A Hybrid Machine Learning and Numeric Optimization Approach to Analog Circuit Deobfuscation

Dipali Jain, Guangwei Zhao, Rajesh Datta, Kaveh Shamsi

THE JONSSON SCHOOL
DEPT OF ELECTRICAL &
COMPUTER ENGINEERING

Department of Electrical And Computer Engineering
University of Texas At Dallas, Richardson, USA

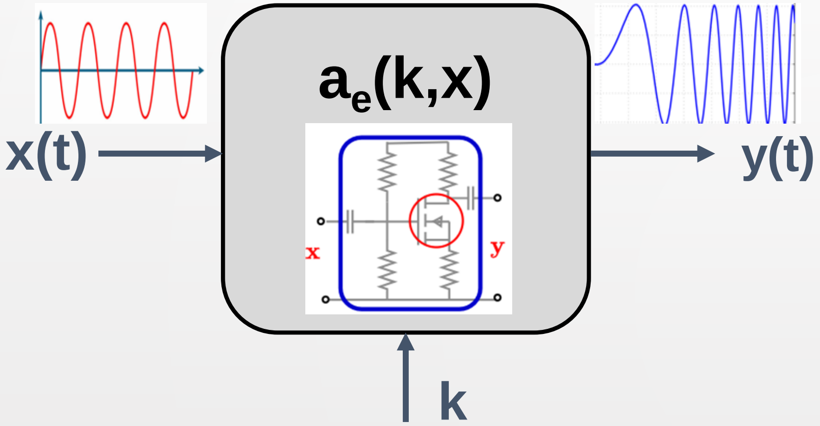
ASIA SOUTH PACIFIC DESIGN AUTOMATION CONFERENCE

22.01.2025

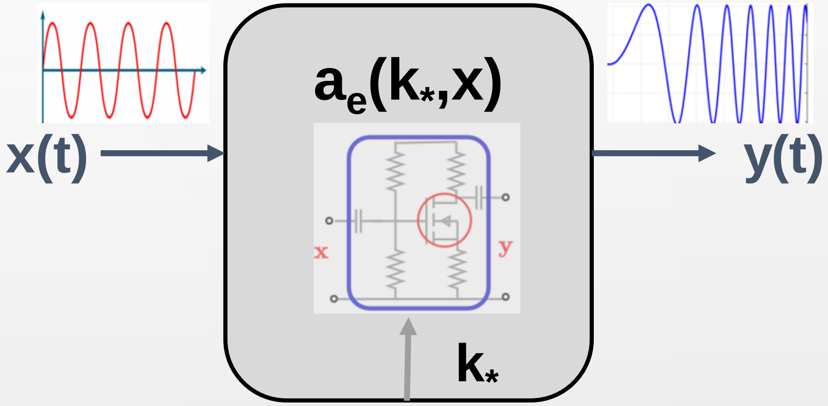
Contents

Oracle Guided Circuit Learning (OGCL)

Obfuscated (Analog) Circuit



Oracle (Analog) Circuit



where k_* = unknown analog values

x = controllable analog inputs

a_e = analog circuit

y = output response

$$a_e(k,x) \equiv a_e(k_*, x)$$

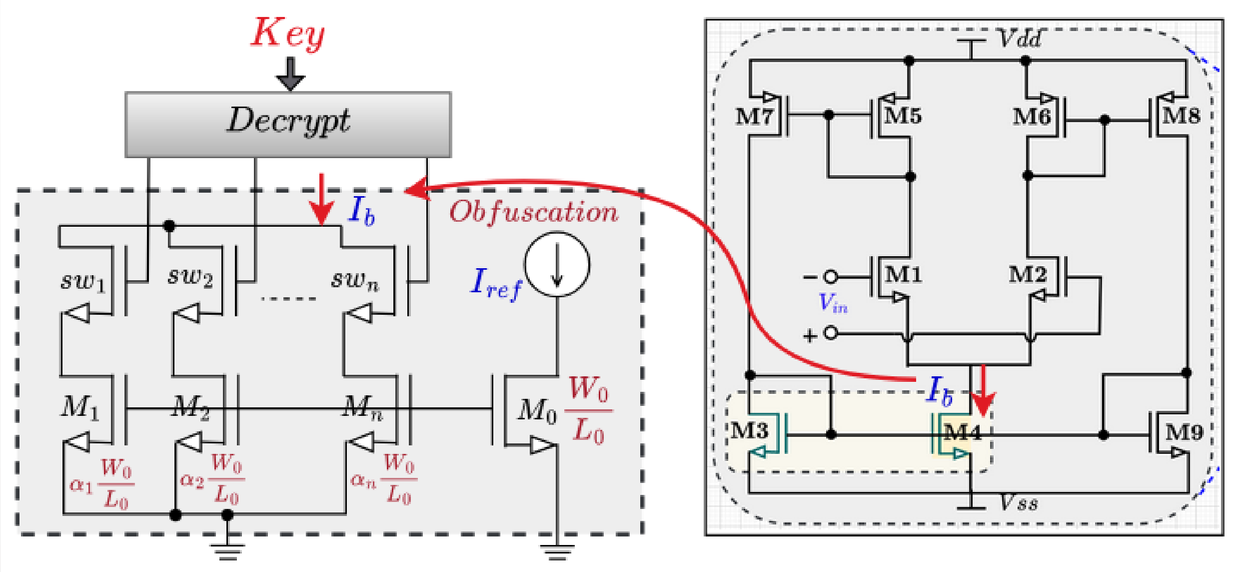
Applications of OGCL

- Security Analysis of circuit obfuscation
- Predicting Unknowns that are obscured in reverse engineering[1]
- Trojan Detection [2]

[1] Jain, Dipali, Guangwei Zhao, Rajesh Datta, and Kaveh Shamsi. "Towards Machine-Learning-based Oracle-Guided Analog Circuit Deobfuscation." In *2024 IEEE International Test Conference (ITC)*, pp. 323-332. IEEE, 2024.
[2] Datta, Rajesh Kumar, Guangwei Zhao, Dipali Jain, and Kaveh Shamsi. "On Hardware Trojan Detection using Oracle-Guided Circuit Learning." In *Proceedings of the Great Lakes Symposium on VLSI 2024*, pp. 198-203. 2024.

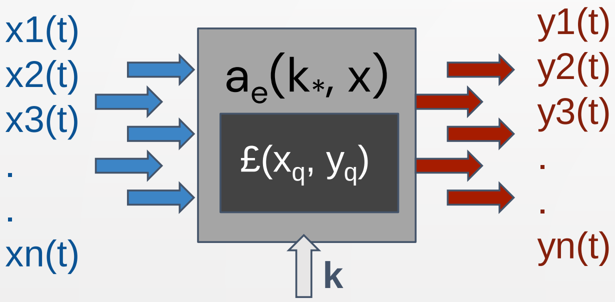
Analog Circuit Obfuscation.

Obfuscating a current mirror circuit used to bias an amplifier



Equation (Analytical) Based Approach

Time-independent analog circuit :



fj (xn, kn, sn) = yn
xn = query inputs
kn = unknowns
sn = intermediate variables
yn = observables

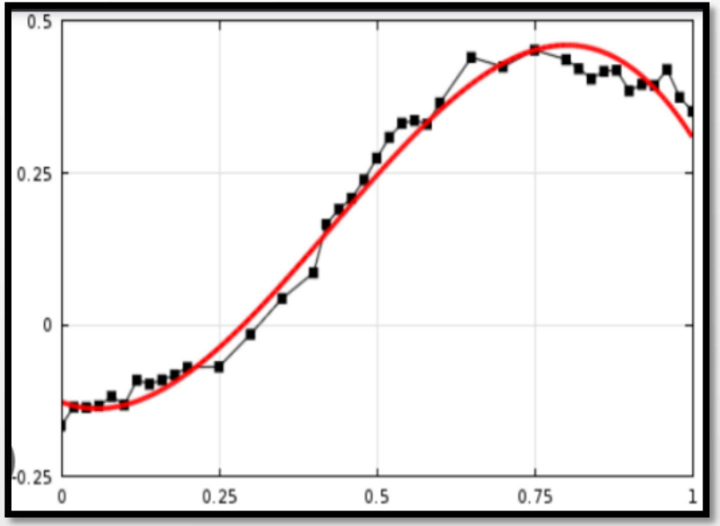
- Static: Nonlinear Programming (NLP)

Time-dependent analog circuit :

- Dynamic: Differential-Algebraic Equations (DAEs), with time derivatives of x(t), y(t), s(t)

Optimal Control Approach

- Finding the optimal value of the set of unknown parameters k,
- Solutions to DAEs are sufficiently close to a set of desired curves.
- Larger systems involve performing numeric integration with new kn created at every time point
- Using NLP to solve expanded system of equations optimize for the desired fitness function



Deobfuscation Via ML and Optimization

Motivation

- Eliminate the need for manual derivation of circuit equations
- Published SMT-based attack:
  - equations are extracted by hand that relate unknowns to specific observations
- Emerging trend in analog EDA domain:
  - use of end-to-end and graph-based machine learning to replace tasks like:
    - SPICE simulation
    - Parameter optimization

Contents

Overall Flow

Dataset Generation

Training

Testing

Obfuscated circuit ($a_e(k,x)$)

for input X_q = {1,2, . . ., **Q**}
for unknown K = {1,2, …, **L**}
for random choice U = {1,2 ,…, i}
{
perform simulation
extract Y_q ={1,2…., **Q**}
}
+

Data Augmentation with PV: U(LQ)

Train ML Model

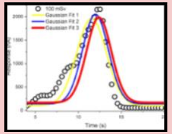
Oracle circuit ($a_e(k^*,x)$)

for input Xq = {1,2, . . ., **Q**}
{
query oracle
extract Yq ={1,2…., **Q**}
}

|k|
Predicted Unknown

|K+|
Initial Guess

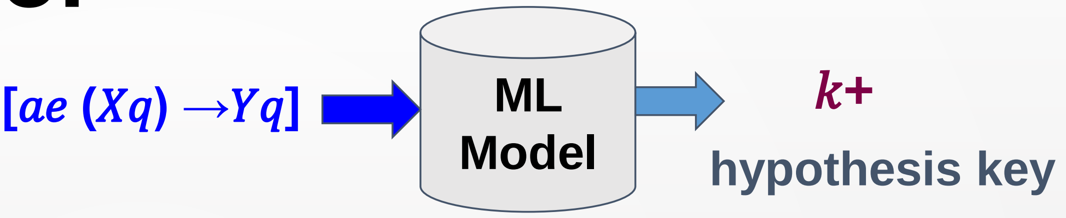
HSpice Optimization
Curve Fit



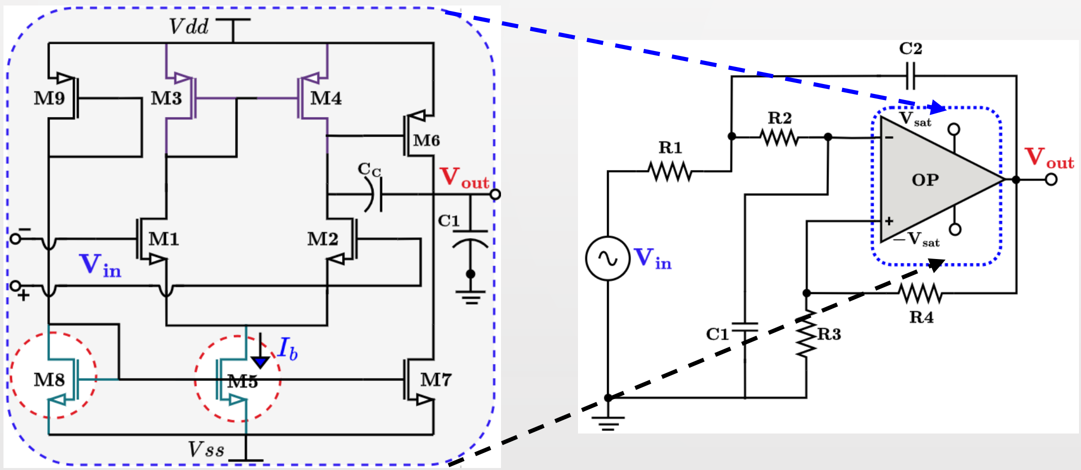
Final unknown value hypothesis

ML Feature Vector

$Xq \rightarrow$ set of Q input queries,
$Yq \rightarrow$ output observations
$ae \rightarrow$ obfuscated circuit



Second-Order Low Pass Filter



Fin : [V_{ac} , V_{dc} , $G_o\,(f_1)$, ...$G_o\,(f_{25})$, f_c, K, R_1, R_2, C_1, W_1/L_1, W_2/L_2 , W_3/L_3, dev_count, num_cap, num_resi, num_trans]

Target Unknown: Wk_5/Lk_5 , Ck_2.

Hybrid Approach

Benefits:

- **Convergence.**
  - Better convergence with the ML-produced initial guess**.**
- **More than One-Shot.**
  - There is no feedback or adjustment over the solution in ML.
- **Dealing with High Dimensionality.**
  - ML approach as a standalone additionally suffers from the curse of dimensionality when the number of unknowns is large.

Contents

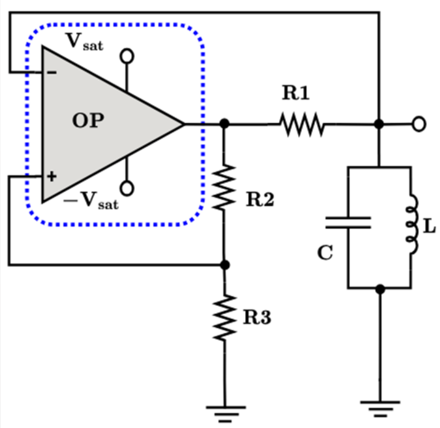
Experiments and Results

- ML Model:
  - RF
  - Figure of merit (PE): $(yp - yt)/yt \times 100$
- Star-HSpice optimization.
  - curve-fitting optimization: 20 points from transient simulations.
  - goal-based optimization: user-defined variables to hit certain targets (settling time, rise time, etc.)
    - requires providing a data file with the target values, an input netlist file, optimization parameters, component limits, and (critical to our hybrid scheme) an initial guess

Benchmark Circuits

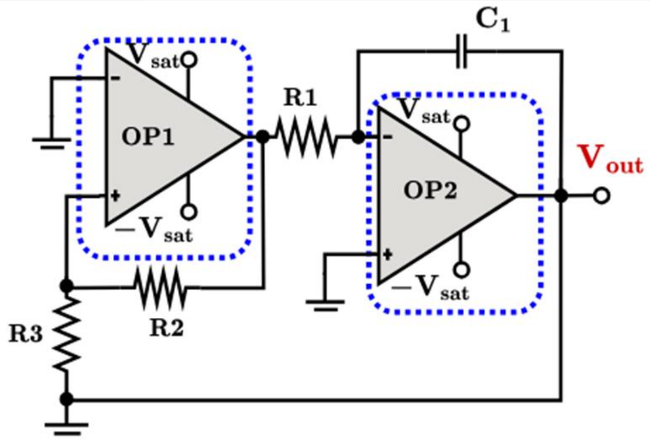
Oscillators

LCO



samples 10000 in 103.6 seconds for 2 unknowns

TWG

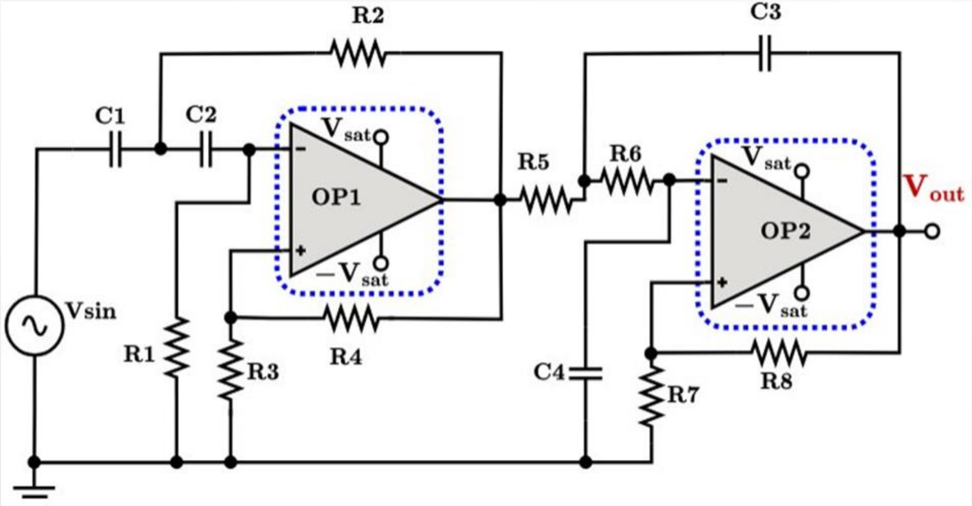


samples 10000 in 119.3 seconds for 3 unknowns

	LCO		TWG		
	U1	U2	U1	U2	U3
ML	0.4	1.3	0.1	1.6	1.8
ML(PV)	4.15	6.87	9.4	12.6	14.8
Spice	1.1	4.7	5.26	6.14	6.32
ML+ Spice	0.01	0.12	0.01	0.25	0.72

Benchmark Circuits

Fourth order Butterworth Sallen Key Band Pass Filter

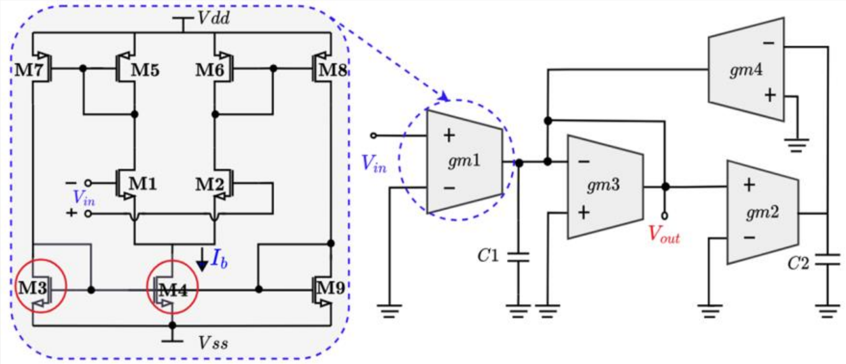


samples 32400 in 535.4 seconds for 4 unknowns

	FOSKBPF			
	U1	U2	U3	U4
ML	3.92	2.83	3.45	4.34
ML(PV)	14.7	20.1	20.0	23.3
Spice	6.49	2.71	4.66	1.24
ML+ Spice	1.33	1.21	1.68	1.91

Benchmark Circuits

Second order Gm-C Band Pass Filter using OTA with CCM supplying the bias current [3]



 # samples 1500 in 103.6 seconds for 6 unknowns

	GmC-BPF					
	U1	U2	U3	U4	U5	U6
ML	3.4	3.7	4.8	5.9	5.8	6.6
ML(PV)	9.2	10.0	12.5	13.8	14.6	16.0
Spice	2.09	1.24	4.55	5.22	9.18	2.89
ML+ Spice	1.8	1.07	1.32	1.66	1.21	1.01

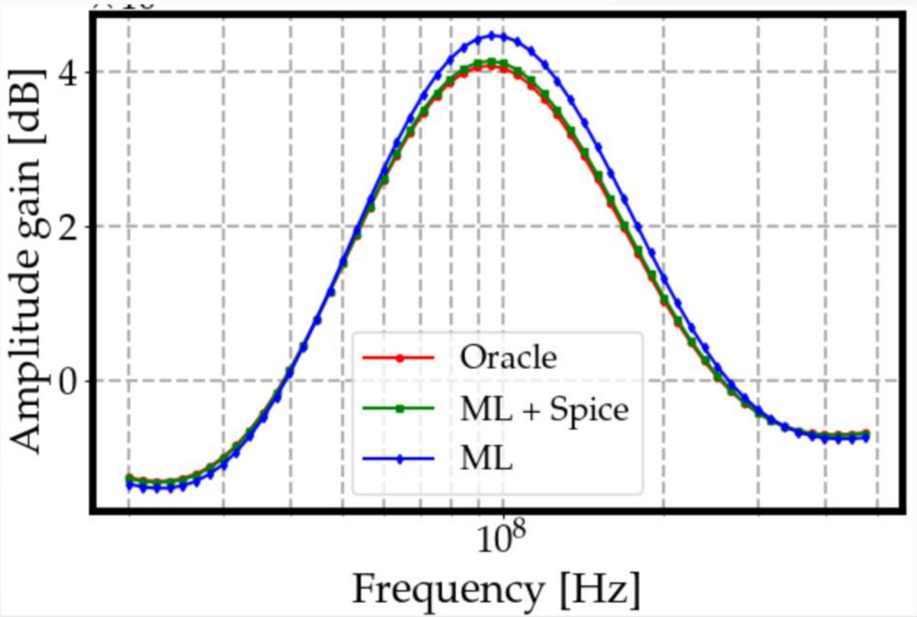
[3] N. G. Jayasankaran et all, "Breaking Analog Locking Techniques via Satisfiability Modulo Theories," 2019 IEEE International Test Conference (ITC), Washington, DC, USA, 2019, pp. 1-10, doi: 10.1109/ITC44170.2019.9000113.

Contents

Discussion

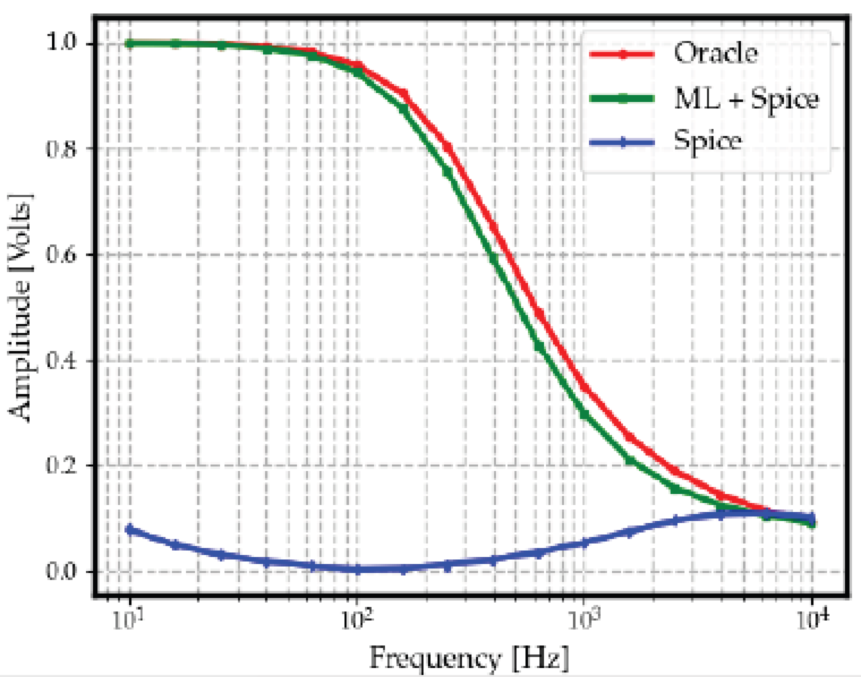
FOGmBPF simulation for Oracle circuit, by substituting

- values of unknowns predicted by ML model
- values of unknowns predicted using a hybrid ML+Spice approach



Discussion

Spice optimizer solution for the **Second Order LPF** when plugged into the circuit diverges from the oracle at lower frequencies indicating a local optima trap. The ML+Spice flow however avoids this.



Contents

- Background and Motivation
- Proposed Method
- Experiments and results
- Discussion
- **Conclusion**

Conclusion

- First use of a hybrid machine learning and optimization approach to automated generic analog circuit deobfuscation
- Does not rely on manual equation extraction and analysis by an expert.
- The automated flow based on using an ML-produced guess as an initial value in a subsequent Newton method optimization loop
- Demonstrated superior accuracy and runtime compared to either method as a standalone, especially in the presence of process variation.

Acknowledgement

This work was supported by a grant from the **National Science Foundation (NSF-2155189)**.

Questions?