vector. We consider the uniform linear array (ULA) antenna with the steering vector being expressed as

$$\mathbf{a}\left(\theta_p\right)=\left[1,e^{-\frac{j2\pi d}{\lambda}\sin\left(\theta_p\right)},\cdots,e^{-\frac{j2\pi d}{\lambda}(M-1)\sin\left(\theta_p\right)}\right]^T, \tag{3}$$

where d is the antenna spacing, and λ is the carrier wavelength.

At any time slot, at the IDET receiver, the L received discrete-time pilot samples $\mathbf{y}\in\mathbb{C}^{1\times L}$ can be expressed as

$$\mathbf{y}=\left(\mathbf{h}\right)^H\mathbf{X}+\mathbf{n}_a, \tag{4}$$

where $\mathbf{h}$ is the channel vector given by (2) and $\mathbf{n}_a \in \mathbb{C}^{1\times L}$ is additive white Gaussian noise (AWGN) with zero mean and variance σ_a^2. The receiver is capable of executing IDET functionality. Specifically, the IDET receiver is equipped with a circuit that can harvest energy and decode data from received signals at the same time by using the power splitting structure [15]. Accordingly, the receiving vector $\mathbf{y}$ can be divided into two parts, one is used for EH and the other for data decoding. As shown in Fig. 1, the received data signal $\mathbf{y}_I$ and energy signal $\mathbf{y}_E$ after power splitting can be modelled respectively as

$$\mathbf{y}_I = \sqrt{1-\rho}\,\mathbf{y}, \tag{5a}$$

$$\mathbf{y}_E = \eta\rho\left|\mathbf{y}\right|^2, \tag{5b}$$

where $\rho \in [0,1]$ is the power splitting ratio and $\eta \in [0,1]$ denotes the energy conversion efficiency. Note that a linear EH model is used in this work. For non-linear energy harvester, we can modified the structure by employing a trained DNN to simulate the actual receiving energy [16]. In conventional approaches [17], CSI is solely extracted from data signal $\mathbf{y}_I$ by using traditional channel estimation methods like MMSE usually [18], [19], and then acquired by the BS.

III. Hypernetwork-aided channel estimation

Different from the conventional methods which extract downlink CSI only from the data signal $\mathbf{y}_I$, in this section, we propose an end-to-end channel estimation approach that also takes the channel information extracted from the harvested energy into consideration, as shown in Fig. 2. Specifically, at the transmitter, a linear layer is employed for pilot training. The received pilot signals at the receiver are first processed by a power splitter, obtaining the data and the energy signal. During the channel estimation process, the proposed Hyper-DNN learns downlink CSI from both data signal $\mathbf{y}_I$ via the main DNN and energy signal $\mathbf{y}_E$ via the hypernetwork. The hypernetwork deals with energy signal $\mathbf{y}_E$, outputting weight parameters to assist channel estimation of the main DNN, hence improving the accuracy of channel estimation while imposing moderate complexity. The hyperDNN structure is comprised with a main DNN and a hypernetwork, which are discussed in Sections III-A and III-B, respectively, and the training details of the proposed HyperDNN for channel estimation in IDET system is discussed in Section III-C.

A. Main DNN

Pilot transmission and power splitting are discussed in Section II. After the power splitting, we build a main DNN for performing channel estimation, which learns downlink CSI embedded in the data signal.

In the main DNN, the channel estimation is implemented by applying a multi-layer fully-connected DNN and the downlink CSI is extracted from data signal $\mathbf{y}_I$ only. We separate the real and imaginary components of $\mathbf{y}_I$ and employ them as inputs to the main DNN, with $\Re\left(\mathbf{y}_I\right)$ and $\Im\left(\mathbf{y}_I\right)$ representing the real and imaginary parts of the elements in $\mathbf{y}_I$ respectively. Further, the complex to real value representation can be denoted as the $2L\times 1$ vector $\mathrm{c2r}\left(\mathbf{y}_I\right) = \left[\Re\left(\mathbf{y}_I\right)^T, \Im\left(\mathbf{y}_I\right)^T\right]^T$.

The procedure of acquiring downlink CSI from $\mathbf{y}_I$ can be carried out through the $M^{(\mathrm{CE})}$-layer fully-connected DNN, which can be expressed as

$$\begin{aligned}\mathrm{c2r}\left(\hat{\mathbf{h}}\right) =& \mathbf{W}^{(\mathrm{CE})}_{M^{(\mathrm{CE})}}\left[\cdots f_{\mathrm{ReLU}}\left(\mathbf{W}^{(\mathrm{CE})}_1 \mathrm{c2r}\left(\mathbf{y}_I\right) + \mathbf{b}^{(\mathrm{CE})}_1\right)\cdots\right] \\ &+ \mathbf{b}^{(\mathrm{CE})}_{M^{(\mathrm{CE})}}, \\ \triangleq& f^{(\mathrm{CE})}\left(\mathbf{y}_I \middle| \mathbf{\Omega}^{(\mathrm{CE})}\right),\end{aligned} \tag{6}$$

where $\mathrm{c2r}(\hat{\mathbf{h}})$ is the output of the main DNN, which represents the channel estimation values acquired through the main DNN processing, with the dimension of $2M \times 1$. $\left\{\mathbf{W}^{(\mathrm{CE})}_1, \mathbf{b}^{(\mathrm{CE})}_1, \ldots, \mathbf{W}^{(\mathrm{CE})}_{M^{(\mathrm{CE})}}, \mathbf{b}^{(\mathrm{CE})}_{M^{(\mathrm{CE})}}\right\}$ is the set of the optimization parameters for the main DNN, where $\mathbf{W}^{(\mathrm{CE})}_m$ are the weight matrixes and $\mathbf{b}^{(\mathrm{CE})}_m$ are the bias vectors of the m-th layer. Denoting $\mathbf{\Omega}^{(\mathrm{CE})} = \left\{\mathbf{W}^{(\mathrm{CE})}_1, \mathbf{b}^{(\mathrm{CE})}_1, \ldots, \mathbf{W}^{(\mathrm{CE})}_{M^{(\mathrm{CE})}}, \mathbf{b}^{(\mathrm{CE})}_{M^{(\mathrm{CE})}}\right\}$ in (6) as the symbol of optimization parameters, and $\ell^{(\mathrm{CE})}_m$, $m = 1, \ldots, M^{(\mathrm{CE})}$ as the number of neurons in the fully connected layers. The dimension of the weight matrixes $\mathbf{W}^{(\mathrm{CE})}_m$ can be expressed as $\ell^{(\mathrm{CE})}_m \times \ell^{(\mathrm{CE})}_{m+1}$ and the dimension of the bias vectors $\mathbf{b}^{(\mathrm{CE})}_m$ are $\ell^{(\mathrm{CE})}_m \times 1$, with the known condition that $\ell^{(\mathrm{CE})}_{M^{(\mathrm{CE})}} = 2M$. $f_{\mathrm{ReLU}}(\cdot)$ is the rectified linear unit (ReLU) activation function of first $\left(M^{(\mathrm{CE})}-1\right)$ fully connected layers, while the output layer produces the estimated channel $\mathrm{c2r}(\hat{\mathbf{h}})$ through the linear (or identity) function, i.e., $\varphi_o(v) = v$.

B. Hypernetwork

While the main DNN deals with the received data signal $\mathbf{y}_I$, we also construct a hypernetwork to extract the amplitude information of the channel from energy signal $\mathbf{y}_E$, assisting the channel estimation. Accordingly, a more precise channel estimation at the IDET transmitter may be obtained.

The input of the hypernetwork is the values of harvested energy $\mathbf{y}_E$, which is a $L \times 1$ vector denoted by $g(\mathbf{y}_E)$. Through the processing of a $M^{(\mathrm{H})}$-layer fully-connected DNN, the output of the hypernetwork represents a parameter for adjusting the weights of the main DNN and ultimately aiding channel estimation. This procedure can be expressed as

$$\begin{aligned}\boldsymbol{\omega} =& \mathbf{W}^{(\mathrm{H})}_{M^{(\mathrm{H})}}\left[\cdots f_{ReLU}\left(\mathbf{W}^{(\mathrm{H})}_1 g\left(\mathbf{y}_E\right) + \mathbf{b}^{(\mathrm{H})}_1\right)\cdots\right] + \mathbf{b}^{(\mathrm{H})}_{M^{(\mathrm{H})}}, \\ \triangleq& f^{(\mathrm{H})}\left(\mathbf{y}_E \middle| \mathbf{\Omega}^{(\mathrm{H})}\right),\end{aligned} \tag{7}$$

where $\boldsymbol{\omega}$ is the common weight parameters at each time slot, which is used to adjust the weights of the main DNN. $\left\{\mathbf{W}^{(\mathrm{H})}_1, \mathbf{b}^{(\mathrm{H})}_1, \cdots, \mathbf{W}^{(\mathrm{H})}_{M^{(\mathrm{H})}}, \mathbf{b}^{(\mathrm{H})}_{M^{(\mathrm{H})}}\right\}$ is the set of the optimization parameters for the hypernetwork, where $\mathbf{W}^{(\mathrm{H})}_m$ denote the weights and $\mathbf{b}^{(\mathrm{H})}_m$ denote the bias. Denoting $\mathbf{\Omega}^{(\mathrm{H})} = \left\{\mathbf{W}^{(\mathrm{H})}_1, \mathbf{b}^{(\mathrm{H})}_1, \cdots, \mathbf{W}^{(\mathrm{H})}_{M^{(\mathrm{H})}}, \mathbf{b}^{(\mathrm{H})}_{M^{(\mathrm{H})}}\right\}$ in (7) as the symbol of optimization parameters, and $\ell^{(\mathrm{H})}_m$, $m = 1, \cdots, M^{(\mathrm{H})}$ as the number of neurons in the fully connected layers, thus, $\mathbf{W}^{(\mathrm{H})}_m$ are $\ell^{(\mathrm{H})}_m \times \ell^{(\mathrm{H})}_{m+1}$ weight matrices, $\mathbf{b}^{(\mathrm{H})}_m$ are $\ell^{(\mathrm{H})}_m \times 1$ bias vectors and the output of the hypernetwork $\boldsymbol{\omega}$ is a $\left(\sum_{m=1}^{M^{(\mathrm{H})}} \ell^{(\mathrm{H})}_m\right) \times 1$ vector. The ReLU activation function is also used in the first

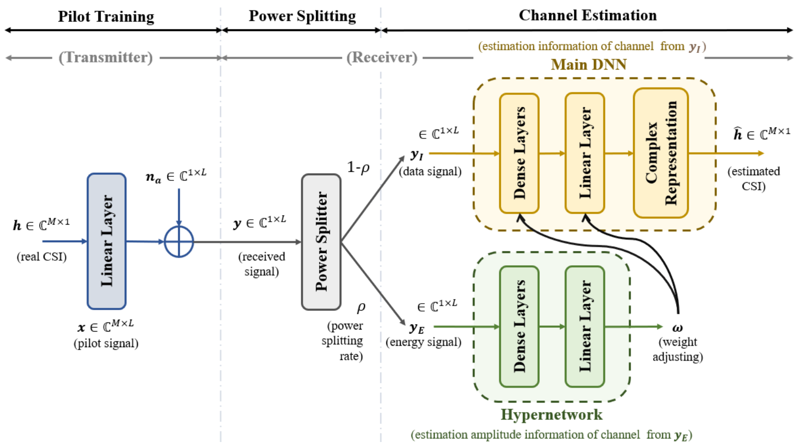

Fig. 2. The proposed HyperDNN architecture for channel estimation of IDET.

$\left(M^{(\mathrm{H})}-1\right)$ fully connected layers, and linear (or identity) function is used in the last layer.

Further, we virtually split the output of the hypernetwork $\boldsymbol{\omega}$ into $\boldsymbol{\omega} = [\boldsymbol{\omega}_1, \cdots, \boldsymbol{\omega}_{M^{(\mathrm{CE})}}]$, where $\boldsymbol{\omega}_m \in \mathbb{C}^{\ell_m^{(\mathrm{CE})}\times 1}$. The adjustment method of $\left(\sum_{m=1}^{M^{(\mathrm{H})}} \ell_m^{(\mathrm{H})}\right) \times 1$ vector $\boldsymbol{\omega}$ to the weights of the main DNN can be expressed as

$$\mathbf{W}_m^{(\mathrm{CE})} = \tilde{\mathbf{W}}_m^{(\mathrm{CE})} \cdot \mathrm{diag}\left\{\boldsymbol{\omega}_m\right\}, \quad m = 1, \cdots, M^{(\mathrm{CE})}, \tag{8}$$

where $\tilde{\mathbf{W}}_m^{(\mathrm{CE})}$ contains the weight parameters of the main DNN to be adjusted. We define $\tilde{\boldsymbol{\Omega}}^{(\mathrm{CE})} = \left\{\tilde{\mathbf{W}}_1^{(\mathrm{CE})}, \mathbf{b}_1^{(\mathrm{CE})}, \cdots, \tilde{\mathbf{W}}_{M^{(\mathrm{CE})}}^{(\mathrm{CE})}, \mathbf{b}_{M^{(\mathrm{CE})}}^{(\mathrm{CE})}\right\}$ as the set of optimization parameters for the channel estimation of the proposed energy signal aided channel estimation structure.

Note that we emphasize the IDET system in this work, in which data signal may not carry complete CSI information due to power splitting and data processing. Therefore, we consider using hypernetwork to extract the amplitude channel information also carried in energy signal, which is exclusive to the IDET system, to help with channel estimation. When using conventional system that received signal is not split, there is no energy signal and no need to introduce the hypernetwork [13].

C. Training

The proposed energy signal aided channel estimation model consists of both the main DNN and the hypernetwork, and is trained in an end-to-end approach that aims to minimize the training squared error between the estimated and real channel. The corresponding optimization problem can be formulated as

$$\min_{\mathbf{x}, \boldsymbol{\Omega}^{(\mathrm{CE})}, \boldsymbol{\Omega}^{(\mathrm{H})}} \mathrm{E}\left[\left\|\hat{\mathbf{h}} - \mathbf{h}\right\|^2\right], \tag{9a}$$

$$\text{s.t. } \hat{\mathbf{h}} = f^{(\mathrm{CE})}\left(\mathbf{y}_I \middle| \tilde{\boldsymbol{\Omega}}^{(\mathrm{CE})}, \boldsymbol{\omega}, \boldsymbol{\Omega}^{(\mathrm{H})}\right), \tag{9b}$$

$$\boldsymbol{\omega} = f^{(\mathrm{H})}\left(\mathbf{y}_E \middle| \boldsymbol{\Omega}^{(\mathrm{H})}\right), \tag{9c}$$

$$\left\|\mathbf{x}_l\right\|^2 \leq P, \ \forall l = 1, \cdots, L, \tag{9d}$$

where $\mathbf{y}_I$ and $\mathbf{y}_E$ are respectively the data and energy parts of the received signal after power splitting; $f^{(\mathrm{CE})}(\cdot)$ represents the training process of the main DNN with input $\mathbf{y}_I$; optimization parameters include $\tilde{\boldsymbol{\Omega}}^{(\mathrm{CE})}$, ω and $\boldsymbol{\Omega}^{(\mathrm{H})}$; $f^{(\mathrm{H})}(\cdot)$ represents the training process of the hypernetwork with input $\mathbf{y}_E$ and the optimization parameters $\boldsymbol{\Omega}^{(\mathrm{H})}$.

The optimization problem of (9) can be implemented offline in practice according to Algorithm 1. Specifically, noise samples are first generated according to noise statistics. Note that the noise considered in this paper is AWGN, with a mean of 0 and a variance of σ_a^2. Secondly, the channel samples are generated according to the channel model of (2). We use the standard P-path multipath channel with AoD θ_p, path gain α_p and fading amplitude β_p, where the AoD θ_p obeys the uniform distribution of $\theta_p \sim \mathcal{U}(-\pi/6, \pi/6)$. Since we assume the quasi-static time-invariant fading channel, the path gain α_p and fading amplitude β_p are independent random values that obey the Gaussian random distribution. After the generation of channel samples, the system model and neural networks structure are constructed. Finally, the samples are sent to the neural networks for training. Pilots and other optimization parameters in this structure are jointly optimized in an end-to-end training form under the premise of transmitter power limitation.

Compared with the basic DL channel estimation scheme, in which the main DNN structure is used and the parameters $\boldsymbol{\Omega}^{(\mathrm{CE})}$ are optimized only, the proposed energy signal aided channel estimation method also optimizes the parameters $\boldsymbol{\Omega}^{(\mathrm{H})}$ of the hypernetwork and the pilot signal $\mathbf{X}$. At the same time, the proposed scheme also learns how to estimate the channel more accurately from the received signal, reducing the pilot overhead of downlink transmission. The training parameters are given in detail in Section IV.

IV. Numerical Results

In this section, we evaluate the channel estimation performance of the proposed HyperDNN. The simulation setting details are given in Section IV-A, while the numerical results are shown in Section IV-B.

A. Simulation Settings

We adopt the 3GPP standardized spatial channel model (SCM) [14] for the proposed structure. Since we assume a

Algorithm 1: Training Procedure of the Proposed Hypernetwork aided Channel Estimation

Input: M, L, ρ, P, σ_a^2 `// system parameters;`
θ_p, α_p, β_p `// channel parameters;`
N_{ep}, N_{batch}, N_b `// numbers of epochs, batches, and batch size;`
$\alpha^{(0)}$ `// initial learning rate;`

Objective: minimize NMSE $= \mathrm{E}[\|\hat{\mathbf{h}} - \mathbf{h}\|^2 / \|\mathbf{h}\|^2]$

Output: $\mathbf{x}, \Omega^{(\mathrm{CE})}, \Omega^{(\mathrm{H})}$ `// pilot, and optimization parameters of DNN;`

Initialization:
  $i \leftarrow 0$;

Training:

while $i < N_{ep}$ **do**
  if $i < N_{ep}/3$ **then**
    $\alpha^{(i)} = 10^{-3}$;
  else
    if $i < 3N_{ep}/4$ **then**
      $\alpha^{(i)} = 10^{-4}$;
    else
      $\alpha^{(i)} = 10^{-5}$;
    end
  end
  $t \leftarrow 0$;
  while $t < N_{batch}$ **do**
    Generate training set $\mathcal{S}^{(t)}$ on batch size N_b;
    Update optimization parameters on set $\mathcal{S}^{(t)}$ by Adam optimizer with the objective of minimizing NMSE, including
    1.Update $\Omega^{(\mathrm{H})}$ of the hypernetwork and output $\boldsymbol{\omega}$ using (7);
    2.Update $\Omega^{(\mathrm{CE})}$ of the main DNN and output $\mathrm{c2r}\left(\hat{\mathbf{h}}\right)$ using (6) and (8);
  end
end

time invariant channel, the channel model uses i.i.d. fading amplitudes β_p in this experiment. For power splitter, we use a fixed energy conversion efficiency η (i.e. η=1) and a power splitting ratio ρ that can be adjusted. The details of the simulation parameters are listed in Table I.

We use the standard DL libraries TensorFlow and Keras to implement the proposed energy signal aided channel estimation structure. The whole network is trained with 10^4 epochs and 1024 mini-batch size, with the learning rate gradually decrease from 10^{-3} to 10^{-5}. For the basic DL approach (basic DNN) only utilizing data signal $\mathbf{y}_I$, we adopt a 4-layer DNN structure, including 3 dense layer and a binary output layer, with $\ell_1^{(\mathrm{CE})} = 1024$, $\ell_2^{(\mathrm{CE})} = 512$, $\ell_3^{(\mathrm{CE})} = 512$ ReLU hidden neurons and $2M = 128$ output nodes. For the proposed energy signal aided channel estimation structure, on the basis of adopting basic DNN, a 3 dense layer one binary layer neural hypernetwork structure is employed, with $\ell_1^{(\mathrm{H})} = 128$, $\ell_2^{(\mathrm{H})} = 512$, $\ell_3^{(\mathrm{H})} = 1024$ ReLU hidden neurons, the number of output nodes is equal to the sum of input nodes and hidden neurons in basic DNN. The parameter initialization of the neural network is realized by Xavier initialization on the TensorFlow. During the optimization process, we adopt the adaptive moment estimation (Adam) optimizer. After each iteration, pilot signal $\mathbf{X}$ is normalized to satisfy the power constraint $\|\mathbf{x}_l\|^2 \leq P$. The objective function uses NMSE to measure the channel estimation performance, which is expressed as NMSE $= \mathrm{E}[\|\hat{\mathbf{h}} - \mathbf{h}\|^2 / \|\mathbf{h}\|^2]$.

TABLE I
SIMULATION PARAMETERS

Parameters	Values
No. of paths (P)	2
AoDs (θ_p)	$\mathcal{U}(-\pi/6, \pi/6)$
No. of transmit antennas (M)	64
Pilot length (L)	$8, 16, 32$
Power splitting Rate (ρ)	$[0.2, 0.9]$
Energy conversion efficiency (η)	1
Learning rate per epoch ($\alpha^{(i)}$)	10^{-3} to 10^{-5}
epochs (N_{ep})	10^4
batches (N_{batch})	1
Mini-batch size (N_b)	1024

B. Numerical Results

We first compare the NMSE of the proposed HyperDNN approach with the basic DNN which only utilizes data signal $\mathbf{y}_I$, the traditional MMSE approach, and the SF-CNN mentioned in [5]. The calculation formula for MMSE estimator can be expressed as

$$\hat{\mathbf{h}}_{\mathrm{MMSE}} = \mathbf{R}_h \left(\mathbf{X}^H \mathbf{X} \mathbf{R}_h + \sigma_a^2 \mathbf{I}\right) \mathbf{X}^H \mathbf{y}_I, \tag{10}$$

where $\mathbf{R}_h$ is the correlation matrix of the real channel $\mathbf{h}$ which can be denoted as $E\left[\mathbf{h}\mathbf{h}^H\right]$, $\mathbf{I}$ is the identity matrix. The NMSE versus signal-to-noise ratio (SNR) is shown in Fig. 3, where power splitting rate ρ is 0.3. We can see that with the increase of pilot lengths L and SNR, the NMSE decreases, because it represents a more accurate channel testing and a more ideal environmental condition. At a pilot length of 16, hyperDNN has a lower NMSE than SF-CNN at a high SNR. Furthermore, the proposed HyperDNN achieves the lowest NMSE, especially compared with basic DNN and MMSE approaches, due to exploiting the energy signal.

In Fig. 4, we study the influence of different power splitting rates ρ on channel estimation NMSE. We can see that a power splitting rate region between 0.2 and 0.4 brings better training performance, achieving a lower NMSE. This is because an intermediate power splitting rate, not too large nor too small, can minimize noise interference to the data signal while allowing the energy signal to contribute to the CSI estimation.

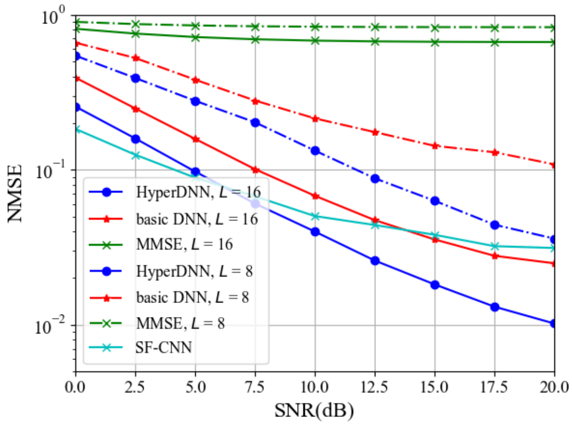

Fig. 3. NMSE of the proposed method, basic DNN, MMSE, and SF-CNN method for channel estimation with $M = 64$, $\rho = 0.3$, with different pilot lengths L (the pilot transmission mode in SF-CNN is fixed [5]).

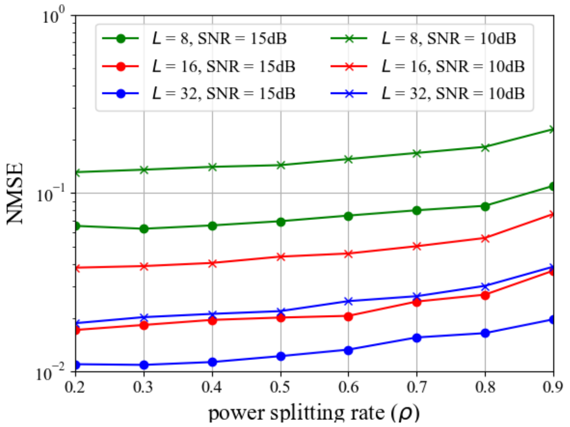

Fig. 4. NMSE of the proposed method over frequency-flat fading channels having different power splitting rate ρ for an $M = 64$ system, with SNR$= 10$ and 15dB, pilot lengths $L = 8, 16$ and 32.

When the power splitting rate is too large, the data signal becomes smaller, making it more susceptible to noise during RF-to-baseband conversion. However, with the improvement of training conditions (such as the increase of pilot length L and/or SNR), when the system itself can have a lower NMSE, the performance gain brought by the power splitting rate is less obvious.

V. Conclusions

In this paper, we proposed an energy signal-assisted channel estimation framework for IDET, referred to as HyperDNN. Unlike existing channel estimation methods, our approach features a DL-based end-to-end structure where pilot training and CSI estimation are jointly optimized. Additionally, we have employed a hypernetwork to extract channel information from the energy signal. The proposed HyperDNN framework introduces additional performance gains, offering valuable insights for future research.

Our future work will consider more practical channel characteristics, e.g., the channel's temporal correlations, Doppler frequency offset. Moreover, we will explore solutions for how to address the channel estimation in the FDD mode, where uplink and downlink channels differ.

References

[1] H. Xiu, Z. Gao, A. Liao, Y. Mei, D. Zheng, S. Tan, M. D. Renzo, and L. Hanzo, "Joint activity detection and channel estimation for massive IoT access based on millimeter-wave/terahertz multi-panel massive MIMO," *IEEE Trans. Veh. Technol.*, vol. 72, no. 1, pp. 1349–1354, 2023.

[2] Y. Liu, I. Al-Nahhal, O. A. Dobre, and F. Wang, "Deep-learning channel estimation for IRS-assisted integrated sensing and communication system," *IEEE Trans. Veh. Technol.*, vol. 72, no. 5, pp. 6181–6193, 2023.

[3] V. Khodamoradi, A. Sali, O. Messadi, A. Khalili, and B. B. M. Ali, "Energy-efficient massive MIMO SWIPT-enabled systems," *IEEE Trans. Veh. Technol.*, vol. 71, no. 5, pp. 5111–5127, 2022.

[4] D. Kumar, P. K. Singya, K. Choi, and V. Bhatia, "SWIPT enabled cooperative cognitive radio sensor network with non-linear power amplifier," *IEEE Trans. Cognit. Commun. Networking*, vol. 9, no. 4, pp. 884–896, 2023.

[5] P. Dong, H. Zhang, G. Y. Li, I. S. Gaspar, and N. NaderiAlizadeh, "Deep CNN-based channel estimation for mmWave massive MIMO systems," *IEEE J. Sel. Topics in Signal Process.*, vol. 13, no. 5, pp. 989–1000, 2019.

[6] J. Guo, C.-K. Wen, and S. Jin, "Canet: Uplink-aided downlink channel acquisition in FDD massive MIMO using deep learning," *IEEE Trans. Commun.*, vol. 70, no. 1, pp. 199–214, 2022.

[7] B. Banerjee, R. C. Elliott, W. A. Krzymień, and H. Farmanbar, "Downlink channel estimation for FDD massive MIMO using conditional generative adversarial networks," *IEEE Trans. Wireless Commun.*, vol. 22, no. 1, pp. 122–137, 2023.

[8] O. L. A. López, D. Kumar, R. D. Souza, P. Popovski, A. Tölli, and M. Latva-Aho, "Massive MIMO with radio stripes for indoor wireless energy transfer," *IEEE Trans. Wireless Commun.*, vol. 21, no. 9, pp. 7088–7104, 2022.

[9] J.-M. Kang, C.-J. Chun, and I.-M. Kim, "Deep-learning-based channel estimation for wireless energy transfer," *IEEE Commun. Lett.*, vol. 22, no. 11, pp. 2310–2313, 2018.

[10] J.-M. Kang, C.-J. Chun, I.-M. Kim, and D. I. Kim, "Deep RNN-based channel tracking for wireless energy transfer system," *IEEE Sys. J.*, vol. 14, no. 3, pp. 4340–4343, 2020.

[11] X. Xu, W. Zhu, S. Yang, J. Bao, W.-P. Zhu, and Z. Liu, "Second order rectified parallel factor model based cascaded channel estimation in IRS-assisted SWIPT system," *IEEE Trans. Veh. Technol.*, vol. 72, no. 10, pp. 13314–13325, 2023.

[12] D. N. Amudala and R. Budhiraja, "Wireless information and power transfer enabled massive MIMO multi-way relaying," *IEEE Trans. Wireless Commun.*, vol. 22, no. 10, pp. 6654–6672, 2023.

[13] Y. Liu and O. Simeone, "Learning how to transfer from uplink to downlink via hyper-recurrent neural network for FDD massive MIMO," *IEEE Trans. Wireless Commun.*, vol. 21, no. 10, pp. 7975–7989, 2022.

[14] 3GPP TR 25.996 V16.0.0, "Spatial channel model for multiple input multiple output (MIMO) simulations (Release 16)," *3rd Generation Partnership Project Std.*, Jul. 2020.

[15] Y. Zheng, S. A. Tegos, Y. Xiao, P. D. Diamantoulakis, Z. Ma, and G. K. Karagiannidis, "Zero-energy device networks with wireless-powered RISs," *IEEE Trans. Veh. Technol.*, vol. 72, no. 10, pp. 13655–13660, 2023.

[16] M. Varasteh, J. Hoydis, and B. Clerckx, "Learning to communicate and energize: Modulation, coding, and multiple access designs for wireless information-power transmission," *IEEE Trans. Commun.*, vol. 68, no. 11, pp. 6822–6839, 2020.

[17] B. Kim, J.-M. Kang, H.-M. Kim, and J. Kang, "Joint channel estimation, training design, Tx power allocation, and Rx power splitting for MIMO SWIPT systems," *IEEE Commun. Lett.*, vol. 25, no. 4, pp. 1269–1273, 2021.

[18] X. Zhou, "Training-based SWIPT: Optimal power splitting at the receiver," *IEEE Trans. Veh. Technol.*, vol. 64, no. 9, pp. 4377–4382, 2015.

[19] G. Dong, H. Zhang, and D. Yuan, "Downlink achievable rate of massive MIMO enabled SWIPT systems over Rician channels," *IEEE Commun. Lett.*, vol. 22, no. 3, pp. 578–581, 2018.