

Learnable Regularization via Padé Activation Units for Flexible Model-Based Beamforming

Christopher Khan

Department of Biomedical Engineering
Vanderbilt University
Nashville, TN, USA
christopher.m.khan@vanderbilt.edu

Ruud J. G. van Sloun

Department of Electrical Engineering
Eindhoven University of Technology
Eindhoven, Netherlands
r.j.g.v.sloun@tue.nl

Brett Byram

Department of Biomedical Engineering
Vanderbilt University
Nashville, TN, USA
brett.c.byram@vanderbilt.edu

Abstract—Aperture Domain Model Image REconstruction (ADMIRE) is a beamformer that suppresses sources of acoustic clutter by performing model-based fits of channel data. Each fit requires solving an ill-posed, inverse problem using regularization. Initially, these fits were performed with the elastic-net constraint, but we recently proposed using a deep neural network (DNN) for model fitting based on the learned iterative shrinkage thresholding algorithm (LISTA). This approach used a sigmoidal soft-thresholding activation function with a trainable threshold to allow more flexible regularization, but the range of learnable proximal operators of regularization priors is still limited. Therefore, in this work, we propose using fully trainable activation functions to allow for a wider range of operators to be learned. Moreover, using simulated data, we demonstrate the ability of this approach to provide further improvements in ultrasound image quality.

Index Terms—ADMIRE, Beamforming, Deep Learning

I. INTRODUCTION

APERTURE Domain Model Image REconstruction (ADMIRE) is an adaptive beamforming method that improves ultrasound image quality by suppressing sources of acoustic clutter, and it can do so in real time via an open-source, GPU-based implementation (<https://github.com/VU-BEAM-Lab/ADMIRE>) [1]–[8]. It involves constructing model matrices of the aperture domain data, localized in time and frequency for different scattering locations, and then fitting these models to the received aperture domain frequency data. Once fit, only predictors within a specified region of interest (ROI) are used to reconstruct the decluttered aperture domain frequency data. However, performing the model fits of ADMIRE requires solving an ill-posed, inverse problem for which linear regression with elastic-net regularization is used to obtain a solution. This solution is one of infinitely many, and it is possible that some other solutions provide further improvements in image quality.

Therefore, to learn one of these other solutions, we previously proposed using a deep neural network (DNN) sparse encoder based on the learned iterative shrinkage thresholding algorithm (LISTA) [9] to perform the model fits of ADMIRE [10], [11]. This is shown in Fig. 2, and it allows for a data-driven approach to model fitting while still preserving the model-based intuition of ADMIRE. The network architecture is shown in Fig. 1, where the input to the network is $\mathbf{X}^T \mathbf{y}$

and the output is $\hat{\beta}$. In this case, $\mathbf{X}$ is the ADMIRE model matrix for a given fit, $\mathbf{y}$ is the corresponding aperture domain frequency data, and $\hat{\beta}$ is the predicted model coefficients. For the activation function, a sigmoidal soft-thresholding function with the equation shown in (1) was used, where z is the function input and γ is a trainable threshold that can be different for each layer.

$$h(z) = \frac{z}{1 + e^{-(|z| - \gamma)}} \quad (1)$$

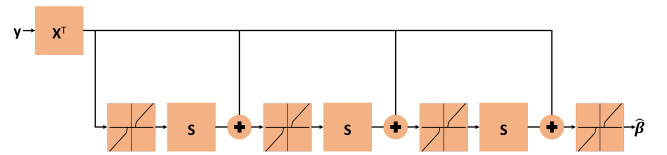


Fig. 1: Diagram of the deep neural network sparse encoder architecture for performing the model fits of ADMIRE. **S** represents a fully connected layer.

Although the aforementioned soft-thresholding activation function allows for more flexible regularization, it is still limited in terms of the range of proximal operators of regularization priors that can be learned. To address this, in this work, we propose replacing the soft-thresholding activation function at each layer of our DNN with a fully trainable activation function. Therefore, a wider range of proximal operators of regularization priors will be able to be learned, which in turn will allow for more flexible model-based beamforming and improved ultrasound image quality.

II. METHODS

To incorporate a fully trainable activation function, we utilized a Padé activation unit (PAU). As shown in (2), PAUs are activation functions that consist of a polynomial in the numerator and a polynomial in the denominator, and the polynomial coefficients are trainable weights. As a result, PAUs have been shown to be able to approximate a wide variety of functions [12].

$$h(z) = \frac{a_0 + a_1 z + a_2 z^2 + \dots + a_m z^m}{1 + b_1 z + b_2 z^2 + \dots + b_n z^n} \quad (2)$$

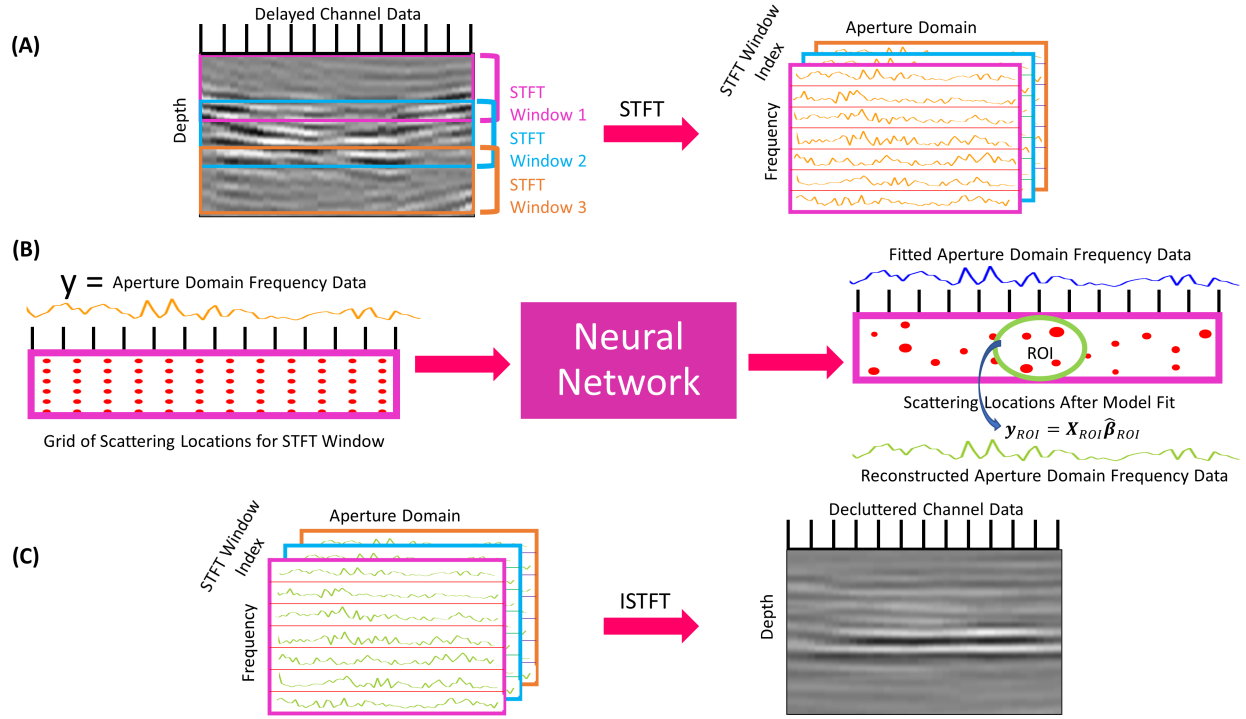


Fig. 2: Overview of ADMIRE using a neural network. (A) Obtain the time-delayed channel data and calculate the short-time Fourier transform (STFT) along the depth dimension for each channel. (B) Obtain the corresponding model matrix for each set of aperture domain frequency data that will be reconstructed in each STFT window in each beam, fit each model matrix to its corresponding set of aperture domain frequency data using a neural network (sizes of red points correspond to how much each scattering location contributes to the aperture domain frequency data), and reconstruct each set of aperture domain frequency data by only using the predictors that correspond to scattering locations that are within a region of interest (ROI). (C) Calculate the inverse short-time Fourier transform (ISTFT) of the reconstructed aperture domain frequency data in order to obtain the decluttered channel data. Note that the scattering locations are not restricted to the depth range of the STFT window. The grid of scattering locations illustrated in (B) corresponds to the first STFT window. For STFT windows that correspond to deeper depths, the scattering locations can also be located in shallower depths because these locations can contribute to off-axis scattering and multipath scattering that affect the aperture domain frequency data for the STFT window. Essentially, as the depths become deeper for subsequent STFT windows, the depth range for possible scattering locations also increases.

For the purposes of DNN training, a modified version of PAUs referred to as safe PAUs was utilized. The equation for one safe PAU is shown in (3). This version of the PAU was used in order to prevent obtaining a value of 0 in the denominator during training, which would cause undefined values. In addition, 5th and 4th order polynomials were used in the numerator and denominator, respectively. As a result, the same DNN architecture shown in Fig. 1 was used except for the replacement of the activation function with a safe PAU at each layer.

$$h(z) = \frac{a_0 + a_1 z + a_2 z^2 + \dots + a_m z^m}{1 + |b_1 z + b_2 z^2 + \dots + b_n z^n|} \quad (3)$$

To perform training, the Field II ultrasound simulation package [13], [14] was used to simulate 10 phantoms that each consisted of nine cysts of varying sizes. Each of these cysts were scaled to have a certain contrast ratio value, where the possible values were anechoic, -70 dB, -60 dB, -50 dB, -40 dB, -30 dB, -20 dB, -10 dB, and 0 dB. Reverberation

clutter simulated using a pseudononlinear approach [15], [16] and simulated off-axis clutter were also added to these 10 phantoms. Different signal-to-clutter ratios (SCRs) with possible values of -5 dB, 0 dB, 5 dB, 10 dB, 15 dB, and 20 dB were used to scale the off-axis clutter and reverberation clutter individually before they were added to the phantoms. Moreover, Gaussian noise yielding a signal-to-noise ratio (SNR) of 50 dB with respect to the combined clean and clutter data was added. In total, 60 unique phantoms were obtained from this process. After this, the short-time Fourier transform of each phantom was computed to obtain the different sets of aperture domain frequency data. In this case, due to the data being simulated, the aperture domain frequency data corresponding to the clean signal without clutter, the clutter signal, and the clean signal with clutter were obtained. The ADMIRE model matrices were also generated, and fourth-order blind identification independent component analysis (FOBI-ICA) [7], [8] was applied to reduce their sizes.

Training was performed in PyTorch [17] using the Adam op-

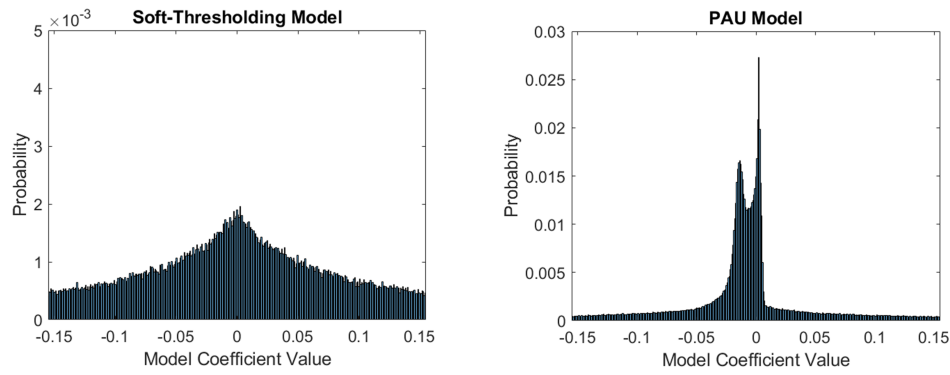


Fig. 3: Distributions of model coefficients predicted by the best-performing sigmoidal soft-thresholding and PAU DNNs. In this example, model coefficients were drawn from a uniform distribution and used with ADMIRE model matrices to create synthetic sets of aperture domain frequency data. Therefore, the distributions in this case primarily represent the regularization prior that each type of activation function imposes on the fitting process. Note that the y-axis is different between the histograms.

TABLE I
MEAN ERROR VALUES FOR IMAGE METRICS (N = 20 FOR MSE AND N = 180 FOR OTHER METRICS)

Activation Function Type	CR Absolute Error (dB)	CNR Absolute Error (dB)	GCNR Absolute Error	Normalized Envelope MSE	Overall Rank Out of 1200 Cases
PAU	17.29	1.79	0.033	0.003	1
Sigmoidal Soft-Thresholding	22.57	1.39	0.049	0.002	171

timizer [18] and three different training schemes. For scheme 1, the model coefficients predicted by the network for a given training iteration were taken and used along with the corresponding ADMIRE model matrices to compute the predicted clutter signal, and a mean-squared error (MSE) loss function was then computed between the predicted signal and the ground-truth clutter signal. Note that, for training, the clutter signal included the added Gaussian noise. For scheme 2, the same operations were performed as scheme 1 except that they were done for the clean signal without clutter, and for scheme 3, the operations were done for both signals resulting in two MSE loss functions. For the DNN architecture, the newly proposed PAU DNN was used as well as the previous DNN that used sigmoidal soft-thresholding. Three fully connected layers were used in both cases. For each architecture and training scheme combination, 100 models were trained with different weight initializations that were achieved by varying the seed for the pseudorandom number generator in PyTorch. This resulted in 600 total models being trained.

Once trained, each model was used to perform the model fits of ADMIRE for 20 cyst phantom data sets. 10 phantoms had 0 dB SCR clutter added to them, and the other 10 had 10 dB SCR clutter added. In addition, all phantoms, with respect to the combined clean and clutter data, had 50 dB SNR Gaussian noise added to them. Once the fits were performed, the decluttered aperture domain frequency data for each phantom was reconstructed by either directly using the ROI predictors or by first computing the clutter signal and then subtracting it from the aperture domain data. This provided two processing cases for each model for a total of 1200 processing cases.

For the cysts of the processed phantoms for each processing case, metrics including contrast ratio (CR), contrast-

to-noise ratio (CNR), and generalized contrast-to-noise ratio (GCNR) were computed. Normalized envelope mean-square error (MSE) was also computed between each processed frame of data and the corresponding ground-truth clean frame of data. Thus, for each processing case, 180 CR, CNR, and GCNR values were obtained (180 total cysts across the phantoms) along with 20 MSE values (20 total phantoms). These metric values were then used to rank the processing cases. This was done by taking the values for a given metric for a given processing case and then computing the mean absolute error (MAE) between those values and the values for the ground-truth clean data. For MSE, the average of the 20 MSE values was taken. Doing so provided four error values for the four metrics for each processing case. For a given metric, the error values across the processing cases were ranked based on lowest error. This resulted in each processing case having four ranks, and the average of these ranks was computed to provide an overall rank to a given processing case.

In addition, to analyze the proximal operators of regularization priors that were learned by the best-performing sigmoidal soft-thresholding DNN and the best-performing PAU DNN, a separate experiment was performed in which model coefficients were generated from a uniform distribution. These coefficients were used along with ADMIRE model matrices to generate aperture domain frequency data rather than simulating aperture domain frequency data using Field II. Model fits were then performed using the two DNNs, and histograms of the predicted model coefficients were created. The purpose of doing this was that due to the coefficients being drawn from a uniform distribution, the histogram of the predicted coefficients should primarily reflect the regularization prior that is imposed on the coefficients.

III. RESULTS

Table I displays the mean error values for the best-performing PAU and sigmoidal soft-thresholding DNN cases. The top performing PAU DNN case was one that was trained with training scheme 2 and that reconstructed the aperture domain frequency data by directly utilizing ROI predictors. The top performing sigmoidal soft-thresholding DNN case was trained with training scheme 1 and reconstructed the aperture domain frequency data by directly utilizing ROI predictors. Fig. 4 shows an example phantom processed with these two cases. In addition, Fig. 3 shows the model coefficient distributions obtained for the best-performing sigmoidal soft-thresholding and PAU DNN models when performing the experiment where model fits were performed with synthetic aperture domain frequency data that was generated from model coefficients drawn from a uniform distribution. As can be seen, it appears as though the sigmoidal soft-thresholding DNN imposes a regularization prior that resembles a Laplacian distribution, as expected, while the PAU DNN imposes a different regularization prior with a bimodal distribution.

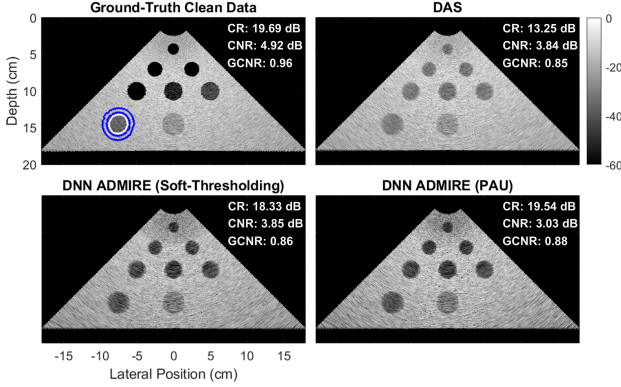


Fig. 4: Example phantom with added 0 dB SCR clutter and 50 dB Gaussian noise processed with the best-performing sigmoidal soft-thresholding and PAU DNN cases. The ground-truth clean data and delay-and-sum (DAS) are also shown as well as example image metrics calculated for one cyst using the displayed cyst (white) and background (blue) masks.

IV. DISCUSSION

When ranking the processing cases based on image metric error, the top 170 ranked cases were all DNNs that used PAUs instead of sigmoidal soft-thresholding for the activation function. This demonstrates that using fully trainable activation functions can be used to achieve further improvements in ultrasound image quality. Moreover, Fig. 3 exhibits that these improvements are obtained due to the ability of PAUs to learn a wider range of proximal operators of regularization priors and therefore allow for more flexible model-based beamforming.

V. CONCLUSION

In this work, we developed a framework for performing flexible model-based beamforming through the incorporation of fully trainable activation functions, and we demonstrated

the ability of this framework to achieve further improvements in ultrasound image quality.

VI. ACKNOWLEDGMENT

The authors would like to thank the staff of the AC-CRE computing resource. This work was supported by NIH grants R01HL156034, R01EB020040, and S10OD016216-01, NAVSEA grant N0002419C4302, and NSF award IIS-1750994.

REFERENCES

- [1] B. Byram and M. Jakovljevic, "Ultrasonic multipath and beamforming clutter reduction: A chirp model approach," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 61, no. 3, pp. 428–440, 2014.
- [2] B. Byram, K. Dei, J. Tierney, and D. Dumont, "A model and regularization scheme for ultrasonic beamforming clutter reduction," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 62, no. 11, pp. 1913–1927, 2015.
- [3] K. Dei and B. Byram, "The impact of model-based clutter suppression on cluttered, aberrated wavefronts," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 64, no. 10, pp. 1450–1464, 2017.
- [4] C. Khan, K. Dei, S. Schlunk, K. Ozgun, and B. Byram, "A Real-Time, GPU-Based Implementation of Aperture Domain Model Image REconstruction," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 68, no. 6, pp. 2101–2116, Feb. 2021.
- [5] C. Khan and B. Byram, "GENRE (GPU Elastic-Net REgression): A CUDA-accelerated package for massively parallel linear regression with elastic-net regularization," *J. Open Source Softw.*, vol. 5, no. 54, p. 2664, 2020.
- [6] K. Dei, S. Schlunk, and B. Byram, "Computationally efficient implementation of aperture domain model image reconstruction," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 66, no. 10, pp. 1546–1559, Oct. 2019.
- [7] J. F. Cardoso, "Source separation using higher order moments," in *Proc. Int. Conf. Acoust., Speech, Signal Process.*, 1989, pp. 2109–2112.
- [8] J. Shlens, "A tutorial on independent component analysis," 2014, *arXiv:1404.2986*. [Online]. Available: <http://arxiv.org/abs/1404.2986>
- [9] K. Gregor and Y. LeCun, "Learning fast approximations of sparse coding," in *Proceedings of the 27th international conference on international conference on machine learning*, 2010, pp. 399–406.
- [10] C. Khan, R. J. G. van Sloun, and B. Byram, "Performing aperture domain model image reconstruction (ADMIRE) using a deep neural network sparse encoder," 2021 IEEE International Ultrasonics Symposium (IUS), 2021, pp. 1–4.
- [11] C. Khan, R. J. G. van Sloun, and B. Byram, "Unfolding model-based beamforming for high quality ultrasound imaging," 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2022, pp. 8682–8686.
- [12] A. Molina, P. Schramowski, and K. Kersting, "Padé Activation Units: End-to-end Learning of Flexible Activation Functions in Deep Networks," *arXiv preprint arXiv:1907.06732*, 2019.
- [13] J. A. Jensen, "FIELD: A program for simulating ultrasound systems," in *Proc. 10th Nordic-Baltic Conf. Biomed. Imag.*, 1996, vol. 4, no. 1, pp. 351–353.
- [14] J. A. Jensen and N. B. Svendsen, "Calculation of pressure fields from arbitrarily shaped, apodized, and excited ultrasound transducers," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 39, no. 2, pp. 262–267, Mar. 1992.
- [15] B. Byram and J. Shu, "Pseudononlinear ultrasound simulation approach for reverberation clutter," *Journal of Medical Imaging*, 3.4 (2016): 046005.
- [16] E. Vienneau, K. Ozgun, and B. Byram, "Spatiotemporal Coherence to Quantify Sources of Image Degradation in Ultrasonic Imaging," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 69, no. 4, pp. 1337–1352, 2022.
- [17] A. Paszke et al., "Pytorch: An imperative style, high-performance deep learning library," in *Advances in Neural Information Processing Systems* 32, H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alch'e-Buc, E. Fox, and R. Garnett, Eds., Curran Associates, Inc., 2019, pp. 8024–8035.
- [18] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.