



HAL
open science

Toward Robust Analog Equilibrium Propagation by Investigating Power Calculation Dynamics

Fatma Zulal Kiraz, Dang-Kiên Germain Pham, Patricia Desgreys

► **To cite this version:**

Fatma Zulal Kiraz, Dang-Kiên Germain Pham, Patricia Desgreys. Toward Robust Analog Equilibrium Propagation by Investigating Power Calculation Dynamics. Colloque du GDR SOC2, Jun 2023, Lyon, France. 2023. hal-04166116

HAL Id: hal-04166116

<https://telecom-paris.hal.science/hal-04166116>

Submitted on 19 Jul 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Toward Robust Analog Equilibrium Propagation by Investigating Power Calculation Dynamics

Fatma Zülâl Kiraz, Dang-Kiên Germain Pham, Patricia Desgreys

C2S Team, ComElec Dept., *LTCI, Télécom Paris, Institut Polytechnique de Paris*, Palaiseau, France

emails: {fkiraz;dpham;desgreys}@telecom-paris.fr

Abstract—In this work, we investigate the robustness and consistency of the Equilibrium Propagation (EqProp) algorithm for training analog neural networks, addressing previous research limitations. We analyze the algorithm’s robustness concerning learning parameter variations and examine the relationship between power calculations and convergence behavior. By accounting for all relevant components and their impact on learning, we aim to develop a reliable and consistent algorithm based on EqProp principles.

Index Terms—Equilibrium Propagation, analog neural networks, non-von Neumann architectures, energy based models, machine learning.

I. INTRODUCTION

The development of Artificial Intelligence has come a long way since its inception in the 1940s, with pioneers like John von Neumann, Alan Turing, and Norbert Wiener. The question "Can machines think?" posed in the 1950s led to the possibility of creating machines that could mimic human cognition [1]. Today, Artificial Intelligence research has made significant progress, with deep neural networks being a prime example of brain-inspired computing models. However, current deep neural networks based on ASICs that mostly rely on von Neumann architecture, are still far from matching the computational power and efficiency of the human brain, mainly due to the von Neumann bottleneck problem and the diminishing returns of Moore’s Law [2], [3].

Inspired by the brain’s inherent efficiency, non-von Neumann architectures offer a promising approach to enhance time and energy efficiency in computing systems by integrating memory and processing units using local variables. Backpropagation, a widely-used supervised learning algorithm, has been pivotal in training ANNs [4]. Recently, implementing the Backpropagation algorithm in the analog domain has gained popularity [5]. However, analog device imperfections introduce significant vulnerability due to the algorithm’s structure. The mismatch arising from implementing gradient computation with non-ideal circuit components results in increased error as the number of layers deepen. Equilibrium Propagation (EqProp) has emerged as a hardware-friendly alternative for training energy-based models (EBMs) [6], [7], [8]. [6] asserts that analog nonlinear neural networks are EBMs, making them suitable for training using the EqProp algorithm.

The previous work [6] did not analyze the algorithm’s robustness concerning learning parameter vari-

ations; in addition, it did not thoroughly investigate the relationship between power calculations and convergence behavior. This work aims to address the limitations by thoroughly analyzing the EqProp algorithm’s robustness and trying to find a better approximation of the (ideal) Backpropagation gradient. By considering all relevant components and their impact on the learning process, our goal is to develop a reliable and consistent algorithm based on Equilibrium Propagation principles.

II. EQUILIBRIUM PROPAGATION ALGORITHM

Unlike Backpropagation, EqProp computes quantities mimicking the gradient calculation through the minimization of "energy", making it ideal for non-von Neumann architectures [7]. [6] proposed a circuit implementation in which EqProp can be applied to perform a learning task.

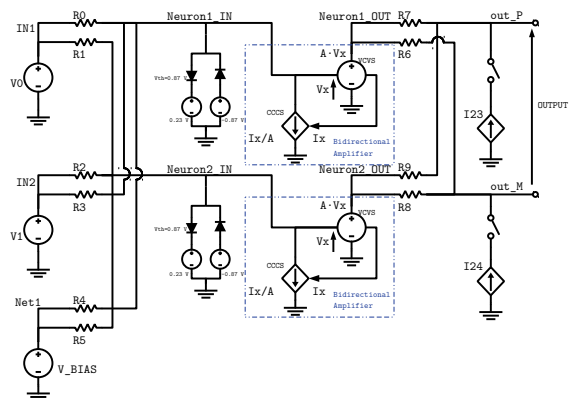


Fig. 1: Equilibrium Propagation Circuit Implementation

Fig. 1 displays the circuit structure used for our analysis, based on [6]. The circuit includes two input DC sources, a bias, one hidden layer with two neurons, and two output nodes. To address the positive conductance constraint, we only double the output nodes and subtract their voltages for potential negative outputs. The activation function is implemented with two antiparallel diodes in series with network (weight) resistors, and bidirectional amplifiers prevent voltage decay in both directions.

The algorithm has two training phases: the free and nudged phases. In the free phase, no current flows through the output nodes, as switches are open. The system reaches a steady state, and the difference

between the target and obtained output yields a loss value to be minimized throughout the iterations.

In the nudged phase, we introduce a constraint on the circuit to move it from its equilibrium state, allowing for the computation of the gradient. A feedback current, proportional to the loss value, is injected by closing the switches.

In Backpropagation, the weights (conductances) are updated with the gradient of the loss function with the following equation:

$$g_{ij} \leftarrow g_{ij} - \alpha \frac{\partial L}{\partial g_{ij}} \quad (1)$$

where α is the learning rate.

In Equilibrium Propagation, according to [6, Theorem 1] (2), the gradient of the loss can be estimated as:

$$\frac{\partial L}{\partial g_{ij}} = \lim_{\beta \rightarrow 0} \frac{1}{2\beta} ((V_{ij}^{\beta})^2 - (V_{ij}^0)^2), \quad (2)$$

where g_{ij} represents the conductance of a linear resistor with terminals i and j ; and V_{ij}^0 and V_{ij}^{β} represent the voltage drop through resistors in the free phase and nudged phase, respectively; and L represents the loss calculated after free phase.

III. PRELIMINARY RESULTS

The implementation represented in Fig. 1 is tested on a simple task, which will be called as *case 1*: while $V_0 = -2V$ and $V_1 = 2V$, $V_{out_target} = 1V$. Then, we observed the impact of variation of α and β on the algorithm performance.

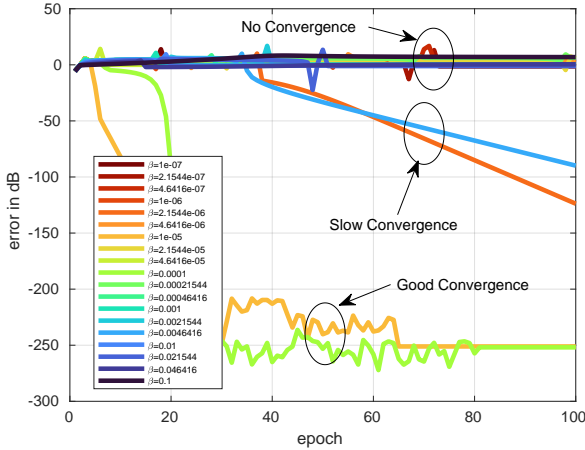


Fig. 2: Evolution of the Error for $\alpha = 5 \times 10^{-4}$ and $10^{-7} \leq \beta \leq 0.1$

Despite observing favorable convergence behaviors like the green and orange lines, most beta values resulted in non-convergence, highlighting the algorithm's sensitivity to learning parameter changes and lack of robustness. Additionally, learning failed for most randomized initial conductance values, indicating inconsistency in the learning process. This discrepancy between theoretical predictions and actual analog circuit learning performance raises questions about

the causes of this mismatch. The convergences issues observed in the EqProp algorithm in Fig.2 would signal the potential for improving gradient estimation.

To better understand this discrepancy and investigate the underlying causes, we delve into the power computation of two-terminal components within the nonlinear resistive networks function as EBMs with their energy function representing the total pseudo-power of the circuit. We employ a variational approach using PySpice to compute the derivative of the power with respect to the resistor values, enabling a deeper examination of the energy perspective in these networks.

IV. CONCLUSION

Equilibrium Propagation is a promising algorithm for energy-efficient, fast analog deep learning circuits. This paper presents the EqProp algorithm in training analog neural networks, highlighting its sensitivity to learning parameter variations and initial conductance values. The mismatch between theoretical predictions and actual performance can be attributed to previously overlooked components in power calculations.

Our overall project focuses on a comprehensive analysis, considering all relevant components and their effects on the optimization process. Future work will aim to optimize the EqProp algorithm for increased stability and robustness and develop gradient estimation equations for the loss function. By achieving these objectives, we aim to advance the understanding of EqProp in analog neural networks and contribute to developing low-power, fast, and reliable machine-learning analog circuits inspired by the human brain's neural network.

REFERENCES

- [1] A. M. TURING, "I.—COMPUTING MACHINERY AND INTELLIGENCE," *Mind*, vol. LIX, no. 236, pp. 433–460, 10 1950. [Online]. Available: <https://doi.org/10.1093/mind/LIX.236.433>
- [2] S. Spiga, A. Sebastian, D. Querlioz, and B. Rajendran, "Chapter 1 - role of resistive memory devices in brain-inspired computing," in *Memristive Devices for Brain-Inspired Computing*, ser. Woodhead Publishing Series in Electronic and Optical Materials, S. Spiga, A. Sebastian, D. Querlioz, and B. Rajendran, Eds. Woodhead Publishing, 2020, pp. 3–16.
- [3] J. Hasler and H. Marr, "Finding a roadmap to achieve large neuromorphic hardware systems," *Frontiers in Neuroscience*, vol. 7, 2013. [Online]. Available: <https://www.frontiersin.org/articles/10.3389/fnins.2013.00118>
- [4] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, pp. 533–536, 1986.
- [5] A. Fasoli, S. Ambrogio, P. Narayanan, H. Tsai, C. Mackin, K. Spoon, A. Friz, A. Chen, and G. W. Burr, "Optimization of analog accelerators for deep neural networks inference," in *2020 IEEE International Symposium on Circuits and Systems (ISCAS)*, 2020, pp. 1–5.
- [6] J. Kendall, R. Pantone, K. Manickavasagam, Y. Bengio, and B. Scellier, "Training end-to-end analog neural networks with equilibrium propagation," 2020.
- [7] B. Scellier and Y. Bengio, "Towards a biologically plausible backprop," *CoRR*, vol. abs/1602.05179, 2016. [Online]. Available: <http://arxiv.org/abs/1602.05179>
- [8] —, "Equilibrium propagation: Bridging the gap between energy-based models and backpropagation," *Frontiers in Computational Neuroscience*, vol. 11, 2017. [Online]. Available: <https://www.frontiersin.org/articles/10.3389/fncom.2017.00024>