Denoising and Voltage Estimation in Modular Multilevel Converters Using Deep Neural-Networks

SAÚL LANGARICA1, (Member, IEEE), GERMÁN PIZARRO1, PABLO M. POBLETE1, (Graduate Student Member, IEEE), FELIPE RADRIGÁN1, (Graduate Student Member, IEEE), JAVIER PEREDA3,1,2, (Member, IEEE), JOSE RODRIGUEZ3, (Life Fellow, IEEE), AND FELIPE NÚÑEZ1, (Member, IEEE)

1Department of Electrical Engineering, Pontificia Universidad Católica de Chile, Santiago 7820436, Chile
2UC Energy Research Center, Pontificia Universidad Católica de Chile, Santiago 7820436, Chile
3College of Engineering, Universidad Andres Bello, Santiago 8370146, Chile

Corresponding author: Felipe Núñez (fenunez@ing.puc.cl)

This work was supported in part by grants ANID PIA ACT192013 and ANID/FONDECYT/1171142. The work of Saúl Langarica was supported by grant ANID/Doctorado Nacional/21200565. The work of Pablo M. Poblete was supported by grant ANID/Doctorado Nacional/21201967.

ABSTRACT Modular Multilevel Converters (MMCs) have become one of the most popular power converters for medium/high power applications, from transmission systems to motor drives. However, to operate properly, MMCs require a considerable number of sensors and communication of sensitive data to a central controller, all under relevant electromagnetic interference produced by the high frequency switching of power semiconductors. This work explores the use of neural networks (NNs) to support the operation of MMCs by: i) denoising measurements, such as stack currents, using a blind autoencoder NN; and ii) estimating the sub-module capacitor voltages, using an encoder-decoder NN. Experimental results obtained with data from a three-phase MMC show that NNs can effectively clean sensor measurements and estimate internal states of the converter accurately, even during transients, drastically reducing sensing and communication requirements.

INDEX TERMS Modular multilevel converters, neural networks, cyber-physical systems.

I. INTRODUCTION

Multilevel converters (MCs) are a cost-effective solution for medium to high-voltage power conversion. MCs have been under intense research and development during the last decades and have shown wide adoption in industrial applications such as high power AC drives, high-voltage direct-current transmission, integration of renewable energy, power filters and static synchronous compensators [1], [2].

Among the different topologies for MCs, Modular Multilevel Converters (MMCs) have become one of the most attractive options for medium/high power applications [3]. The MMC stands out from other topologies due to its high modularity, impressive scalability, fault tolerance and high output power quality, which means very low harmonic distortion with reduced passive filters [4]. MMCs operate at high voltages with standard power semiconductors, but with a high voltage DC-link and a floating capacitor in each sub-module (SM). Nowadays, the MMC is widely found in commercial solutions, such as the ABB HVDC light, Alstom MaxSine, Siemens Sinamics SM120N and the Benshaw M2L series.

MMCs contain up to thousands of SMs depending on the application, which poses several challenges for its operation and control. Power converters are a natural source of electromagnetic interference (EMI) due to the high \( \frac{dv}{dt} \) and \( \frac{di}{dt} \) generated by power semiconductors. EMI can lead to high frequency spikes in measurements and it also may interfere with adjacent communications, monitoring and control systems. Hence, its proper management is imperative [5].

A second challenge comes from the fact that the modular structure of the MMC is not exploited in the control scheme,
which usually is centralized and requires a large number of measurements and signals, including at least one voltage measurement per each SM. This increases the complexity and cost of the converter, and reduces reliability and noise tolerance [6]. Several efforts have been made to reduce the number of voltage sensors in MMCs, in order to simplify the hardware requirements and increase reliability and noise immunity [6]. The Kalman filter has been proposed as an estimator for capacitor voltage in [7], but the formulation uses an extra voltage sensor per stack and its slow dynamic response may affect control performance during transients and faults. In [8], an adaptive linear neural network is proposed for SM voltage estimation, adding sensors to the stack inductors and output voltage of each phase. However, this method presents a considerable tracking error.

This work proposes the use of Deep Neural Networks (DNNs) to support the operation of MMCs. Particularly, it is proposed to combine Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) in an encoder-decoder architecture, for: i) denoising measurements from the MMC, with the aim of providing cleaner signals, with negligible delay, to the control system; and ii) estimating the capacitor voltage of each SM of the MMC, with the aim of eliminating most of the sensors required to operate the MMC. The proposed denoiser is able to blind-denoise the signals, i.e., denoise without knowing any prior information of the noise, while the proposed estimator is capable of estimating the capacitor voltages in different states of operation, even during transients. The effectiveness and performance of the proposed techniques are evaluated using real operational data from a three-phase MMC, with 4 sub-modules per stack, obtained under both steady-state and transient operation.

II. BACKGROUND
A. NOTATION AND BASIC DEFINITIONS
In this work, \( \mathbb{R} \) denotes the real numbers, \( \mathbb{Z}_{>0} \) the positive integers, \( \mathbb{R}^n \) the Euclidean space of dimension \( n \), and \( \mathbb{R}^{n \times m} \) the set of \( n \times m \) matrices with real coefficients. For \( a, b \in \mathbb{Z}_{>0}, [a; b] \) denotes their closed interval in \( \mathbb{Z} \). For a vector \( \nu \in \mathbb{R}^n, \nu^i \) denotes its \( i \)th component. For a matrix \( A \in \mathbb{R}^{n \times m}, A_i \) denotes its \( i \)th column and \( A^i \) its \( i \)th row. For an \( n \)-dimensional real-valued sequence \( \alpha : \mathbb{Z}_{>0} \rightarrow \mathbb{R}^n \), \( \alpha(t) \) denotes its \( t \)th element and \( \alpha([a; b]) \) its restriction to the interval \([a; b] \), i.e., a sub-sequence. For an \( n \)-dimensional sub-sequence \( \alpha([a; b]), M(\alpha([a; b])) \in \mathbb{R}^{n \times (b-a+1)} \) is a matrix whose \( i \)th column is equal to \( \alpha(a+i-1) \), with \( i \in [1; b-a+1] \). The same notation applies to an \( n \)-dimensional finite-length sequence \( \alpha : [1; \bar{T}] \rightarrow \mathbb{R}^n \) provided \([a; b] \subseteq [1; \bar{T}]\) holds. Given an \( n \)-dimensional sequence \( \alpha \), its \( T \)-depth window is a matrix-valued sequence \( \beta : \mathbb{Z}_{>0} \rightarrow \mathbb{R}^{n \times T} \), where \( \beta(t) = M(\alpha([t-T+1; t])) \).

B. MODULAR MULTILEVEL CONVERTERS
The schematic representation of the MMC under study is shown in Figure 1. The three-phase MMC consists of three legs connected in parallel to a common dc-bus \( V_{DC} \). Each leg is divided in two stacks, upper \((u)\) and lower \((l)\), while each stack is formed by the series connection of \( N \) SMs with a stack inductor \( L_s \) and an equivalent resistance \( R_s \). The midpoint of each leg is connected to an AC output terminal \((a, b, c)\).

The number of SMs per stack depends on the voltage of the application. In a typical HVDC application, the number of SMs per stack can be over 500 [6], with a voltage of ±400 kV. On the other hand, motor drive applications need between 5 – 20 SMs for an operating voltage of 3.3 – 13.8 kV [9].

There are plenty of SMs topologies currently proposed in the literature. Nevertheless, the basic and most commonly used one is the half-bridge (HB) [10], due to its high efficiency, simplicity and straightforward SM capacitor voltage balancing. The HB consists in two transistors operating in a complementary manner, and a single floating capacitor \((C)\), generating two possible voltage levels \((v_c \text{ and } 0)\). The variation of the capacitor voltage is a function of the stack current and the switching state of the power semiconductors.

C. ENCODER-DECODER NEURAL NETWORKS
RNNs have been extensively used in applications where time series have to be processed, as modeling of dynamical systems or fault detection, mainly because of its recurrent structure, similar to a dynamical system in state-space form. However, recent studies, have brought to light problems of the classical RNN architecture, such as difficulties to retain long-term information and learn complex patterns. To address these problems, in [11] an architecture consisting in two cascaded RNNs was proposed, where the first network

![FIGURE 1. Schematic representation of an MMC.](image-url)
(the encoder) maps a variable length input sequence to a fixed-length vector representation in an intermediate space and then, the second network (the decoder) maps this vector to the target sequence space. This architecture is known as encoder-decoder.

The encoder-decoder architecture has gained popularity and variations to the initial formulation have been proposed lately, as attention mechanisms [12] or the use of non-recurrent networks. In particular, encoder-decoder CNNs have shown similar performance to the original RNN formulation in sequences-related tasks [13], but with a fraction of the training and processing time required by RNNs. Recently, the benefits of combining both type of networks in an encoder-decoder structure have been exposed [14], where CNNs process longer sequences in less time and extract local spatial features, while RNNs find temporal relationships among input signals.

A particular instance of encoder-decoder networks is the Autoencoder (AE) [15], which is an unsupervised NN capable of learning efficient representations of the inputs by trying to reconstruct them. Formally, the AE maps an input vector \( y \in \mathbb{R}^m \) to its latent representation \( z \in \mathbb{R}^m \) where typically \( m < n \). This mapping is done by an encoder \( f_{\text{enc}} \), parameterized by \( \theta_e \), as \( z = f_{\text{enc}}(y) \). The latent vector \( z \) can be mapped back to the input space by the decoder \( f_{\text{dec}} \), parameterized by \( \theta_d \), as \( \hat{y} = f_{\text{dec}}(z) \). Both the encoder and the decoder can be any type of NN depending on the application. Nonetheless, in all cases the same optimization problem is solved, which amounts to minimizing the difference between input and decoded vectors.

### III. DENOISING AND VOLTAGE ESTIMATION USING ENCODER-DECODER NEURAL NETWORKS
#### A. DENOISING PROBLEM FORMULATION
Denoising based on autoencoders started with the appearance of the Denoising Autoencoder (DAE) [16], which is fed with inputs corrupted by noise and is forced to reconstruct the clean version of the inputs. DAEs have been successfully used in many applications, particularly image processing [17]. However, DAEs work under the assumption that the clean version of the inputs is available and the noise nature is known, which in general is not true. Blind denoising autoencoders (BDAEs) [18] address the aforementioned issue. BDAEs are able to denoise inputs by restricting the latent space, without knowledge about the noise nature nor access to the clean inputs. In [19], a BDAE with a dual encoder structure (DBDAE), one RNN encoder to exploit temporal relationships among inputs and one CNN encoder to exploit spatial correlations, was designed to blind-denoise time series data from dynamical systems. In this brief, we propose to use this DBDAE structure to filter sensory data in MMCs. Figure 2 illustrates the DBDAE and its building blocks are explained in the following.

Consider a dynamical system \( \Sigma \) and a sampling process generating a sequence of noisy measurements \( \tilde{y} \), a T-depth window \( \tilde{Y} \) can be constructed, so that

\[
\tilde{y}(n) = [\tilde{y}(n-T+1), \tilde{y}(n-T+2), \ldots, \tilde{y}(n)] \in \mathbb{R}^N \times T. \tag{1}
\]

In this setting, \( \tilde{Y}_k(n) \in \mathbb{R}^N \) is a column vector containing the \( N \) measurements at time instant \( n-T+k \), and \( Y_k(n) \in \mathbb{R}^l \) is a row vector containing the last \( T \) samples from sensor \( k \).

1) RECURRENT ENCODER
It generates its latent representation \( h_R^{Lr}(j) \), which is a local \( Q_L\)-dimensional finite-length sequence of length \( T \), by sequentially processing \( \tilde{Y}(n) \), namely,

\[
h_R^{Lr}(j) = f_{\text{RE}}(\tilde{Y}_j(n), h_R^{Lr}(j-1)), \tag{2}
\]

where, \( f_{\text{RE}} \) represents the recurrent network, with \( L_R \) layers and \( Q \) neurons in each layer \( l \). Considering Gated Recurrent Units (GRU) [11], the operations at each layer \( l \) are given by

\[
z_l(j) = \sigma(W_{zl}^p l(j) + U_{zl}^p h_R^{Lr}(j) + b_z^l) \tag{3}
\]

\[
r_l(j) = \sigma(W_{rj}^p l(j) + U_{rj}^p h_R^{Lr}(j) + b_r^l) \tag{3}
\]

\[
n_l(j) = \text{tanh}(W_{nl}^p l(j) + r_l(j) \circ (U_{nl} h_R^{Lr}(j) + b_n^l)) \tag{3}
\]

\[
h_R^{Lr}(j) = (1 - z_l(j)) \circ n_l(j) + z_l(j) \circ h_R^{Lr}(j) \tag{3}
\]

where \( h_R^{Lr}(j) \in \mathbb{R}^{Q_L} \) is the hidden state of layer \( l \) at iteration \( j \), \( r_l, z_l \) and \( n_l \) are \( Q \)-dimensional finite-length sequences representing the value of reset, update and new gates of layer \( l \) respectively, \( \sigma \) and \( \text{tanh} \) are the sigmoid and the hyperbolic tangent activation functions, \( \circ \) is the Hadamard product and \( p_l(j) \) is the input of layer \( l \) at iteration \( j \). For \( l = 1 \), the input, \( p_1(j) \), is given by \( \tilde{Y}_j(n) \), while for \( l > 1 \) is given by \( h_R^{Lr-1}(j) \).

The latent representation of this encoder is \( h_R^{Lr}(T) \in \mathbb{R}^{Q_L} \), which is the last hidden state of the last layer. The final output of the encoder \( h_R(n) \in \mathbb{R}^Q \) is obtained by projecting \( h_R^{Lr}(T) \) with a linear layer to obtain the dimension of the latent space \( Q \). Note that \( h_R \) is an infinite-length sequence with one value per dataset \( \tilde{Y}(n) \), hence is indexed by \( n \).

2) CONVOLUTIONAL ENCODER
It generates a latent representation in parallel as

\[
h_C^{Lc}(n) = f_C(\tilde{Y}(n)), \tag{4}
\]
where \( f_c \) represents the convolutional encoder network with \( L_C \) layers. The output of each layer \( i \) is composed by \( K_i \) concatenated output channels \( C_{\text{out}}^i(n) \in \mathbb{R}^2 \), with \( k \in \{1; K_i\} \), given \( I \) input channels. Each of the resulting output channels for layer \( l \) at time step \( n \) is given by:

\[
C_{\text{out}}^l(n) = b_{\text{L}}^l + \sum_{i=1}^{l} W_{\text{L}}^l(i) \star p_{\text{L}}^i(n),
\]

where \( b_{\text{L}}^l \) is a bias vector for channel \( k \) at layer \( l \), \( W_{\text{L}}^l(i) \) is the \( i \)-th channel of the \( l \)-channel kernel corresponding to output channel \( k \), \( p_{\text{L}}^i(n) \) is the \( i \)-th channel of the input of layer \( l \) and \( \star \) is the cross-correlation operator. For \( l = 1 \), \( p_{\text{L}}^1(n) \) is given by \( \tilde{Y}(n) \), while for \( l > 1 \) \( p_{\text{L}}^1(n) \) is given by \( C_{\text{out}}^{l-1} \).

The convolutional encoder produces its latent representation \( \mathbf{h}_{\text{L}}(n) \in \mathbb{R}^{L_C \times K_L} \) by concatenating all the output channels of layer \( L_C \). The final output of the encoder \( \mathbf{h}_{\text{L}}(n) \) is obtained by projecting \( \mathbf{h}_{\text{L}}(n) \) with two linear layers to match the requirements of the latent space dimension \( Q \).

The final latent representation of the input at time step \( n \) is just the sum of the latent representations of both encoders.

\[
\mathbf{h}_E(n) = \mathbf{h}_R(n) + \mathbf{h}_C(n)
\]

3) RECURRENT DECODER
Is a RNN, \( f_{RD} \), with \( L_D \) layers and \( Q_l \) neurons in each layer \( l \). Its hidden state, \( \mathbf{h}_{\text{L}}^{DL}(n) \), is a \( Q_l \)-dimensional finite-length sequence of length \( T \), given by

\[
\mathbf{h}_{\text{L}}^{DL}(j) = f_{RD}(p_j(n), \mathbf{h}_{\text{L}}^{DL}(j - 1)),
\]

where \( p_j(n) \in \mathbb{R}^N \) is the input for each decoding step of the network, which is either the delayed target \( \tilde{Y}_{j-1}(n) \) or the previous network estimate \( \hat{Y}_{j-1}(n) \), following the scheduled sampling technique [20]. The network estimates, \( \hat{Y}_{j}(n) \) are calculated by projecting \( \mathbf{h}_{\text{L}}^{DL}(j) \) with a linear layer.

The final output of the DBDAE is \( \hat{Y}_T(n) \in \mathbb{R}^N \), which contains the denoised measurements of \( \Sigma \) at time step \( n \).

It should be noted that the decoder starts once \( \mathbf{h}_E(n) \) is available, which involves \( T \) previous iterations of the encoder.

4) CONSTRAINING THE LATENT SPACE
To stop learning, the simple heuristic proposed in [19] is used. The heuristic consists in using principal component analysis (PCA) to transform the latent space of the DBDAE during training and to stop the learning process when the derivative of the reconstruction error is maximum. This heuristic was successfully applied in [19] to a variety of systems and will be used in this work.

B. VOLTAGE ESTIMATION PROBLEM FORMULATION
Modeling of dynamical systems using encoder-decoder RNNs as predictors in input-output form has recently gained popularity [21], [22]. However, when systems exhibit a combination of slow and fast dynamics, models increase in complexity as the input sequences become longer and longer. In this setting, we propose the use of a CNN as encoder to process long sequences aiming to capture the slow dynamics, and a RNN as decoder taking a short sequence as input aiming to capture the fast dynamics. Figure 3 illustrates the structure of the encoder-decoder network proposed as voltage estimator, the building blocks are explained in the following.

Given an \( N \)-dimensional measurement sequence \( X \) sampled from a dynamical system \( \Sigma \), a \( T_C \)-depth window \( X_C \) and a \( T_R \)-depth window \( X_R \) are constructed, with \( T_C \gg T_R \).

1) CONVOLUTIONAL ENCODER
It generates at each time step \( n \) a latent representation of \( X_C(n) \) using the same structure as the convolutional encoder used in the DBDAE, with \( \mathbf{h}_{\text{C}}(n) \in \mathbb{R}^Q \) as the latent representation of its inputs \( X_C(n) \in \mathbb{R}^{N \times T_C} \).

2) RECURRENT DECODER
It uses as initial hidden state the latent representation \( \mathbf{h}_{\text{C}}(n) \) and as input the data matrix \( X_C(n) \in \mathbb{R}^{N \times T_C} \). The decoder is a GRU-based network with \( L_R \) layers and \( Q_l \) neurons in layer \( l \), that generates its latent representation \( \mathbf{h}_{\text{R}}^{DL}(j) \), a finite-length sequence of length \( T_R \), as

\[
\mathbf{h}_{\text{R}}(j) = f_{RD}(X_R(j), \mathbf{h}_{\text{R}}^{DL}(j - 1)),
\]

After \( T_R \) iterations, \( \mathbf{h}_{\text{R}}^{DL}(T_R) \in \mathbb{R}^{O_{tR}} \) is available, which contains information about all the input sequences of the encoder and the decoder and is projected by a linear layer to produce \( \hat{y}(n) \in \mathbb{R}^D \), the final output.

IV. EXPERIMENTAL RESULTS
A. EXPERIMENTAL SETUP
As case study, the three-phase MMC shown in Figure 4 is used. The experimental setup includes an OPAL-RT OP4510 system acting as controller (Intel Xeon E3 v5 CPU and two Kintex-7 FPGAs) and a passive RL load with \( L_o = 1 \) mH and \( R_o = 28 \) Ω. To obtain experimental data for training and testing of the denoiser and estimator, the MMC was configured according to the operating conditions summarized in Table 1.

The MMC was operated in open-loop, with no control for balancing the voltage of the capacitors, nor the output voltages and currents. Open-loop operation was chosen since it exhibits richer dynamic response, which favors the generalization of the DNNs. In such operating conditions, capacitor

FIGURE 3. Estimator structure consisting in a CNN as encoder and a RNN as decoder. The CNN processes a sequence of length \( T_C \) to obtain the latent representation \( \mathbf{h}_C \), while the RNN combines \( \mathbf{h}_C \) with a sequence of length \( T_R \), with \( T_C \gg T_R \), to estimate the voltages of the capacitors.
voltages may diverge over time, compromising the operation of the MMC [3]. To overcome this situation, a phase-shifted pulse-width-modulation (PS-PWM) was used since it distributes the SM power evenly, keeping the capacitor voltages balanced [23].

Several experimental datasets were obtained both under steady-state operation and during transients due to step load variations. The datasets contain output voltages, capacitor voltages, stack currents and switching signals, all recorded synchronously. Figure 5 illustrates a portion of a dataset in steady-state operation. A considerable amount of noise can be observed in the measured variables.

### B. DENOISING IN AN MMC

The noise in the measurement data-buses of the MMC is mainly due to the high EMI generated by the huge number of power semiconductors switching at high voltage and high speed. Reducing noise in MMC signals is important for a proper operation of the equipment.

1) TRAINING

To train the DBDAE to denoise the MMC signals, 40 seconds of experimental data sampled at $T_s = 75 \times 10^{-6}$ s were used. The objective of the network is to denoise all the important signals in each phase, that is, all the capacitor voltages, both stack currents and the output voltage. Therefore three DBDAEs with 11 inputs each are trained.

A latent space dimension of 44 and a window length of 20 were selected for each network. As for the loss function, the MSE (mean squared error) between the reconstructed signals and the original inputs was aimed to be minimized until the stopping criterion, introduced in Section III, was reached. Adam optimizer with a learning rate of $1 \times 10^{-4}$ was used. Implementation details of the DBDAE are given in Table 2.

2) RESULTS

Since we are dealing with real data, the clean versions of the signals are not available, hence is not possible to calculate any metric comparing the results with the ground-truth. However, we can visually inspect the results.

Figure 6 shows the denoising results for some of the MMC signals. It can be seen that the DBDAE removes noise and switching effects from the on-off operation of power semiconductors, which are not required for control purposes and can deteriorate closed-loop performance. The denoised
signals show negligible error nor lag and follows the fast dynamic of the different signals, unlike a low pass filter.

C. CAPACITOR VOLTAGE ESTIMATION IN AN MMC

1) TRAINING

The encoder-decoder network was trained to estimate the 4 capacitor voltages of the same stack. The data used to train the network involves 9.6 seconds of 33 signals: the input signals of each SMs (24 signals), the current of each stack (6 signals) and the output voltage per phase (3 signals). In this case, no capacitor voltage signals were used as inputs (only as targets during training), as the objective is to completely remove voltage sensors in the SMs. The dataset was acquired running experiments that capture both, slow and fast dynamics at a sampling time of $25 \times 10^{-6}$ s. Adam optimizer was used with a learning rate of $1 \times 10^{-3}$ and as for the loss function, MSE was used. Details of the encoder-decoder network are given in Table 3.

2) RESULTS

Results obtained are presented in Figure 7 and Figure 8. Figure 7 illustrates estimated and measured values of the voltages during steady state and transient operation. The estimated time series follows closely the behavior of the voltages in the entire experiment, including the oscillation at the fundamental frequency (50 Hz), the second harmonic, and the slower convergence of the mean value of the voltages upon a load change. Transient operation is highlighted in the right column of Figure 7, where estimation during a step change in the load at time 1.57 s is shown. There is no delay or appreciable mismatch in the estimates, even during changes in the load.

Figure 8 shows a zoomed image where it can be appreciated how the encoder-decoder network also estimates the switching effects in the capacitor voltages. These switching signals have four times the carrier frequency (4 kHz) due to the PS-PWM. It should be noted that although this switching effect in the voltage signal can be harmful for the controller, as it was mentioned in the motivation of the DBDAE-based application, it is useful to validate the correct internal operation of the converter. The estimator is capable of reproducing the switching effect, which if not desired can be filtered.

Table 4 summarizes the numerical errors between the estimated voltages and the real targets, in terms of the root mean square error (RMSE). It can be seen that all the estimated signals present a RMSE below 0.54 volts, which given the operating point around 100 volts, is negligible.

The proposed encoder-decoder network has several advantages. First, no pre-processed data is required to estimate the output, this becomes important when running at real-time, given the high frequency at which the MMC operates.

| Capacitor 1 (Vcap1) | 0.403 |
|---------------------|-------|
| Capacitor 2 (Vcap2) | 0.534 |
| Capacitor 3 (Vcap3) | 0.427 |
| Capacitor 4 (Vcap4) | 0.427 |

TABLE 3. Details of the encoder-decoder network.

| Layer | Convolutional Encoder | Recurrent Decoder |
|-------|-----------------------|-------------------|
| Layers | 3                     | 2                 |
| Channels per layer | 33, 66, 132          |                   |
| Number of parameters | 31,928               | 11,430            |
| Sequence length    | 200                   | 30                |

TABLE 4. Estimation error in terms of RMSE.
Second, the convolutional encoder is an advantage over a recurrent encoder for capturing the slow dynamics, avoiding using long sequence recursions in a RNN. Finally, although the proposed network has many parameters, this architecture allows parallel computation, facilitating its implementation.

V. CONCLUDING REMARKS
The use of neural networks is proposed to support the operation of modular multilevel converters (MMCs). Particularly, a dual-blind denoising autoencoder (DBDAE) is proposed to filter sensor measurements, and an encoder-decoder network is proposed to estimate the voltages of each sub-module. A three-phase MMC with 4 sub-modules per stack was used as case study and setup to obtain datasets of measurements corrupted by electromagnetic interference (EMI) during both steady state and transient operation.

Experimental results show that the DBDAE removed practically all the noise and power switching artifacts in the signals with a negligible lag. Regarding the estimator, experimental results show a high performance and accuracy under steady state and transient operation, estimating the wide spectrum of the capacitor voltage behavior, from the fundamental oscillation of 50 Hz to the switching ripple introduced by the commutations of power semiconductors.
The proposed methods potentially increase the EMI-tolerance and reliability of the MMC, and allow the elimination of voltage sensors in each sub-module, thus reducing the overall cost and communication requirements.

As future work, the filter and estimator should be implemented in the FPGA of the controller of the MMC to test their performance when the MMC operates in closed-loop under aggressive dynamic operation and failures.

REFERENCES

[1] S. Kouro, M. Malinowski, K. Gopakumar, J. Pou, L. G. Franquelo, B. Wu, J. Rodriguez, and J. I. Leon, “Recent advances and industrial applications of multilevel converters,” IEEE Trans. Ind. Electron., vol. 57, no. 8, pp. 2553–2580, Aug. 2010.

[2] H. Abu-Rub, J. Holtz, J. Rodriguez, and G. Baoming, “Medium-voltage multilevel converters—State of the art, challenges, and requirements in industrial applications,” IEEE Trans. Ind. Electron., vol. 57, no. 8, pp. 2581–2596, Dec. 2010.

[3] S. Debnath, J. Qin, B. Bahrami, M. Saedelifard, and P. Barbosa, “Operation, control, and applications of the modular multilevel converter: A review,” IEEE Trans. Power Electron., vol. 30, no. 1, pp. 37–53, Jan. 2015.

[4] M. A. Perez, S. Bernet, J. Rodriguez, S. Kouro, and R. Lizana, “Circuit topologies, modeling, control schemes, and applications of modular multilevel converters,” IEEE Trans. Power Electron., vol. 30, no. 1, pp. 4–17, Jan. 2015.

[5] R. Zhu, N. Lin, V. Dinavahi, and G. Liang, “An accurate and fast method for conducted EMI modeling and simulation of MMC-based HVdc converter station,” IEEE Trans. Power Electron., vol. 35, no. 5, pp. 4689–4702, May 2020.

[6] G. Konstantinou, H. R. Wickramasinghe, C. D. Townsend, S. Ceballos, and J. Pou, “Estimation methods and sensor reduction in modular multilevel converters: A review,” in Proc. 8th Int. Conf. Power Energy Syst. (ICPES), Dec. 2018, pp. 23–28.

[7] O. S. H. M. Ahubasha, M. S. A. Dahidah, S. M. Gadoue, and D. J. Atkinson, “Submodule voltage estimation scheme in modular multilevel converters with reduced voltage sensors based on Kalman filter approach,” IEEE Trans. Ind. Electron., vol. 65, no. 9, pp. 7025–7035, Sep. 2018.

[8] M. Abdelsalam, M. Marei, S. Tennakoon, and A. Griffiths, “Capacitor voltage balancing strategy based on sub-module capacitor voltage estimation for modular multilevel converters,” IEEE J. Emerg. Syst. Power Electron., vol. 2, no. 1, pp. 65–73, Mar. 2016.

[9] A. Dekka, B. Wu, V. Yaramasu, R. Lizana Fuentes, and N. R. Zargari, “Model predictive control of high-power modular multilevel Converters—An overview,” IEEE J. Emerg. Sel. Topics Power Electron., vol. 7, no. 1, pp. 168–183, Mar. 2019.

[10] G. Konstantinou, J. Zhang, S. Ceballos, J. Pou, and V. G. Agelidis, “Comparison and evaluation of sub-module configurations in modular multilevel converters,” in Proc. IEEE 11th Int. Conf. Power Electron. Drive Syst., Jun. 2015, pp. 958–963.

[11] K. Cho, B. van Merrienboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, “Learning phrase representations using RNN Encoder–Decoder for statistical machine translation,” in Proc. Conf. Empirical Methods Natural Lang. Process. (EMNLP), 2014, pp. 1724–1734.

[12] D. Bahdanau, K. Cho, and Y. Bengio, “Neural machine translation by jointly learning to align and translate,” 2014, arXiv:1409.0473. [Online]. Available: http://arxiv.org/abs/1409.0473

[13] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, “Attention is all you need,” 2017, arXiv:1706.03762. [Online]. Available: http://arxiv.org/abs/1706.03762

[14] Z. Zuo, B. Shuai, G. Wang, X. Liu, X. Wang, B. Wang, and Y. Chen, “Learning contextual dependence with convolutional hierarchical recurrent neural networks,” IEEE Trans. Image Process., vol. 25, no. 7, pp. 2983–2996, Jul. 2016.

[15] G. E. Hinton, “Reducing the dimensionality of data with neural networks,” Science, vol. 313, no. 5786, pp. 504–507, Jul. 2006.

[16] P. Vincent, H. Larochelle, Y. Bengio, and P.-A. Manzagol, “Extracting and composing robust features with denoising autoencoders,” in Proc. 25th Int. Conf. Mach. Learn., 2008, pp. 1096–1103.

[17] A. Creswell and A. A. Bharath, “Denoising adversarial autoencoders,” IEEE Trans. Neural Netw. Learn. Syst., vol. 30, no. 4, pp. 968–984, Apr. 2019.

[18] A. Majumdar, “Blind denoising autoencoder,” IEEE Trans. Neural Netw. Learn. Syst., vol. 30, no. 1, pp. 312–317, Jan. 2019.

[19] S. Langarica and F. Nñez, “Dual blind denoising autoencoders for industrial process data filtering,” 2020, arXiv:2004.06806. [Online]. Available: http://arxiv.org/abs/2004.06806

[20] S. Bengio, O. Vinyals, N. Jaitly, and N. Shazeer, “Scheduled sampling for sequence prediction with recurrent neural networks,” in Proc. 28th Int. Conf. Neural Inf. Process. Syst., vol. 1, 2015, pp. 1171–1179.

[21] F. Nunez, S. Langarica, P. Diaz, M. Torres, and J. C. Salas, “Neural network-based model predictive control of a paste thickener over an industrial Internet platform,” IEEE Trans. Ind. Informat., vol. 16, no. 4, pp. 2859–2867, Apr. 2020.

[22] S. Langarica and F. Nñez, “Neuroevolutionary control of industrial processes through mapping elites,” IEEE Trans. Ind. Informat., early access, Aug. 27, 2020, doi: 10.1109/TII.2020.3019846.

[23] Y. Li, Y. Wang, and B. Q. Li, “Generalized theory of phase-shifted carrier PWM for cascaded H-Bridge converters and modular multilevel converters,” IEEE J. Emerg. Sel. Topics Power Electron., vol. 4, no. 2, pp. 589–605, Jun. 2016.

SÁUL LANGARICA (Member, IEEE) was born in Guatemala City, Guatemala, in 1994. He received the B.Sc. and M.Sc. degrees in electrical engineering from the Pontificia Universidad Católica de Chile, in 2018 and 2019, respectively, where he is currently pursuing the Ph.D. degree in electrical engineering. His professional interests include artificial intelligence, intelligent control, and supervision systems for industrial processes.

GERMÁN PIZARRO was born in Santiago, Chile. He received the B.Sc. degree in electrical engineering from the Pontificia Universidad Católica de Chile, Santiago, Chile, in 2019, where he is currently pursuing the M.Sc. degree in electrical engineering. His research interests include control theory, distributed control, artificial intelligence, model predictive control, and coordination of self-driving cars.

PABLO M. POBLETE (Graduate Student Member, IEEE) was born in Rancagua, Chile. He received the B.Sc. degree in electrical engineering from the Pontificia Universidad Católica de Chile, Santiago, in 2018, where he is currently pursuing the Ph.D. degree in electrical engineering. His research interests include control theory, distributed control, artificial intelligence, model predictive control, and coordination of self-driving cars.
FELIPE RADRIGÁN (Graduate Student Member, IEEE) was born in Santiago, Chile. He received the B.Sc. degree in electrical engineering from the Pontificia Universidad Católica de Chile, Santiago, in 2016, where he is currently pursuing the M.Sc. degree. His research interests include power electronics, multilevel converters, and electrical drives.

JAVIER PEREDA (Member, IEEE) received the B.Sc. (Eng.) degree (Hons.) in electrical engineering and the M.Sc. and Ph.D. degrees in electrical engineering from the Pontifical Catholic University of Chile, Santiago, Chile, in 2009 and 2013, respectively. He has been an Assistant Professor with the Department of Electrical Engineering, Pontificia Universidad Católica de Chile, since 2013. From 2014 to 2016, he was an Associate Researcher with the Control and Power Group, Department of Electrical and Electronic Engineering, Imperial College London. He is currently an Associate Researcher with the Solar Energy Research Center, and the UC Energy Research Center, Chile. He is also a Principal Investigator of the Electric Vehicle Laboratory and the Power and Energy Conversion Laboratory (PEClab), Pontificia Universidad Católica de Chile. His research interests include power electronics and control applied to electric vehicles, ac and dc electric networks, microgrids, renewable energy, multilevel converters, industrial applications, and motor drives.

JOSE RODRIGUEZ (Life Fellow, IEEE) received the Engineering degree in electrical engineering from the Universidad Tecnica Federico Santa Maria, Valparaiso, Chile, in 1977, and the Dr.-Ing. degree in electrical engineering from the University of Erlangen, Erlangen, Germany, in 1985. He has been with the Department of Electronics Engineering, Universidad Federico Santa Maria, since 1977, where he was full a Professor and President. Since 2015, he was the President and since 2019, he has been a Full Professor with Universidad Andres Bello, Santiago, Chile. He has coauthored two books, several book chapters, and more than 400 journal and conference articles. His main research interests include multilevel inverters, new converter topologies, control of power converters, and adjustable-speed drives. He has received a number of best paper awards from journals from the IEEE. He is a member of the Chilean Academy of Engineering. In 2014, he received the National Award of Applied Sciences and Technology from the government of Chile. In 2015, he received the Eugene Mittelmann Award from the Industrial Electronics Society of the IEEE. From 2014 to 2019, he has been included in the list of Highly Cited Researchers published by Web of Science.

FELIPE NÚÑEZ (Member, IEEE) was born in Santiago, Chile. He received the B.Sc. and M.Sc. degrees in electrical engineering from the Pontificia Universidad Católica de Chile in 2007 and 2008, and the Ph.D. degree in electrical and computer engineering from the University of California at Santa Barbara, Santa Barbara, in 2014. In 2015, he was with the UCSB/MIT/Caltech Institute for Collaborative Biotechnologies. In 2016, he joined the Department of Electrical Engineering, Pontificia Universidad Católica de Chile, where he is currently an Assistant Professor. His research interests include control theory, cyber-physical systems, control of networks, sensor and computer networks, industrial automation, intelligent control systems, and mineral processing.