

# Inverse Analog IC Sizing and Exploration through Diffusion Models and Structural Knowledge

Filipe Azevedo\*, Markus Leibl\*, Ricardo Martins and Helmut Graeb

**Abstract**—In the realm of analog integrated circuit sizing, quick and effortless exploration of design spaces is of major importance in the face of rapid development cycles, varying specifications and increasingly complex circuits. It is also well known, that the exploitation of slack in specifications by making adjustments to transistor device sizes can lead to improved yield. We address those challenges by combining two state-of-the-art machine learning approaches to achieve accurate generation of optimal performance points, while also allowing for the possibility to explore nearby sizings that might lead to more robust designs. To this end, a diffusion model is combined with an algorithm that analyzes the circuit and splits the problem into simpler subproblems. The models are tested on a set of typical operational amplifiers.

## I. INTRODUCTION

Since the development of SPICE, the industry-accepted method for designing analog circuits has not changed much. While the SPICE models keep becoming more complex [1], the approach of manual device tuning persists [2]. Key challenges of making analog CAD tools useful in practice are robustness, ease of use and applicability to a high range of use cases. All those aspects are not easy to grasp on a fundamental level. On the other hand, machine learning (ML) approaches excel in modeling nonlinear black box phenomena. Their combination with already existing conventional techniques has a high potential to bridge remaining gaps that prevent the adoption of tools in industry.

Analog design is inherently an inverse problem. While the calculation of performances, yield etc., is satisfactorily possible via simulation, the inverse, i.e., getting a circuit with optimal parametrization, requires experience, trial and error and complex numerical routines. As shown in previous works [3], [4], machine learning (ML) methods can be used to accurately reproduce sizings on (Pareto) optimal fronts. Furthermore, it has been shown [5] that diffusion models are well adaptable to the inverse sizing problem. Other ML based

works attempted to solve this same problem. In [6], artificial neural networks (ANNs) are used to guess circuit sizings, utilizing only one ANN for modeling an entire netlist, leading to a higher learning task complexity and thus large sample sizes (in [6] 16600 or roughly 16x the sample of this work) are required to learn the inverse problem within a smaller performance range. In [7], the problem is broken down into smaller tasks, and one ANN is trained to sequentially size a single device, in the same vein as [3], feeding the outputs of the previous ANNs to the input of the next, producing a cascade of networks that can be sorted in different ways until an optimal order is found. The drawback is that for more complex circuits of, e.g., 15 devices, optimally ordering the ANNs for sizing leads to a complexity of 15!.

Other works follow the traditional approach of tackling the direct sizing problem, treating it as an optimization problem [8]–[11]. Reinforcement learning was also studied, with [12]–[14] replacing the optimizer with an ANN, leading to improvements in prediction speed, accuracy, and scalability. However, to the best of our knowledge, none of them show strengths in all four categories, speed, accuracy, sample efficiency and exploration.

In this work, we aim for a combination of all those virtues by applying denoising diffusion probabilistic models (DDPM) to the sizing of OpAmps, utilizing the approach in [3] of subdividing the circuits into simpler structures, which has shown high accuracy for low sample sizes. DDPMs have proven their potential for generating strongly varied, but realistic results in other fields, OpenAI’s DALL-E and Stable Diffusion being among the most prominent examples. As stated, [5] proved these model’s efficacy in the inverse sizing problem.

Unlike most ML models, DDPMs can suggest different sizings for the same problem by repeatedly sampling them. As shown in Section IV, this approach to the inverse problem can produce a more varied set of possible sizings, which is very helpful when exploring sizing options in the vicinity of performance optima. This work contributes a ML approach that:

- can suggest valid sizings with resulting performance in the vicinity of the required metrics, allowing for exploration of sizings in order to pursue further optimal performance regions, while still fulfilling specifications,
- is highly data efficient, requiring a fraction of the dataset size used in [5] and [6],
- has an accurate expectation of the generated sizings,
- generates sizing solutions at push-button speed.

Filipe Azevedo and Ricardo Martins are with Instituto de Telecomunicações and Instituto Superior Técnico - Universidade de Lisboa, (email: filipeazevedo@gmail.com, ricmartins@lx.it.pt). Markus Leibl and Helmut Graeb are with Technical University of Munich, Chair of Electronic Design Automation (email: markus.leibl@tum.de, helmut.graeb@tum.de).

\* Both authors contributed equally to this work.

Parts of this work have been done within the collaborative project HoLoDEC funded by the Federal Ministry of Education and Research Germany (BMBF) under the funding code 16ME0705. and also funded by Fundação para a Ciência e Tecnologia – Ministério da Ciência, Tecnologia e Ensino Superior (FCT/MCTES) through national funds and, when applicable cofounded by European Union (EU) funds under the projects UIDB/50008/2020 (DOI identifier 10.54499/UIDB/50008/2020), ACTON (DOI identifier 10.54499/2023.11981.PEX), and also, by Sony Semiconductor Solutions (project GENERALISE).

## II. FROM STRUCTURAL ANALYSIS TO CIRCUIT SIZING VIA MACHINE LEARNING

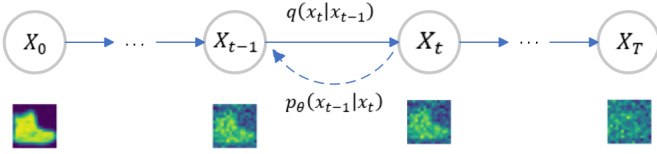


Fig. 3: Correlation between forward and reverse processes [5].

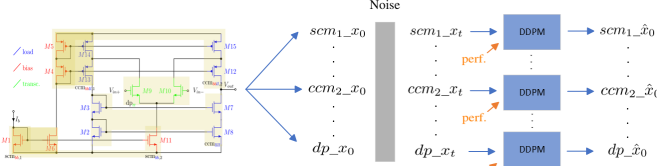


Fig. 4: Training phase for sizing an OpAmp. After the decomposition into sub-blocks, noise is added to the respective sizings. A DDPM is trained for each sub-block to predict the original sizing values. The respective performances of the original sizings are given as guidance.

3) *Sampling Process*: Finally, after the model is trained, the sampling process is used to generate new data. The model learned to reverse the forward process, and so, when giving it pure Gaussian noise, it tries to reverse the noise addition until arriving at the original data. Except this time, there is no actual data to begin with, marking this a form of generative artificial intelligence. This stage is similar to the training phase shown in fig. 4, except random noise is given as input to the network instead of the noisy sizings.

It was shown by [5] that DDPMs produce interesting results when applied to the inverse sizing problem. In this work, a similar implementation is considered, with a cosine schedule for  $\beta$  and classifier-free guidance [16], but with 2 key differences:

- instead of predicting the noise  $\epsilon$ , we implement a model that predicts a velocity equation  $v_t = \alpha_t \epsilon - \sigma_t x$  which has been shown to enable a true signal-to-noise ratio of 0 at the last timestep  $T$  (meaning that at the last timestep the input of the model is pure Gaussian noise), as opposed to other DDPM parametrizations. This removes an important discrepancy between training and inference [17];
- the backbone of the model implemented is a transformer, with an architecture similar to the “adaLN-Zero” model of [18], as opposed to the simple multi-layer perceptrons implemented by [5]. To predict a more complex value in  $v$ , a stronger architecture was required.

#### IV. EXPERIMENTAL RESULTS

We trained our approach on two different OpAmps, the FCOA previously studied in [3] shown in fig. 2 and the Miller OpAmp (MOA) in fig. 5. After training, we sampled the models for all ground truth sizings in the test sets of both OpAmps, corresponding to 118 for the FCOA and 77 for the Miller OpAmp. After simulating the sampled values, the resulting performances were compared with simulations of the ground truth values. All models have been trained with a

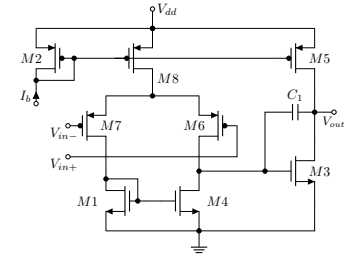
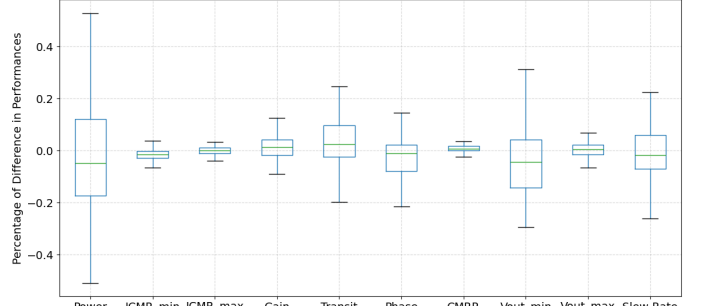
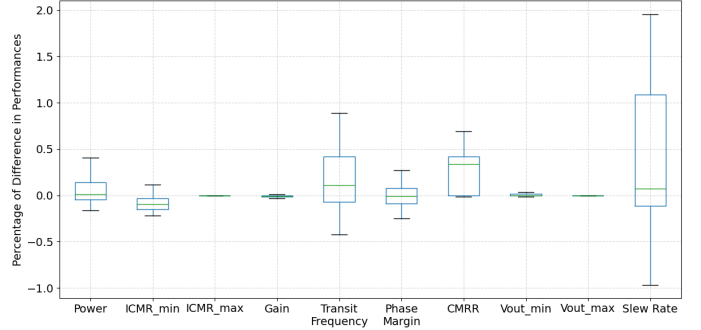


Fig. 5: Miller OpAmp (MOA)



(a) FCOA.



(b) MOA.

Fig. 6: Boxplots of the percentage of differences in performances for the test set of both tested OpAmps, after removal of outliers. The innermost line represents the median, the box 50% of the data.

sample size between 700 and 1000 elements. While slightly less efficient as [3], the price for added flexibility, the sample sizes are very low.

Boxplots of the relative performance differences are presented in fig. 6, bar outliers. For the FCOA in fig. 6a, most boxes are centered around 0 (i.e. the ground truth), but some performances show a tendency of improvement. For Power, ICMR<sub>min</sub> and Vout<sub>min</sub> median and box extend downwards (the predictions’ performances are lower than the labels), while Gain and Transit Frequency median and box extend upwards (the predictions’ performances are greater than the ground truth sizings). Similar observations can be made for the MOA in fig. 6b, with the CMRR and Slew Rate having the biggest improvement. Here, the only performance that showed a tendency toward deterioration was Power. This shows that the models are able to accurately learn the distribution of the circuits’ sizing and respective performances, sometimes finding interesting improvements that might be attributed to

TABLE I: Estimated mean value and standard deviation of W parameter sizes of the FCOA for 100 samples and same performance specification.

device	mean value	standard deviation
0	8.52	2.0
1	43.26	19.0
2	41.74	17.0
3	14.4	1.0
4	90.98	8.0
5	88.47	23.0
6	41.74	17.0
7	43.26	19.0
8	97.88	55.0
9	97.88	55.0
10	114.78	35.0
11	367.21	22.0
12	367.21	22.0
13	244.11	13.0
14	244.11	13.0

learned trade-offs between sizes of the different sub-blocks.

In a second experiment we sampled the model for the FCOA 100 times, while keeping the same performance specification. The results can be seen in fig. 7 and table I. Most distributions in the performance space in fig. 7 exhibit a gaussian behaviour, centered around the ground truth performance marked by a red star. Table I indicates that the model learns to estimate the different contributions and sensitivities for the individual parameters. It can be seen that the standard deviation varies heavily depending on the device, indicating that the model learns to scale the effects of parameter deviations.

## V. CONCLUSION

We have shown a machine learning approach to not only reproduce optimal sizings for a given specification, but also to help explore points in the nearby performance space. In the scenario of sampling multiple times for the same specification, the sampling can be interpreted as adding random noise to the ground truth sizing. The model learns the corresponding covariance matrix. As the standard deviation is not equal for all devices, this behaviour cannot be mimicked by simply adding noise to an optimal sizing.

## REFERENCES

- [1] C. Gatermann and R. Sommer, “Teaching the mosfet: A circuit designer’s view,” in *Int. Conf. on SMACD*, July 2022, 2022, pp. 1–4.
- [2] G. Gielen, “Analog synthesis 3.0: Ai/ml to synthesize and test analog ics: hope or hype ?” in *2023 ACM/IEEE 5th Workshop on Machine Learning for CAD (MLCAD)*, 2023, pp. 1–1.
- [3] M. Leibl and H. Graeb, “Optimizer-free sizing of opamps leveraging structural and functional properties,” in *Int. Conf. on SMACD*, July 2024, 2024, pp. 1–4.
- [4] N. Lourenço, E. Afacan, R. Martins, F. Passos, A. Canelas, R. Póvoa, N. Horta, and G. Dundar, “Using polynomial regression and artificial neural networks for reusable analog ic sizing,” in *Int. Conf. on SMACD*, July 2019, 2019, pp. 13–16.
- [5] P. Eid, F. Azevedo, R. Martins, and N. Lourenço, “Solving the inverse problem of analog integrated circuit sizing with diffusion models,” in *Int. Conf. on SMACD*, July 2024, 2024, pp. 1–4.
- [6] N. Lourenço, J. Rosa, R. Martins, H. Aidos, A. Canelas, R. Póvoa, and N. Horta, “On the exploration of promising analog ic designs via artificial neural networks,” in *Int. Conf. on SMACD*, July 2018, 2018, pp. 133–136.

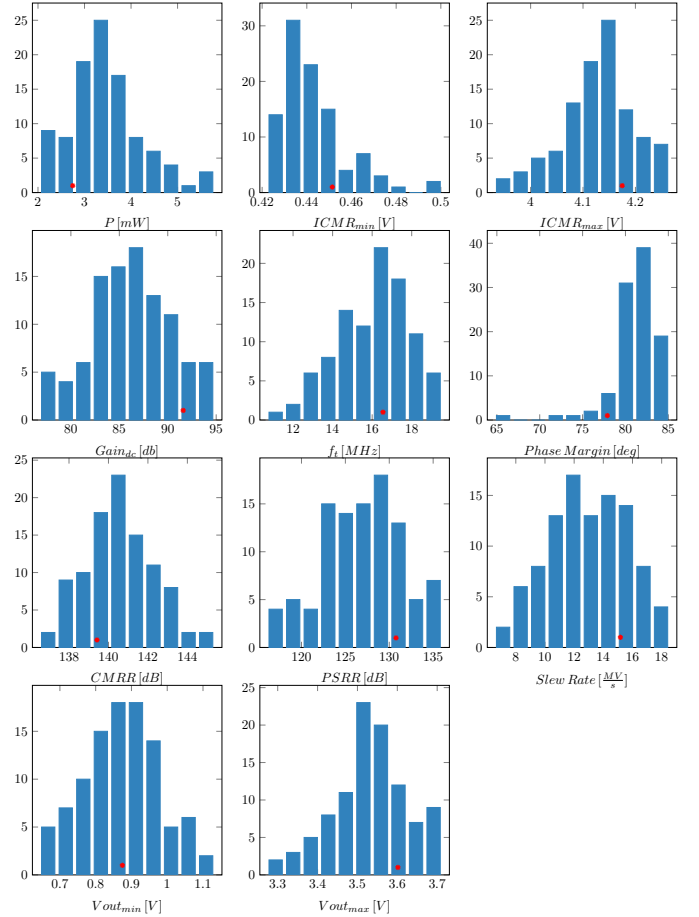


Fig. 7: Histograms for 100 samples of the trained FCOA model with the **same** performance specification. Most distributions are well shaped around the ground truth from the test set. The performance value of the ground truth sizing is marked by a red star.

- [7] P.-O. Beaulieu, E. Dumesnil, F. Nabki, and M. Boukadoum, “Analog rf circuit sizing by a cascade of shallow neural networks,” *IEEE TCAD*, vol. 42, no. 12, p. 4391–4401, 2023.
- [8] M. Fayazi, M. T. Taba, E. Afshari, and R. Dreslinski, “Angel: Fully-automated analog circuit generator using a neural network assisted semi-supervised learning approach,” *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 70, no. 11, pp. 4516–4529, 2023.
- [9] A. F. Budak, D. Smart, B. Swahn, and D. Z. Pan, “Apostle: Asynchronously parallel optimization for sizing analog transistors using dnn learning,” in *Proc. of ASP-DAC*, Jan. 2023, 2023, pp. 70–75.
- [10] G. Wolfe and R. Vemuri, “Extraction and use of neural network models in automated synthesis of operational amplifiers,” *IEEE TCAD*, vol. 22, no. 2, pp. 198–212, 2003.
- [11] K. Hakhamaneshi, M. Nassar, M. Phielipp, P. Abbeel, and V. Stojanovic, “Pretraining graph neural networks for few-shot analog circuit modeling and design,” *IEEE TCAD*, vol. 42, no. 7, pp. 2163–2173, 2023.
- [12] J. Gao, W. Cao, and X. Zhang, “Rose: Robust analog circuit parameter optimization with sampling-efficient reinforcement learning,” in *2023 60th ACM/IEEE Design Automation Conference (DAC)*, 2023, pp. 1–6.
- [13] K. Settaluri, Z. Liu, R. Khurana, A. Mirhaj, R. Jain, and B. Nikolic, “Automated design of analog circuits using reinforcement learning,” *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 41, no. 9, pp. 2794–2807, 2022.
- [14] J. Zhang, J. Bao, Z. Huang, X. Zeng, and Y. Lu, “Automated design of complex analog circuits with multiagent based reinforcement learning,” in *2023 60th ACM/IEEE Design Automation Conference (DAC)*, 2023, pp. 1–6.

- [15] J. Ho, A. Jain, and P. Abbeel, "Denoising diffusion probabilistic models," 2020. [Online]. Available: <https://arxiv.org/abs/2006.11239>
- [16] J. Ho and T. Salimans, "Classifier-free diffusion guidance," 2022. [Online]. Available: <https://arxiv.org/abs/2207.12598>
- [17] S. Lin, B. Liu, J. Li, and X. Yang, "Common diffusion noise schedules and sample steps are flawed," 2024. [Online]. Available: <https://arxiv.org/abs/2305.08891>
- [18] W. Peebles and S. Xie, "Scalable diffusion models with transformers," 2023. [Online]. Available: <https://arxiv.org/abs/2212.09748>