A 4-8 GHz Cryo-CMOS LNA with Decoupled Noise-Power Matching and Reduced Self-Heating Achieving 12.3-K NT and 13 dB Return-Loss

Jiawei Li, Yatao Peng, Yue Wu, Chen Deng, Jun Yin, Rui P. Martins, Pui-In Mak
State Key Lab. of Analog and Mixed-Signal VLSI, IME/ECE of FST, University of Macau, China
yataopeng@um.edu.mo

***Abstract*—A broadband CMOS low noise amplifier (LNA) implemented in 28-nm bulk CMOS and operating under cryogenic temperature (CT) is presented. Besides source degeneration, the cascode topology together with a frequency-dependent LC tank load introduces an additional degree of design freedom. Thus, the effective Miller capacitance contributed by C_{gd} becomes frequency-dependent and decreases as the frequency increases, which facilitates wideband simultaneous noise and input matching (SNIM). An electro-thermal simulation method is proposed to quantify the self-heating effect (SHE) and a dual-attenuator cryogenic noise measurement method is introduced to remove the uncertainty caused by the attenuator temperature. The LNA achieves a measured gain S_{21} of 23.1±2.4 dB, return loss better than 10 dB, and a noise figure (NF) of 1.43–2.94 dB over 3.4–8.4 GHz at 300 K. At 8.8 K, it shows a measured gain of 29.8±2.3 dB and noise temperature (NT) of 12.3-31 K from 4 to 8 GHz.**

***Index Terms*—Cryogenic CMOS, low noise amplifier, self-heating effect, measurement uncertainty, quantum readout**

I. Introduction

In superconducting quantum state readout (QSR) chains, the cryogenic low noise amplifier (cryo-LNA) with low noise temperature (NT) at cryogenic temperature (CT) is essential to amplify the extremely weak readout signals for subsequent processing [1]. Currently, the cryo-LNA is mainly realized by InP HEMTs. In recent years, SiGe HBTs have also been exploited for cryo-LNA implementation. However, to enable monolithic integration of the high-circuit-density QSR receiver, realizing cryo-LNA with sufficiently low NT in CMOS technology—particularly in bulk CMOS—remains an ongoing pursuit. Nevertheless, cryo-CMOS LNAs still suffer from relatively high NT and degraded input return loss at cryogenic temperatures [2]–[4]. The primary reasons are twofold: (1) the increased mismatch between the optimum noise impedance ($Z_{opt}(\omega)$) and the conjugate of the input impedance ($Z_{in}^{*}(\omega)$) at low temperatures (see Fig. 1(a)); and (2) the dramatically reduced thermal conductivity of silicon and silicon dioxide at CTs (see Fig. 1(b)), which exacerbates the self-heating effect (SHE) and results in excessively high channel temperature.

In this work, we present a cryo-CMOS LNA in a 28-nm CMOS process that alleviates the aforementioned design challenges. The drain-loaded impedance of the cascode stage is exploited to optimize return loss, while the gate matching network provides simultaneous noise and input matching (SNIM). Although the capacitive loading has been previously used for input power matching [2], [5], those designs typically require off-chip RF chokes or LC tanks between the common-source (CS) and cascode (CG) stages. This often induces a degraded NF and potential instability problems, which are avoided in this work. To handle SHE, we introduce an electro-thermal co-simulation methodology using CAD tools to evaluate SHE, thus guiding layout design, which reduces the average channel temperature by ~35%. In addition, a dual-attenuator de-embedding technique that eliminates the uncertainty associated with the attenuator temperature is proposed for accurate NT measurement at 4 K.

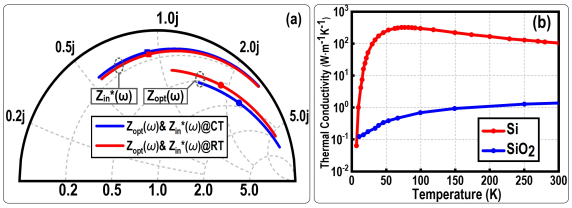

Fig. 1. Design challenges in cryo-LNAs: (a) increased mismatch between $Z_{opt}(\omega)$ and $Z_{in}^{*}(\omega)$, and (b) reduced thermal conductivity of Si and SiO_2.

II. Analysis of The Proposed LNA

The proposed LNA topology and its small-signal equivalent circuit for noise and input impedance analysis are shown in Fig. 2(a) and (b), respectively. It consists of an input low-noise stage and a gain-boosting stage, followed by a unity-gain output buffer for impedance matching. The first stage adopts a CS-CG cascode structure with inductive degeneration and an inductive L-type input matching network. The C_{gd} path of M_1 couples the impedance seen from the source node of M_2 into the input impedance with minimal impact on noise matching, achieving SNIM [2]. A transformer formed by L_g and L_d augments the contribution of the drain impedance to the input matching via C_{gd}, thereby introducing additional design flexibility.

A. Noise Analysis for Cryogenic Operation

It has been proven that the channel noise of a MOSFET does not necessarily scale linearly with ambient temperature

This work is supported by the Science and Technology Development Fund (FDCT) of Macau under Grant 0007/2025/EQP and 0001/2025/NRP , by National Natural Science Foundation of China under 62304004, Guangdong Provincial Quantum Science Strategic Initiative under GDZX-2306005, by University of Macau under Grant MYRG-GRG2025-00329-IME.

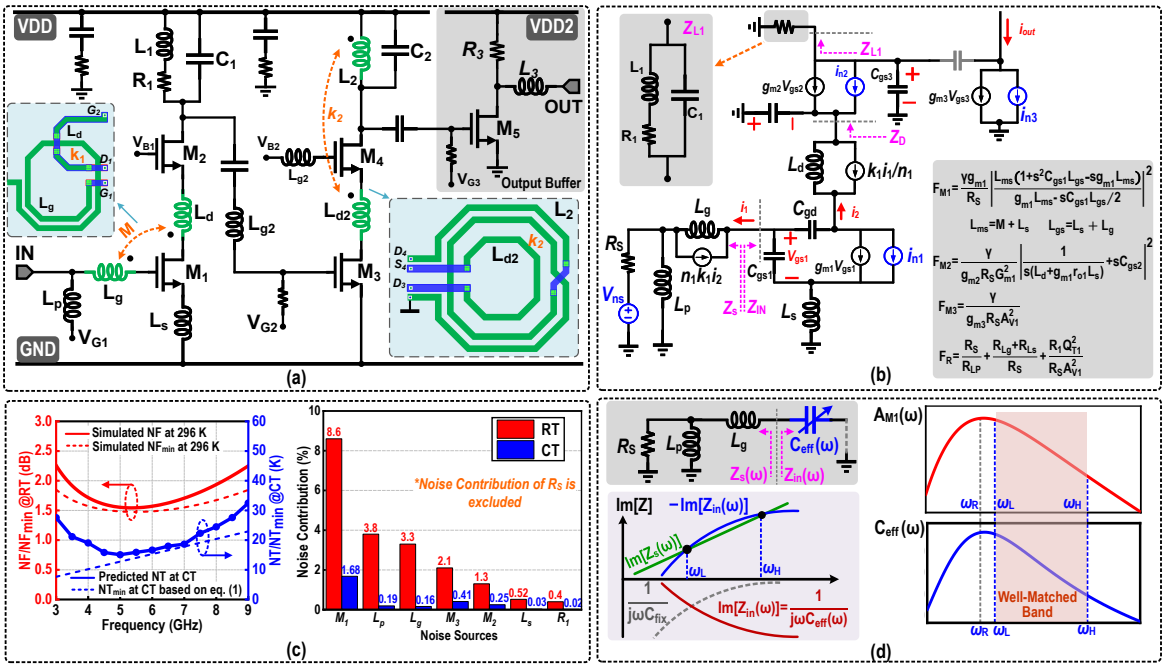

Fig. 2. (a) Schematic of the proposed LNA. (b) Small signal equivalent circuit. (c) noise performance simulation results. (d) Wideband input matching realization.

(T_a) due to the SHE and the prominent shot-like noise at CTs. The minimum noise temperature of M_1 accounting for the temperature influence can be expressed as [6], [7]

$$NT_{\min} \approx 2\frac{f}{f_t}\sqrt{\frac{qFI_{DS}r_{gs}T_c}{2k}}, \tag{1}$$

where f_t is the cutoff frequency, r_{gs} is the gate–source resistance, q is the elementary charge, k is Boltzmann's constant, F is the Fano factor that captures the degree of shot noise suppression, and T_c is the channel temperature. (1) indicates that decreasing T_c can lower noise temperature. In our design, T_c is evaluated from the electro-thermal simulation. The key parameters in the above model (i.e., f_t and r_{gs}) can be approximated using room temperature (RT) models together with proper cryogenic scaling [6]. The 50 nm rather than 28 nm channel length is selected for M_1 to suppress potential shot-noise-related contributions [6]. With the simulated T_c and scaled device parameters, an order of magnitude estimate of $NT_{\min}$ can be obtained even without compact RF models. For example, at T_c=60 K, the estimated $NT_{\min}$ is scaled by a factor α_N=1/7.6, reaching ~16 K. According to the small-signal model shown in Fig. 2(b), the overall noise factor can be expressed as

$$F_{LNA} = 1 + \alpha_M(F_{M1} + F_{M2} + F_{M3}) + \alpha_R F_R, \tag{2}$$

where F_{M1}, F_{M2}, and F_{M3} denote the noise contribution of M_1, M_2 and M_3, respectively, and F_R accounts for the noise of the lossy passive components L_p, L_g, L_s and R_1, and α_M and α_R are the temperature-dependent scaling factors that equal to one at RT. In our analysis, $\alpha_M = \alpha_N$ and $\alpha_R = T_a/296$ K. A cryogenic model for NT simulation is built based on the above analysis. The simulation noise results and contribution breakdown are illustrated in Fig. 2(c), in which L_p and L_g realize the optimum noise matching at ~6.5 GHz for CT operation with an ~1.5 GHz frequency offset compared with RT.

B. Input Return Loss Analysis

The reactance of input impedance seen into gate of M_1, i.e., $Z_{in}(\omega)$, must be conjugate to $Z_s(\omega)$, the impedance looking out from the gate, across the operating band to minimize the reflections, which can improve the quantum state readout fidelity. Due to R_S and noise matching inductors L_p and L_g, the imaginary part of $Z_s(\omega)$ increases ~1.5× from 4 to 8 GHz, while the real part increases slightly (~10%). Thus, to realize a broadband matching, the effective capacitance ($C_{eff}(\omega)$) loaded $Z_s(\omega)$ must be frequency-dependent and decrease rapidly with frequency so as to track the Im[$Z_s(\omega)$] profile across the band (see Fig. 2(d)). In this design, such a characteristic is realized through the frequency-dependent Miller effect of M_1. The drain node of M_1 is loaded by a CG stage with a parallel LC tank load ($Z_{L1}(\omega)$), whose resonant frequency (ω_R) is designed at lower than the lower edge frequency (ω_L) of the operating band. Thus, for the operating frequency $\omega>\omega_L$, the magnitude of $Z_{L1}(\omega)$ decreases monotonically till the upper edge ω_H. Then, $Z_{L1}(\omega)$ is almost linearly scaled by $g_{m2}r_{o2}$ [8], reaching $Z_D(\omega)$, serving as the load of M_1. Consequently, the gain of M_1 $|A_{M1}(\omega)|$ also decreases with frequency from ω_L to ω_H. The effective input capacitance that includes the Miller effect becomes

$$C_{eff}(\omega) = C_{gs1} + (1 + |A_{M1}(\omega)|)C_{gd}, \tag{3}$$

which decreases with frequency and can neutralize the reactance of $Z_s(\omega)$ in a wideband. In conclusion, by designing $\omega_R<\omega_L$ and leveraging a sufficiently large C_{gd} along with an appropriate

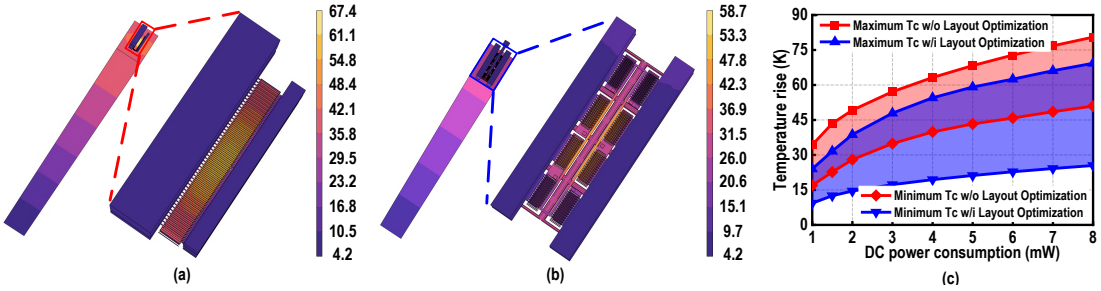

Fig. 3. Channel temperature distributions of layout at CT (a) without and (b) with optimization. (c) The channel temperature rise versus power consumption.

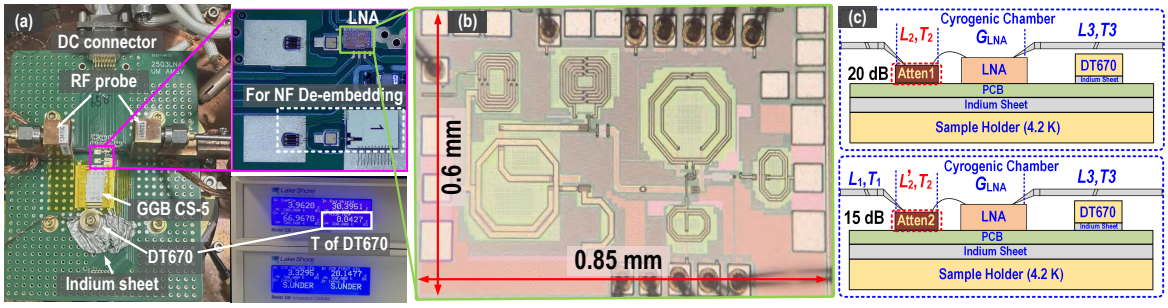

Fig. 4. (a) Cryogenic testing setup for the LNA. (b) Microphotograph of the LNA die. (c) Diagram of the dual-attenuation cryogenic noise temperature test setup.

$g_{m2}r_{o2}$, frequency-dependent capacitive load of M_2 can be effectively replicated to the input. Furthermore, the real part of $Z_{in}(\omega)$ also exhibits an increasing trend with frequency, consistently aligning with $\mathrm{Re}[Z_s(\omega)]$. This concurrent tracking of both real and imaginary components ensures robust conjugate matching throughout the entire operating range.

C. Self-heating effect simulation and suppression

To enable pre-silicon optimization, we applied an electro-thermal simulation flow using CAD tools to quantify real channel temperature due to SHE and guide layout design for M_1, which is designed with a relatively large dimension (W_{eff}= 208 um) and large power consumption. Two layout styles are evaluated: a deliberately extreme layout with 104-finger single-cell (Layout-1) and a distributed 2×4 sub-cell array (Layout-2) (Fig. 3(a)(b)). Simulation results at 4.2 K ambient temperature with 4.1 mW dissipation show that Layout-2 effectively decouples heat sources, reducing the peak temperature rise from 63.2 to 54.5 K (Fig. 3(c)). Notably, Layout-2 achieves an additional 18.5 K reduction in channel margin temperature rise compared to Layout 1, demonstrating its superior thermal uniformity and effectiveness in suppressing SHE. In addition, the source of M_1 is designed with multiple bondwires connected to the ground of the PCB, providing excellent thermal conduction to further reduce the channel temperature.

III. Measurement Results

A. S-parameters

Fig. 4(a) shows the cryogenic testing setup (inside a Lakeshore cryogenic probe station) for the LNA, where the DT670 is installed for monitoring the temperature of the PCB and GGB CS-5 is used for on-wafer calibration. The die microphotograph is drawn in Fig. 4(b). Fig. 5(a-c) show the simulated and measured S-parameters at RT and CT. At RT, the circuit achieves a maximum gain of 25.5 dB, and the input return loss S_{11} remains below -10 dB from 3.4 to 8.4 GHz. Within this operating range, the gain varies from 20.3 to 25.5 dB. The output port is well matched, and S_{22} is below −15 dB owing to the output buffer. At 3.9 K of the station's sample holder (8.8 K for PCB), the peak gain increases to 32.1 dB under 17.1 mW dissipation due to the increase in transconductance, and S_{11} remains below -12.7 dB from 4 to 8 GHz.

B. Noise Measurement

The measured NF at RT is shown in Fig. 5(d), ranging from 1.43 to 2.94 dB across the 3.4–8.4 GHz band. For cryogenic noise characterization, a cold 20-dB attenuator is generally inserted before the LNA to suppress uncertainty contributions (±0.175 K) from the temperature gradient along the input cable and the ENR uncertainty of the RT noise source [2], [6]. However, the attenuator's own temperature T_2 remains a critical uncertainty source, as a 1 K error in T_2 translates almost directly to a 1 K error in the extracted LNA noise temperature (NT_{LNA}) [6]. Since bulky commercial sensors often fail to capture the precise temperature of the attenuator core, this work proposes a dual-attenuator method to eliminate this uncertainty by performing two measurements with different insertion losses.

As illustrated in Fig. 4(c), the total measured NTs with attenuators $L_2 \approx 20$ dB and $L_2' \approx 15$ dB are expressed as:

$$NT_{tol1} \approx NT_1 + T_2(L_2 - 1)L_1 + NT_{LNA}L_1L_2 \quad (4)$$

$$NT_{tol2} \approx NT_1 + T_2(L_2' - 1)L_1 + NT_{LNA}L_1L_2' \quad (5)$$

where NT_1 and L_1 represent the noise temperature and loss of the input cable, respectively. By solving (4) and (5), the

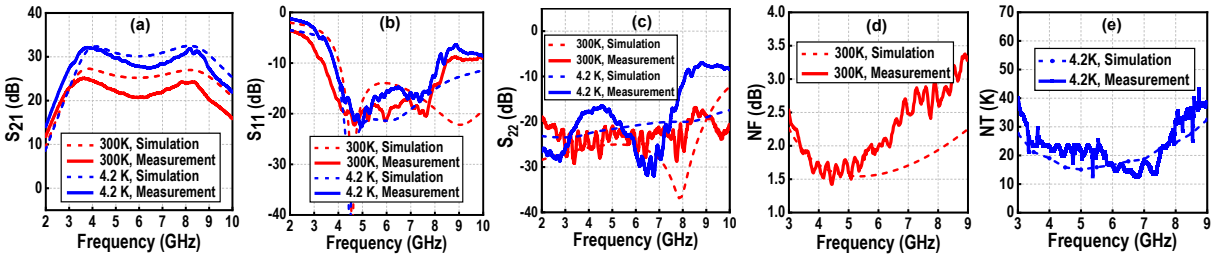

Fig. 5. Simulated and measured S-parameters of the LNA at RT and CT. (a) S_{21}. (b) S_{11}. (c) S_{22}. Simulated and measured noise performance at (d) RT. (e) CT.

TABLE I
PERFORMANCE SUMMARY AND COMPARISON WITH PRIOR ART CRYO-LNAS

	Frequency (GHz)	Temperature (K)	Power (mW)	Gain (dB)	NT (K)	S_{11} (dB)	Technology
[1] JSSC'17	0.1~0.5	4.2	91	50~58	6.8~62.7	-7~-3	160 nm CMOS
[9] ISSCC'24	0.1~2.2	4.2	10.3	42	15	<-10	40 nm CMOS
[11] IMS'22	3.9~5.3	16	10.6	35	10.2	<-5	65 nm CMOS
[10]JSSC'24	4.5~7	4.2	N/A	52	77.6	<-15	65 nm CMOS
[12] RFIC'22	5.8~8.3	4.1	2.57	13.4	37.6~40.7	<-5	14 nm FinFET
[3] RFIC'22	4.2~9.2	16	21	31.4~34.7	4.7~20.7	-8.9~-5.6	22 nm FDSOI
[4] JSSC'22	6~8	4.2	4.2	N/A	28~50.7	<-4	28 nm CMOS
[2] JSSC'21	4.6~8.2	4.2	39	39.2~44.8	15.8~48.2	-26~-5.8	40 nm CMOS
This Work	4~8	8.8	17.1	27.5~32.1	12.3~31	-22.6~-12.7	28 nm CMOS

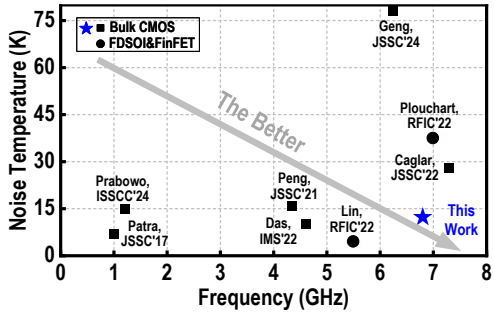

Fig. 6. Benchmarks of Noise Temperature with cryogenic LNAs.

noise temperature of LNA and the attenuator temperature are decoupled as:

$$NT_{\mathrm{LNA}} = \frac{(NT_{\mathrm{tol2}} - NT_1)(L_2 - 1) - (NT_{\mathrm{tol1}} - NT_1)(L_2' - 1)}{L_1(L_2 - L_2')} \tag{6}$$

$$T_2 = \frac{(NT_{\mathrm{tol1}} - NT_1)L_2' - (NT_{\mathrm{tol2}} - NT_1)L_2}{L_1(L_2 - L_2')}. \tag{7}$$

Using this method, the extracted NT_{LNA} at CT is measured between 12.3 K and 31 K (see Fig. 5(e)), showing good agreement with simulations based on [6] and the electro-thermal analysis in Section II-C. The calculated attenuator temperature T_2 is 10.2 ± 0.2 K. Despite careful thermal anchoring, the DT-670 sensor indicates a lower temperature of 8.8 K (see Fig. 4(a)), highlighting the necessity of the proposed dual-attenuator method for accurate cryogenic noise extraction. Table I and Fig. 6 summarize the measured performance and compare the proposed design with state-of-the-art wideband cryogenic LNAs [1]–[4], [6], [9]–[12].

IV. CONCLUSION

By leveraging SHE-aware electro-thermal simulation and layout optimization, the proposed cryo-LNA achieves competitive cryogenic noise performance. An electro-thermal co-simulation methodology is introduced to quantify self-heating and guide layout optimization, and a dual-attenuator cryogenic noise measurement technique is proposed to eliminate uncertainty caused by attenuator temperature.

REFERENCES

[1] B. Patra, R. M. Incandela, J. P. Van Dijk, H. A. Homulle, L. Song, M. Shahmohammadi, R. B. Staszewski, A. Vladimirescu, M. Babaie, F. Sebastiano *et al.*, "Cryo-cmos circuits and systems for quantum computing applications," *IEEE Journal of Solid-State Circuits*, vol. 53, no. 1, pp. 309–321, 2017.

[2] Y. Peng, A. Ruffino, and E. Charbon, "A cryogenic broadband sub-1-db nf cmos low noise amplifier for quantum applications," *IEEE Journal of Solid-State Circuits*, vol. 56, no. 7, pp. 2040–2053, 2021.

[3] B. Lin, H. Mani, P. Marsh, R. Al Hadi, and H. Wang, "A 4.2-9.2 ghz cryogenic transformer feedback low noise amplifier with 4.5 k noise temperature and noise-power matching in 22nm cmos fdsoi," in *2022 IEEE Radio Frequency Integrated Circuits Symposium (RFIC)*. IEEE, 2022, pp. 23–26.

[4] A. Caglar, S. Van Winckel, S. Brebels, P. Wambacq, and J. Craninckx, "Design and analysis of a 4.2 mw 4 k 6–8 ghz cmos lna for superconducting qubit readout," *IEEE Journal of Solid-State Circuits*, vol. 58, no. 6, pp. 1586–1596, 2022.

[5] L. Belostotski and J. W. Haslett, "Sub-0.2 db noise figure wideband room-temperature cmos lna with non-50 ω signal-source impedance," *IEEE Journal of Solid-State Circuits*, vol. 42, no. 11, pp. 2492–2502, 2007.

[6] S. Das, S. Raman, and J. C. Bardin, "Experimental characterization of the mosfet fano factor at cryogenic temperatures for accurate cryo-cmos rf modeling," *IEEE Transactions on Microwave Theory and Techniques*, 2025.

[7] X. Chen, "Noise characterization and modeling of nanoscale mosfets," Ph.D. dissertation, University of Waterloo, 2017.

[8] B. Razavi, *Design of analog CMOS integrated circuits*. , 2005.

[9] B. Prabowo, O. Pietx-Casas, M. A. Montazerolghaem, G. Scappucci, L. M. Vandersypen, F. Sebastiano, and M. Babaie, "29.3 a cryo-cmos receiver with 15k noise temperature achieving 9.8 db snr in 10μs integration time for spin qubit readout," in *2024 IEEE International Solid-State Circuits Conference (ISSCC)*, vol. 67. IEEE, 2024, pp. 474–476.

[10] Y. Geng, H. Lin, B. Wang, and C. Wang, "Cryo-cmos dual-qubit homodyne reflectometer array with degenerate parametric amplification," *IEEE Journal of Solid-State Circuits*, vol. 59, no. 10, pp. 3290–3306, 2024.

[11] S. Das, S. Raman, and J. C. Bardin, "Design and implementation of a 3.9-to-5.3 ghz 65 nm cryo-cmos lna with an average noise temperature of 10.2 k," in *2022 IEEE/MTT-S International Microwave Symposium-IMS 2022*. IEEE, 2022, pp. 719–722.

[12] J.-O. Plouchart, D. Yilma, J. Timmerwilke, S. Chakraborty, K. Tien, A. Valdes-Garcia, and D. Friedman, "A 2.57 mw 5.9-8.4 ghz cryogenic finfet lna for qubit readout," in *2022 IEEE Radio Frequency Integrated Circuits Symposium (RFIC)*. IEEE, 2022, pp. 27–30.