UCLA
UCLA Electronic Theses and Dissertations

Title
Niobium-based Superconducting Silicon Interconnect Fabric for Future Cryogenic Applications

Permalink
https://escholarship.org/uc/item/600993hd

Author
Yang, Yu-Tao

Publication Date
2022

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA

Los Angeles

Niobium-based Superconducting Silicon Interconnect Fabric for Future Cryogenic Applications

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Electrical and Computer Engineering

by

Yu-Tao Yang

2022

© Copyright by

Yu-Tao Yang

2022

ABSTRACT OF THE DISSERTATION

Niobium-based Superconducting Silicon Interconnect Fabric for Future Cryogenic Applications

by

Yu-Tao Yang

Doctor of Philosophy in Electrical and Computer Engineering

University of California, Los Angeles, 2022

Professor Subramanian Srikanteswara Iyer, Chair

To build up large-scale quantum systems, four challenges in the existing hardware realization need to be addressed: (1) I/O management, (2) latency and phase matching in a large system footprint, (3) the heat and power dissipation as the qubit count scales up, and (4) transmission of delicate quantum information. To minimize these issues, there are efforts using different architectures, such as CryoCMOS, 3D stacking, and Josephson-Junction-based (JJ-based) approach. Superconducting Silicon Interconnect Fabric (Superconducting-IF) is yet another promising approach, which is an advanced fine-pitch and wafer-scale cryogenic packaging platform. The proposed architecture, which aims to co-integrate superconducting qubit control/readout with qubits in a compact way, and the demonstration of the Superconducting-IF are detailed. The key enabler of Superconducting-IF is the low-temperature assembly technology, the Au interlayer bonding method, which is demonstrated to be fine-pitch (≤ 10

μm), mechanically robust (> 30 MPa), electrically reliable, and quantum-compatible (< 150°C). The Au interlayer with Nb superconducting interconnects has a transition temperature at 9 K and is demonstrated to be a low-temperature-compatible process having a high critical current for integrated superconducting processing applications. On the Superconducting-IF platform, heterogeneity and flexibility are both demonstrated regarding the die sizes, the inter-die spacing, and the configuration. For future transmission of delicate quantum information, Superconducting RF optimization, including insertion loss and crosstalk, is implemented. All the simulated and measured results on short superconducting links (≤ 500 μm) with 2 and 5 μm line/space are far below the quantum limitation, meaning the superconducting interconnects are suitable to carry future inter-dielet delicate quantum communication.

This dissertation achieves the following intellectual contributions: (1) the first combination of heterogeneous integration and advanced packaging with quantum applications, (2) the first proposal of the architecture of integrated system-on-wafer quantum hardware, (3) the first demonstration of the Au interlayer bonding to be fine-I/O pitch, mechanically robust, electrically reliable, and quantum compatible, and (4) the first achievement of low-loss, low-crosstalk, and 20 GHz-broadband RF capability on superconducting short links through advanced packaging.

To sum up, it is promising to use the Superconducting-IF for future large-scale quantum systems to realize the full strength of quantum computing.

The dissertation of Yu-Tao Yang is approved.

Chee Wei Wong

Dwight Christopher Streit

Kang L. Wang

Subramanian Srikanteswara Iyer, Committee Chair

University of California, Los Angeles

2022

This dissertation is dedicated to my parents.

Table of Contents

Chapter 1 Introduction 1

1.1 Quantum computing 1

1.2 Challenges in quantum computing 5

1.3 Objective of this Dissertation 10

1.4 Organization of this Dissertation 11

Chapter 2 State of the art in the architectures of superconducting quantum computing 12

2.1 Conventional room-temperature CMOS architecture 12

2.2 Cryogenic CMOS (CryoCMOS) architecture 15

2.3 3D stacking architecture 20

2.4 Josephson-Junction-based (JJ-based) architecture 21

2.5 Summary 27

Chapter 3 Overview of packaging technologies 28

3.1 Packaging technology evolution 28

3.1.1 Conventional System Integration 28

3.1.2 Advanced System Integration 31

3.2 Packaging in quantum application 36

3.3 Silicon-Interconnect Fabric (Si-IF) introduction 39

3.4 Summary 42

Chapter 4 Concept of integrated quantum system on cryogenic Si-IF – Superconducting Silicon Interconnect Fabric (Superconducting-IF) 43

4.1 Superconducting-IF introduction 43

4.2 Proposed qubit integrated system on Superconducting-IF 46

4.3 Comparison with existing scheme 54

4.4 Summary 57

Chapter 5 Wafer-level Superconducting-IF realization 58

5.1 Deposition and fabrication of Nb interconnects 59

5.1.1 Nb oxidation issue and top passivation layer 59

5.1.2 Selection of proper flip-chip bonding medium 62

5.2 Au interlayer technology on Nb interconnects 65

5.2.1 Pre-bonding check 65

5.2.2 Au-Au bonding Design of Experiment (DOE) 69

5.2.3 Electrical tests of large-pitch Kelvin structure 74

5.2.4 Electrical tests of fine-pitch daisy chain structure 82

5.3 Collaboration with MIT Lincoln Lab 98

5.4 Wafer-level assembly on Superconducting-IF 103

5.5 Summary 106

Chapter 6 RF characterization on Superconducting-IF 108

6.1 Superconducting coplanar waveguide simulation 110

6.1.1 Simulation assumption & approximation 110

6.1.2 Simulation model establishment 112

6.1.3 Simulation results & analysis 114

6.2 RF characterization of the superconducting interconnects on the Superconducting-IF.... 120

6.2.1 Dilution fridge setup 120

6.2.2 PCB sample holder design 125

6.2.3 Passive superconducting dielets and -IF deisgn and fabrication 127

6.2.4 Pre-check and measurement results and discussion 132

6.3 Passive die with resonator integrated on Superconducting-IF 140

6.4 Summary 147

Chapter 7 Conclusions 149

7.1 Conclusions 149

7.2 Outlook.. 152

References 155

Table of Figures

Figure 1: Existing control/readout scheme in a quantum computer. Reprinted with permission © 2018 American Physical Society [17] 12

Figure 2: Google CryoCMOS architecture. Reprinted with permission © 2019 IEEE [20]–[22]15

Figure 3 Microsoft CryoCMOS architecture. Reprinted with permission © 2021 Nature [23] .. 16

Figure 4 Intel CryoCMOS architecture. Reprinted with permission © 2021 IEEE [24] 17

Figure 5 Intel CryoCMOS architecture (cont.). Reprinted with permission © 2021 IEEE [24] . 18

Figure 6 3D packaging on a qubit chip. Reprinted with permission [27] 20

Figure 7. Architecture for qubit controlling with JJ-based devices. Reprinted with permission © 2019 American Physical Society [37] 23

Figure 8 SEM image of integrated control with a qubit. Reprinted with permission © 2019 American Physical Society [37] 24

Figure 9 Qubit manipulation (a) JJ-based SFQ pulse (b) coherent control of qubits with time-domain pulse interval matching the qubit resonant frequency (c) discrete rotation of the qubit state vector in the Bloch sphere. Reprinted with permission © 2019 American Physical Society [37] 25

Figure 10 solder extrusion. Reprinted with permission © 2015 IEEE [48] 29

Figure 11. Apple A15 processor System on Chip (SoC) design with multiple function blocks. (Picture source [51] 31

Figure 12. Apple M1 Ultra on Si interposer.(picture source [69]) 34

Figure 13. Wiring bonding to qubit chip. Reprinted with permission © 2012 AIP [77], [78] 36

Figure 14. Flip-chip packaged qubit chip through the interposer technology. Reprinted with permission © 2019 nature [12] 37

Figure 15. 10 um interconnect pitch on Si-IF. Reprinted with permission © 2021 IEEE [89] 39

Figure 16 ~ 370 dies (1×1 mm^2 to 5×5 mm^2) at ≤ 100μm inter-dielet spacing on a 100 mm Si-IF. Reprinted with permission © 2018 IEEE [90] 40

Figure 17. Superconducting-IF integration scheme with fine-pitch interconnects and proximal inter-die spacing. © IOP Publishing. Reproduced with permission. All rights reserved. 45

Figure 18. Thermal design with qubits based on the generated power of the control/readout circuits [94]. 48

Figure 19 architecture comparison between this work and the convention method [111] 55

Figure 20. EDX elemental composition analysis on (a) a source pellet and (b) blanket Nb structure on a Si wafer. The oxygen content in the Nb film is twice of the oxygen content in the pellet... 61

Figure 21. Preliminary Tc test of Nb/Ir samples with the protection of Ir against Nb oxidation during heating up. The Tc is maintained at 9K without significant degradation......................... 62

Figure 22. Nb-Nb bonding test with the control and experiment set. No bonding reaction happens in both conditions.. 63

Figure 23. Successful In-Sn-Au solid-liquid interdiffusion bonding of superconducting dielets: (a) top view of the samples, (b) SEM image of the cross section of In-Sn-Au solid-liquid interdiffusion bonding at bonding interface.. 64

Figure 24. The worst-case scenario of Au atom diffusion into Nb film below 150 ℃ without Ir film in detween. © 2020 IEEE... 67

Figure 25. Adhesion test between the Au and the Nb/Ir film through tape-adhesive test. © 2020 IEEE.. 67

Figure 26. Surface roughness study: (a) surface roughness measurement on the samples after

each film deposition, (b) Surface scanning image on a sample with 50 nm Ti/500 nm Au (a deposition rate of 2 Å/s). The grain size of Au is around 50-100 nm wide. © IOP Publishing. Reproduced with permission. All rights reserved. 69

Figure 27 Images on the substate side after shear test with various bonding temperature in DOE experiments 70

Figure 28. Images on the substate side after shear test with various bonding duration in DOE experiments 71

Figure 29. Optimization of the Au-Au bonding conditions with respect to shear strength tests (2 x 2 mm^2 dies): (a) bonding duration (b) bonding pressure (c) bonding-interface temperature (d) optimized parameters for low-temperature Au-to-Au bonding DOE results (e) image during shear test (f) the schematic blanket structure for shear tests with a die area of 2 × 2 mm^2. © 2020 IEEE 72

Figure 30. Mechanical strength comparison between different low temperature integration technologies and Au-Au LT TCB in this work. © IOP Publishing. Reproduced with permission. All rights reserved. 73

Figure 31. Low-temperature measurement results: (a) the transition temperature of Nb before any process, (b) Critical field of Nb before any process, (c) Tc check on Nb/Ir samples after one month, three month, and one year storage, (d) Nb Tc check after numerous processes using blanket samples, (e) Room-temperature-measurement image in a manual probe station under microscope before sending samples to PPMS, (f) The image of Physical Property Measurement System (PPMS) [123]. © 2020 IEEE 74

Figure 32. Details of the Kelvin structure: schematic view of the Nb/Ir interconnects and Au bonding sites on the (a) die side and (b) substrate side. Microscopic images of the structures

fabricated on (c) die and (d) substrate with the interconnect, Au bonding sites, and supporting structures after dicing. (e) The schematics of the continuous electrical path after flip-chip bonding (f) Microscopic top-view image of the complete Kelvin structure with a die flip-chip bonded to a substrate with four probing pads for four-probe measurement. (g) The image after the bonded die is sheared by the test bonder. © 2020 IEEE 76

Figure 33. The fabrication process of the Kelvin structure in both the die and the substrate sides for electrical characterization: 500 nm thermal oxide is first grown, followed by Nb/Ir deposition in an ultra-high vacuum system without breaking the vacuum. Patterns are defined through lithography using AZ 5214 photoresist. The whole wafers are etched using Cl dry etching. Au thin film is deposited and patterned at the bonding sites through the AZ 5214 EIR lift-off process. © 2020 IEEE 77

Figure 34. (a) Resistance-temperature measurement of the Kelvin structure (200-µm-width Nb interconnects with Au joints) before and after 10 times of temperature cycling between 300 K and 2 K. The Tc of the whole structure is at 9K and the resistance below Tc is around 0.02 Ω. © IOP Publishing. Reproduced with permission. All rights reserved, (b) Voltage-Current measurements of the Kelvin structure before and after temperature cycling between 300 K and 2 K.. 79

Figure 35. (a) Schematic view of flip-chip bonded Kelvin structure. (b) Zoom-in image of the bonded structure with an Au island structure in the die side. (c) Zoom-in image of the bonded structure with a continuous Au film structure in the die side. © IOP Publishing. Reproduced with permission. All rights reserved. 81

Figure 36. (a) R-T curve with two die side structures: the Au island structure and the continuous Au structure. (b) Zoom in image of the highlight part in (a). More detail information is in the

text. © IOP Publishing. Reproduced with permission. All rights reserved. 81

Figure 37. (a) Superconducting-IF side before flip-chip bonding. The bonding area is 2 × 2 mm^2. (b) Die side before flip-chip bonding. (c) Zoom in picture of (b). The pad side is 17 µm × 7 µm (d) Top view image of Superconducting-IF after shear test is made on a bonded die. Over 100 N is required to shear the die. This image shows a precise die-to-substrate alignment. © IOP Publishing. Reproduced with permission. All rights reserved.. 83

Figure 38. (a) Image of three dies assembled on Superconducting-IF and measured in a row. (b) Schematic view of the cross-section view. The pad size is 17 µm and the bonding pitch is 10 µm. © IOP Publishing. Reproduced with permission. All rights reserved 84

Figure 39. (a) R-T measurement of Ti/Au interconnects and joints at low temperature. (b) zoom-in of (a) between 2 K to 40 K. (c) R-T curve after 10 times of temperature cycling between 300 K and 2 K. (d) zoom-in of (c) between 2 K to 40 K. (e) samples on measurement a puck. © IOP Publishing. Reproduced with permission. All rights reserved.. 86

Figure 40. (a) Cross section image on the Au-to-Au bonding sample with fine-pitch demonstration. (b) Zoom-in of (a). © 2020 IEEE. (c) Cross-section image on the bonding interface before temperature cycling. © 2020 IEEE. (d) Cross-section image on the bonding interface after temperature cycling on another sample. 87

Figure 41. Image of different bonding areas of Au daisy chain samples on the die side with a die size of 5×5 mm^2: (a) 2×2 mm^2 (b) 3×3 mm^2 (c) 4 × 4 mm^2 (d) 5 × 5 mm^2 (e)&(f) zoom- in image of alignment keys on (d). (g) zoom in image of bonding pads. The die pad size is 12 × 5 µm^2 and the bonding pitch is 10 µm. 88

Figure 42. Image of different bonding areas of Au daisy chain samples on the substrate side with a die size of 5×5 mm^2: (a) 2×2 mm^2 (b) 3×3 mm^2 (c) 4 × 4 mm^2 (d) 5 × 5 mm^2 (e)&(f) zoom- in

image of alignment keys on (d). (g) zoom in image of bonding pads. The pad size is 17 × 7 μm^2 and the bonding pitch is 10 μm. 89

Figure 43. Image of different bonding areas of Au daisy chain samples after flip-chip bonding with a die size of 5×5 mm^2: (a) 2×2 mm^2 (b) 3×3 mm^2 (c) 4 × 4 mm^2 (d) 5 × 5 mm^2 (e) zoom- in image of alignment keys on (d). (f) Image for a good alignment. © 2020 IEEE 90

Figure 44. (a) Resistance-temperature electrical measurement on different die bonding area (2 × 2 mm^2, 3 × 3 mm^2, 4 × 4 mm^2, and 5 × 5 mm^2) down to 2 K, (b) Zoom-in (a) between 2 K to 50 K. (c) Bonded samples on a measurement puck. © 2021 IEEE. 92

Figure 45. Electrical measurements of the Au interlayer (temperature cycling from 300 K to 2 K on various sizes of dies: 2×2 mm^2, 3×3 mm^2, 4×4 mm^2, 5×5 mm^2. 93

Figure 46. Electrical measurements of superconducting Nb interconnects with the Au interlayer (temperature cycling from 300 K to 2 K). Tc is around 8 K. © 2021 IEEE. 95

Figure 47. Critical current measurements at 4.2K without any applied magnetic field. (a) In the superconducting interconnects with the Au interlayer, the critical current is 30 mA (critical current density: 4MA/ cm^2, 14% of the theoretical value) (b) Nb interconnects control experiment without the Au interlayer flip-chip bonding. The critical current is 100 mA (critical current density: 13.2MA/ cm^2, 48% of the theoretical value). © 2021 IEEE. 95

Figure 48. Cross-section image of bonded superconducting dies: (a) 10-μm-pitch Au interlayer on superconducting interconnects, (b) Zoom-in at the bonding interface of (a), (c) Bonded dies on the Superconducting-IF. © 2021 IEEE. 96

Figure 49. MIT Lincoln lab samples: (a) 10-μm pitch of Au bonding pads with a diameter of 6.5 μm and 2 μm wide Nb interconnects, (b) 15-μm pitch of Au bonding pads with a diameter of 10 μm and 2 μm wide Nb interconnects, (c) 25-μm pitch of Au bonding pads with a diameter of 15

μm and 6 μm wide Nb interconnects, (d) Top view of the dies and the substrates with alignment marks for die-to-IF flip-chip bonding. 98

Figure 50. MIT Lincoln lab samples: (a) Top view of 5 × 5 mm^2 daisy-chain bonded dielets with pitches of 10/15/25 μm, (b) Cross-section scheme of pads on the die 99

Figure 51. (a) Configuration of probing pads after a 5 X 5 mm^2 dielet flip-chip bonded Si substrate. (b) Image of four-probe measurements at 300 K under microscope. 101

Figure 52. MIT Lincoln lab samples wire-bonded to designated PCB for low-temperature measurements: (a) 10-μm pitch of bonded dielets, (b) 15-μm pitch of bonded dielets, (c) 25-μm pitch of bonded dielets 101

Figure 53. (a) R-T measurements of the dies with 10-μm, 15-μm, and 25-μm pitches from 300K cooling down to 2K. (b) zoom-in of (a) from 40 K to 2 K. (c) One set of the measurement results probing all pads on one die with the 10-μm pitch: all of the results show superconducting behavior 102

Figure 54. (a) SoW demonstration of integrated superconducting (SC) surrogating dielets with die sizes from 2 mm by 2 mm to 5 mm by 5mm. (b) Zoom-in image of highlight area in Fig. 4. The SC surrogating dies are assembled through the Au interlayer technology with an inter-dielet spacing of 60 μm. © 2021 IEEE 104

Figure 55. (a) Detail of the assembly process. In the Au interlayer technology, one bonding cycle is 8.5 second (die transfer of 2 seconds from a tray to the bonding head, die-to-wafer alignment of 2.5 seconds, temperature ramping up and down of 2.5 seconds, and bonding of 1.5 seconds). (b) Pick & place bonding tool. © 2021 IEEE 105

Figure 56. Simulation model: (a) Isomeric view of GSG 125 μm link length with a L/S 2/2 μm.

(b) GSG cross section at the coplanar waveguide part with material stack from bottom to top: 500 μm Si (not to full scale) , 500 nm SiO_2, 200 nm Nb/Ir. (c) GSG cross section at the bonding part with a material stack from bottom to top: 500 μm Si (not to full scale), 500 nm SiO_2, 200 nm Nb/Ir, 50 nm Ti, 1000 nm Au, 50 nm Ti, 200 nm Nb/Ir. (d) Isomeric view of GSSG 125 μm link length with a L/S 2/2 μm. (e) GSSG cross section at the coplanar waveguide part with a material stack from bottom to top: 500 μm Si (not to full scale) , 500 nm SiO_2, 200 nm Nb/Ir. (f) GSSG cross section at the bonding part with material stack from bottom to top: 500 μm Si (not to full scale) , 500 nm SiO_2, 200 nm Nb/Ir, 50 nm Ti, 1000nm Au, 50 nm Ti, 200 nm Nb/Ir. © 2022 IEEE 113

Figure 57. Simulated insertion loss results (a) L/S 2/2 μm with link lengths of 125 μm, 500 μm, and 1750 μm. (b) L/S 5/5 μm with link lengths of 125 μm, 500 μm, and 1750 μm. © 2022 IEEE 114

Figure 58. Results of simulated insertion loss with 125 and 1750 μm link lengths and different materials (all Cu, all superconductor, and Au interlayer with superconducting wires in this thesis). © 2022 IEEE 115

Figure 59. Simulated crosstalk results (a) L/S 2/2 μm with link lengths of 125 μm, 500 μm, and 1750 μm. (b) L/S 5/5 μm with link lengths of 125 μm, 500 μm, and 1750 μm. © 2022 IEEE . 117

Figure 60 Crosstalk versus the length of interdie length 118

Figure 61. A series of cabling and connectors to connect from VNA 8720es at room temperature to PCB at 4 K: (1) VNA terminated with SMA female, (2) coaxial cables with 3.5 mm male and 2.4 mm female, (3) adaptors with 2.4 mm male and 2.4 mm male, (4) fridge interface [142] with 2.4 mm female at room temperature and 2.4 mm male at 4 K, (5) three-inch coaxial cables with

2.4 mm female and 3.5 mm male, and (6) SMA female on PCB. 122

Figure 62. Main dilution refrigerator used in this thesis: (a) vacuum preparation of the cylinder before loading the sample to fridge chamber. (b) Cylinder inside the fridge chamber. (c) Cylinder unloading from the chamber. 123

Figure 63. BLUEFORS LD dilution refrigerator: (a) during operation, (b) chamber inside, and (c) sample loading... 124

Figure 64. (a) Oxygen-Free High thermal Conductivity (OFHC) Cu holder with screw hole to fix PCB holder. (b) PCB sample holder on Cu holder. 124

Figure 65. BLUEFORS LD dilution fridge in a test run with progressively reduced temperature in each cryostat. 125

Figure 66. The design specs of the PCB carrier with 50 Ω coplanar waveguides with ground plane in the KiCAD PCB EDA.. 125

Figure 67. Two-layer FR-4 PCB carrier with a sample mounting area, signal path and wire-bonding 50 Ω traces, out-of-plane SMA soldering area, and an universal ground. Test die and IF design. 126

Figure 68. Insertion loss characterization: (a) three lengths of links (125 μm, 500 μm, and 1750 μm) with two L/S (2/2 and 5/5 μm) are cascaded on 2 X 2 mm^2 dielets, (b) zoom-in of (a), (c) flip-chip bonded multiple 2 X 2 mm^2 dielets, (d) zoom-in of (c). 128

Figure 69. FEXT characterization: (a) three lengths of links (125 μm, 500 μm, and 1750 μm) with two L/S (2/2 and 5/5 μm) are cascaded on 2 X 2 mm^2 dielets, (b) zoom-in of (a), (c) flip-chip bonded multiple 2 X 2 mm^2 dielets, (d) zoom-in of (c). 129

Figure 70. NEXT characterization: (a) three lengths of links (125 μm, 500 μm, and 1750 μm) with two L/S (2/2 and 5/5 μm) are cascaded on 2 X 2 mm^2 dielets, (b) zoom-in of (a), (c) flip-

chip bonded multiple 2 X 2 mm^2 dielets, (d) zoom-in of (c). 130

Figure 71. (a) the sheet resistance measurement on a blanket wafer with Ti/Pd 3/25 nm. (b) resistance drop as the temperature cools down to 2 K. 131

Figure 72. Tweaking of the pattern resistance of Ti/Pd with respect to different Pd thickness of 25/50/75/100 nm. 132

Figure 73. Preliminary DC electrical test result down to 2 K: (a) an insertion loss sample with Tc at 9 K and clear resistance drop below Tc, (b) a XT sample with Tc at 9K and clear resistance drop below Tc. 133

Figure 74. Preliminary DC electrical test result down to 0.25 K: (a) an insertion loss sample with Tc at 9K and clear resistance drop below Tc, (b) zoom-in of (a) between 0.25 K and 0.6 K. ... 133

Figure 75. Image of the XT samples. A cryogenic adhesion problem is observed between the Ti film and the Nb/Ir film in the links with L/S 2/2 μm. © 2022 IEEE 134

Figure 76. Image of the XT samples: (a) at the GSSG bonding site, (b) GSSG bonding site with the surrounding supporting structure, (c) zoom-in of the supporting bonding structure. 134

Figure 77. Image of the XT samples: (a) layer-to-layer lithography alignment overlay <1 μm, (b) assembly mechanical bonding overlay <1 μm. 135

Figure 78. (a) image of flip-chip bonded samples, (b) image of flip-chip bonded samples adhered on PCB and with wire-bonding. © 2022 IEEE 136

Figure 79. (a) PCB carrier installed on the fridge cylinder, (b) front side image. (c) back side image. © 2022 IEEE 137

Figure 80. (a) 4 K measurement of insertion loss with samples of L/S 5/5 μm and link lengthes of 125/500/1750 μm. © 2022 IEEE (b) impedence mismatching in the samples due to fabrication process. 137

Figure 81. 4 K measurement of (a) FEXT (b) NEXT with samples of L/S 5/5 μm and link lengthes of 125/500/1750 μm. (c) Impedence mismatching due to fabrication process. © 2022 IEEE 138

Figure 82. Detailed die information with blanket Au bonding pads, 20 μm cylinder Au bonding pads, and 50 μm circular Au bonding pads. 140

Figure 83. (a) Cross section view for a 31.6 Ω ungrounded coplanar waveguide. The width of the center signal path and the gap between the signal path and the ground plane are 28 μm and 2 μm. (b) Top view of the ungrounded coplanar waveguide in a GSG configuration. (c) Simulation result shows a characteristic impedance of around 31.6 Ω. Each curve indicates the impedance in each port. 141

Figure 84. (a) Top view of the design quarter-wavelength-coplanar-waveguide impedance transformer on the Superconducting-IF. (b) Cross section view of the coplanar-waveguide with a 28-μm center signal path and a 2-μm gap between the signal path and the ground plane. 143

Figure 85. Simulation results of S parameter with the varying signal width from 24 to 30 μm. The resonance is at 13 GHz. 143

Figure 86. Simulation results of S parameter with varying gap width from 1.8 to 2.6 μm. The resonance is at 13 GHz. 144

Figure 87. (a) Fabricated Superconducting-IF with quarter-wavelength-coplanar-waveguide impedance transformers. (b) SeeQC dies. (c) flip-chip bonded Superconducting-IF 144

Figure 88 (a) Optical and (b) SEM images of the quarter-wavelength-coplanar-waveguide impedance transformers on Superconducting-IF. 145

Figure 89. DC measurements of flip-chip-bonded Superconducting-IF with blanket Au pads. Tc exists in all the samples at 9K with smooth curves 146

Figure 90 DC measurements of flip-chip-bonded Superconducting-IF with 20-um Au pads. Tc exists in all the samples at 9K with smooth curves 146

Table of Tables

Table 1 Comparison between the conventional scheme and the work of Google, Microsoft, and Intel 19

Table 2 Estimated power consumption of JJ-based family operating at 4 K with a clock frequency of around 5GHz........ 22

Table 3. Estimated power consumption of JJ-based family operating at 20 mK with a clock frequency of around 5GHz........ 23

Table 4. Comparison between the existing flip-chip-demonstrated superconducting assembly methods 38

Table 5. Potential way to achieve one-million qubits integrated on the superconducting IF 53

Table 6 comparison between the concept proposed in this work, the conventional scheme and the work of Google, Microsoft, and Intel 55

Table 7. Comparison between this work and the existing flip-chip-demonstrated superconducting assembly methods. © IOP Publishing. Reproduced with permission. All rights reserved. 73

Table 8 Parameters for dies with bonding area of 2 × 2 mm^2, 3 × 3 mm^2, 4 × 4 mm^2, and 5 × 5 mm^2 91

Table 9. Comparison between dielets with 10-µm/15-µm/25-µm pitch of Au bonding pads...... 99

Table 10 Case study for Nb-based and Cu based interconnect........ 116

Acknowledgment

I would like to thank my advisor, Prof. Subramanian Iyer, for his generous support, guidance, and instruction over my Ph.D. career. This was a fantastic journey for me to explore a very wide range of research topics with Subu and think out of the box. This dissertation would not be possible without his vision.

I am very grateful for the help, encouragement, and discussion from my Ph.D. committee members: Prof. Chee Wei Wong, Prof. Dwight Streit, and Prof. Kang L. Wang.

I would like to thank Prof. Ni Ni, Dr. Chaowei Hu, Tiema Qian, and Jasmine Green for low-temperature DC equipment.

I would like to thank Prof. Kang L. Wang, Dr. Su Kong Chong, Dr. Gang Qiu, Dr. Christopher Eckberg, Dr. Yang Cheng, Dr. Peng Zhang, Hanshen Huang, and Haoran He for low-temperature dilution refrigerator RF measurements.

I would like to thank Prof. Chee Wei Wong, Dr. Yujin Cho, Murat Can Sarıhan, Jin Ho Kang, and Kai-Chi Chang for low-temperature dilution refrigerator RF measurements.

I would like to thank Dr. Rabindranath Das (MIT Lincoln Labs) and Dr. Deborah van Vechten (ONR) for valuable discussions.

I would like to thank SeeQC collaborators: Dr. Oleg Mukhanov, Igor Vernik, Naveen Katam, and Daniel Yohannes.

I enjoyed my graduate school years with many labmates in CHIPS. I am very fortunate to work as a member of the Silicon Interconnect Fabric team with Dr. Umesh Mogera, Dr. Siva Jangam, Dr. Niloofar Shakoorzadeh, Dr. Kannan K. T., Dr. Ujash Shah, Haoxiang Ren,

Krutikesh Sahoo, Golam Sabbir, Ankit Kuchhangi, Pranav Ambhore, Shelly Ku, Eric Sorensen, Arpan Dasgupta, Po-Chang (Steve) Shih, Ko-Ching Hou, Meng-Hsiang (Andy) Liu, Yu-Pei (Anne) Huang, Arshiya Vohra, Anthony Kim. I also thank my other CHIPS colleagues: Dr. Boris Vaisband, Dr. Samatha Benedict, Dr. Zhe (Frank) Wan, Dr. Saptadeep Pal, Guangqi Ouyang, Henry Sun, Randall Irwin, Siyun Qiao, Steven Moran, Premsagar Kittur, Xuefeng Gu, and Jonathan Cox, Sepideh Nouri and among many others. We are lucky to have Kyle Jung and Athena Sung-Miller in CHIPS, both of who takes care of everything we need in many aspects.

I would like to acknowledge the financial support from Office of Naval Research (ONR) Award N00014-18-1-2638 and in part by the Semiconductor Research Corporation (SRC), DARPA award N6600-12-C-4001, UCOP MRP-17-454999, and the UCLA CHIPS consortium, all of whom provided sufficient funding for my Ph.D. research. Special thanks to the technical support from ISNC, EICN, nanolab, ECE department staff.

I am extremely lucky to have too many good friends to support during my Ph.D. career: Dr. Huei- Bin Wang, Po-Han Chao, Yen-Ru Lin, Hao-Yu Chien, Dr. Siang Chen, Dr. Hung-Yu Yang, Dr. Ken-Fu Liang, Elizabeth Luo, Ivy Wong, Scott Huai and his family.

I would like to acknowledge publishers to grant copyright of a few images to this dissertation and parts of this dissertation are adapted from several publications during my doctoral studies.

Last but not least, I would like to thank the immense love and support from my parents and family in Taiwan.

Curriculum Vitae

EDUCATION:

2018 — Ph.D. candidate, Electrical and Computer Engineering, UCLA

2016 — 2017 Master of Science, Electronics Engineering, NCTU

2012 — 2016 Undergraduate Honors Program of Nano Science and Engineering, NCTU

PATENT:

Yu-Tao Yang, and Subramanian S. Iyer, "Apparatus and Method of Making Superconducting Silicon Interconnect Fabric" U.S. Provisional Application No. 63/032,980, filed on June 1, 2020.

JOURNAL ARTICLES

Yu-Tao Yang, Chaowei Hu, Peng Zhang, Niloofar Shakoorzadeh, Ni Ni, Kang L. Wang, and Subramanian S. Iyer, "Nb-based Superconducting Silicon Interconnect Fabric for cryogenic electronics," *Quantum Science and Technology,* vol. 6, no. 2, pp. 1-14, 2021.

Haoxiang Ren, **Yu-Tao Yang**, Guangqi Ouyang, and Subramanian S. Iyer "Mechanism and Process Window Study for Die-to-wafer (D2W) Hybrid Bonding," ECS Journal of Solid State Science and Technology, no. 6, 2021.

SELECTED CONFERENCE PRESENTATIONS

Yu-Tao Yang, Chaowei Hu, Peng Zhang, Haoxiang Ren, Ni Ni, Kang L. Wang, and Subramanian S. Iyer, "Fine-Pitch (≤ 10 μm) Nb-based Superconducting Silicon Interconnect Fabric for Large-Scale Quantum System Application," ***IEEE International Electron Devices Meeting (IEDM)*****,** San Francisco, Dec. 2021, pp. 873-876.

Yu-Tao Yang, and Subramanian S. Iyer, "Large-Scale Quantum System Design on Nb-based Superconducting Silicon Interconnect Fabric," *International Symposium on Electronic Quality Design*, virtual conference, USA, Apr. 202, pp. 315. (**Best paper award)**

Yu-Tao Yang, Haoxiang Ren,Chaowei Hu, Peng Zhang, Ni Ni, Kang L. Wang, and Subramanian S. Iyer, "Emerging Applications Using Niobium-based Superconducting Silicon Interconnect Fabric," *7th Annual Critical Material Conference (CMC)*, Arizona, Apr. 2022. (**Invited talk**)

Yu-Tao Yang, Haoxiang Ren, Su Kong Chong, Gang Qiu, Shu-Yun Ku, Yang Cheng, Chaowei Hu, Tiema Qian, Kuan-Neng Chen, Ni Ni, Kang L. Wang, and Subramanian S. Iyer, "RF Characterization on Nb-based Superconducting Silicon Interconnect Fabric for future large-scale quantum applications," 72th IEEE Electronic Components and Technology Conference

(ECTC), San Diego, USA, May 2022, pp. 949-955.

Yu-Tao Yang, Chaowei Hu, Peng Zhang, Haoxiang Ren, Ni Ni, Kang L. Wang, and Subramanian S. Iyer, "Fine-Pitch (≤ 10 μm) Nb-based Superconducting Silicon Interconnect Fabric for Large-Scale Quantum System Application," *2021 UCLA CHIPS Workshop,* Los Angeles, USA, Dec. 2021.

Yu-Tao Yang, Chaowei Hu, Peng Zhang, Niloofar Shakoorzadeh, Ni Ni, Kang L. Wang, and Subramanian S. Iyer, "Nb-based Superconducting Silicon Interconnect Fabric for cryogenic computing applications," *71th IEEE Electronic Components and Technology Conference (ECTC)*, pp. 1577-1582, virtual conference, USA, May 2021.

Yu-Tao Yang, Chaowei Hu, Jazmine Green, Peng Zhang, Niloofar Shakoorzadeh, Pranav Ambhore, Umesha Mogera, Ni Ni, Kang L. Wang, and Subramanian S. Iyer, "Demonstration of Superconducting Interconnects on the Silicon Interconnect Fabric Using Thermocompression Bonding," *in Proc. 70th IEEE Electronic Components and Technology Conference (ECTC)*, Orlando, USA, May 2020.

Yu-Tao Yang, Chaowei Hu, Jazmine Green, Niloofar Shakoorzadeh, Pranav Ambhore, Peng Zhang, Umesha Mogera, Ni Ni, Kang L. Wang, and Subramanian S. Iyer, "Superconducting Interconnect Fabric (Superconducting-IF) for Low-temperature Electronics," *2019 UCLA CHIPS Workshop,* Los Angeles, USA, Nov. 2019.

Haoxiang Ren, **Yu-Tao Yang**, Subramanian S. Iyer, "Recess Effect Study and Process Optimization of Sub-10 μm Pitch Die-to-wafer Hybrid Bonding" 72th IEEE Electronic Components and Technology Conference (ECTC), San Diego, USA, May 2022, pp. 149-156.

Yousef Safari, **Yu-Tao Yang**, Subramanian S. Iyer, Toshifumi Nakatani, Neal Levine, and Boris Vaisband, "Split-Fabric: A Novel Wafer-Scale Hardware Obfuscation Methodology using Silicon Interconnect Fabric, " 72th IEEE Electronic Components and Technology Conference (ECTC), San Diego, USA, May 2022, pp. 660-667.

Niloofar Shakoorzadeh, Randall Irwin, **Yu-Tao Yang**, and Subramanian S. Iyer, "Reliability Considerations for Wafer Scale Systems" 71th IEEE Electronic Components and Technology Conference (ECTC), pp. 84-89, virtual conference, USA, May 2021.

Krutikesh Sahoo, Niloofar Shakoorzadeh, Saptadeep Pal, **Yu-Tao Yang**, Puneet Gupta, Subramanian S. Iyer, "Copper to gold thermal compression bonding in heterogenous wafer-scale systems," 71st IEEE Electronic Components and Technology Conference (ECTC), 2021, pp. 487-493, virtual conference, USA, May 2021.

Nillofar Shakoorzadeh Chase, Krutikesh Sahoo, **Yu-Tao Yang**, and Subramanian S. Iyer, "Atomic Layer Deposited Al2O3 Encapsulation for the Silicon Interconnect Fabric" in Proc. 70th IEEE Electronic Components and Technology Conference (ECTC), Orlando, USA, May 2020.

Chapter 1

Introduction

In this chapter, general introduction to quantum computing is first introduced in the section 1.1, followed by major challenges to realize a large-scale quantum computing in the section 1.2 and objective of this dissertation to resolve the challenges through heterogeneous integration and advanced packaging in the section 1.3. At the last of this chapter is the organization of this dissertation.

1.1 Quantum computing

Quantum computing is one of the emerging technologies that can fulfill some requirements of high performance computing. Compared to traditional deterministic computing through CMOS and BJT transistors, non-deterministic quantum computing uses quantum mechanical principals (superposition states, quantum coherence and quantum entanglement) to perform computations in a parallel and efficient way [1], [2] and to reach exponentially-grew capability of computation. For one transistor, it stays in the bit "1" or in the bit "0" at a time; while for a quantum bit (qubit), it can simultaneously in the bit "1" and "0" with a certain probability distribution. This feature of superposition of two states allows a quantum computer, within the qubit coherent time, to compute all the possible combination between "1" and "0" in one cycle parallelly, instead of the serial computing in classical circuits. Quantum entanglement means that in a two-qubit system, the state of two qubits correlates with each other, meaning the state of one qubit cannot be described independently without the state of the other.

The physics behind quantum computing is to manipulate quantum particles between the lowest

energy state, as the state “0”, and the one-step-higher state, as the state “1”. There are various physical systems that can constitude two-level quantum states and can potentially realize a large-scale quantum computing[3]: superconductors, photons, trapped atoms, and spin. These types of qubits are briefly introduced in the following paragraphs.

Superconducting qubits use the principle of quantum harmonic oscillator, a quantum system of one single particle residing in a parabolic potential. With a boundary set by the potential energy and the kinetic energy within a particle, a quantum harmonic oscillator with equi-spacing between each state can be described through the Schrödinger equation. This harmonic system can be physically realized using circuit designs with proper inductance (potential energy, or magnetic energy) and capacitance (kinetic energy, or electrostatic energy) in a low-temperature superconducting region. Circuit designs can also be used to couple two nearby qubits. Without using superconductors, qubits decohere easily due to the I^2 X R Joule heating. To utilize the quantum oscillator for computing, anharmonicity is needed to lock the transition of the particle within the two lowest states. The anharmonicity can be introduced through the key component in superconducting qubits: the Josephson junction, which is a tri-layer tunneling junction with a thin insulation layer separating sections of a superconductor. The quantization of the charge tunnelling through the junction leads to a cosine term to the parabola in the potential energy, which is proportional to the junction critical current. The excitation energy of one qubit between the two lowest states is determined macroscopically by the inductance and the capacitance; the frequency is designed at 5 – 10 GHz, which is high enough to minimize thermal effects at the low temperature environment available in dilution refrigerators (~10 mK; $k_BT/h \sim 0.2$ GHz).

Photonic qubits are realized through the polarization state of a photon and are intrinsically free of decoherence. One-qubit gates (polarization state rotations) can easily be done using

‘waveplates’ made of birefringent material. One of the challenges is to achieve the required interactions between photons for universal multi-qubit control. Knill, Laflamme, and Milburn [4] report a scalable quantum computing scheme using single-photon sources, detectors, and linear optical circuits. This scheme relies on quantum interference with auxiliary photons at a beam splitter and single-photon superconducting nanowire detection to induce interactions non-deterministically. Regardless of the approach used for photon sources and detectors, this method leads to major photon loss challenge (~0.1 dB/cm) and degrades the coherence time close to matter-based qubits.

Trapped atom qubits are based on isolated atomic systems, confined in free space with nanometer precision using appropriate electric fields from nearby electrodes[5]. This provides excellent coherence properties of certain energy levels within atoms. Entangling two-qubit quantum gates can be realized through a laser-induced coupling of the spins mediated by a collective mode of harmonic motion in the trap. Although the coherence time is long enough and there are several attempts for a scalable system, such as collections of Coulomb-coupled ions through photonic interactions [6], and optical lattice [7], the scaling scheme and the relevant technology are still not mature.

The spin qubits use a small semiconductor nanostructure (GaAs, Si, Ge), and impurity/impurity complex (Nitrogen vacancy) to binds one or more electrons or holes into a localized potential with discrete energy levels, similar to an electron bound to an atomic nucleus. Quantum logic would be accomplished by changing gate voltages to physically move electrons closer and farther from each other, activating and deactivating the exchange interaction. The coherence time of spin qubits is demonstrated up to seconds, but the scaling of coupled spin is still challenging.

Although the quantum coherent time is still a challenge and beyond the scope of this thesis, the key advantage of superconducting qubit is its scalability by microwave circuit designing and Si fabrication as well as advanced packaging technologies. Therefore, in this dessertation, superconducting qubits are the main focus for applications and concepts establishment.

1.2 Challenges in quantum computing

Although parallel quantum computation can attain exponential growth in the computing speed, before it comes into real applications, four main challenges that need to be addressed are elaborated in the following paragraphs.

The first challenge is I/O management. Between transistors and qubits, in addition to fundamentally-different working principals, there are two differences during circuit realization that affect I/O management: one is fan-out/fan-in routing; the other one is cloning mechanism [8]. Regarding the circuit fan-out routing, transistor circuits enable output signals of the logic gate in the previous stage to feed into multiple gates in the next stage (and vice-versa for fan-in routing). The degree of fanning out is specified in Rent's rule [9]: a relationship between the external I/O count and the number of internal logic gates or transistors. As an example, in the processor of Iphone X, there are 4.3 billion transistors made with the 10 nm FinFET technology but only a few hundreds of I/O are required to exchange information with the outside world. Quantum circuits, on the other hand, do not share the same characteristic as classical circuits. Quantum information within one qubit does not directly control the state of another qubit. Even for two-qubit entangled logic gates, the classical readout and control sub-systems are required to mediate the interactions as well as to prepare and readout the state of the qubits. As for information cloning, it is doable in classical memory circuits through charge sharing between two capacitors; while for qubits, there is no fundamental cloning theorem in quantum mechanics [10]. This indicates no exact copy of each quantum state can be achieved, ensuring the fan-out/fan-in of information should happen effectively within classical interface outside qubits and implying that the number of I/O should be as large as the count of qubits for distinct I/O signals.

The I/O count and the interconnect wiring both outspread significantly as the count of qubits increases since the required I/O count is at least twice of the qubit number. For a large-scale of qubit array (> one millions of qubits), at least two millions of I/O are required. With a huge demand of I/O, advance packaging technologies are needed to come into the picture for a large-scale of qubit array.

The second challenge is the system footprint. Quantum computers nowadays occupy at least an area of ~ 10 m^2. This area includes mounting of racks of room-temperature microwave control/readout electronics and interfaces with milli-Kelvin qubits at the bottom of a dilution fridge. This distributed system poses two major problems in signaling: one is phase shift; the other is time-of-flight latency. The wavelength of the qubit controlling microwave signals (bwteen 5 – 10 GHz) is in the order of a few to 10s centimeter. This is comparable to the system size, causing currents and voltages to be the function of space and time and leading to additional constraints in impedance and phase matching network. For a large-scale qubit array following the conventional scheme, which will be detailed in the subsection of Chapter 2, millions of transmission lines that are comparable to or longer than the wavelength will be used for signal transmission, causing appreciable phase shifts in signals between different lines. It is challenging to synchronize and equalize broadband signals having different levels of phase shift for precise phase and amplitude of a microwave tone. For the time-of-flight latency between room-temperature electronics and qubits at the deep cryogenic environment, the bottleneck is limited by the readout time (in the order in 100s nanosecond to microseconds) compared to the typical values (10s nanosecond) in control time. This problem will be amplified when manipulating multiple qubits in future large-scale array. An instinct way to reduce the latency is to design an ultra-compact cryogenic system with proximal control/readout electronics at the expense of

increased thermal load. To reduce latency and to remedy the increased thermal load, a concept to integrate ultra-low-power superconducting electronics as the qubit control/readout on a wafer-level system is proposed and is elaborated in Chapter 4.

The third challenge is related to the heat and power dissipation as the qubit count scales up. In the conventional scheme, the heat and power dissipation increase linearly with the qubit number because of the previously-mentioned independent I/O issue. As an example, to manipulate 53 qubits, it is required to install 53 cryogenic isolators, nine quantum-limited amplifiers, nine high electron mobility transistor amplifiers, and about 200 coaxial cables inside a fridge [11] [12]. In a Bluefors XLD400 dilution refrigerator, the 53 qubits and their periphery electronics already consume one third of its overall cooling power around 20 μW [13]. Through installing proper shields to reduce active components, adopting superconducting cables to minimize heat dissipation, and tolerating a chamber temperature at 30 mK (20 mK is the norm for superconducting qubits) without raising much noise, there is a chance to achieve one thousand qubits by this brute-force method using one cooling unit. With multiple cooling units, it is possible to achieve a large-scale qubit array but the footprint issue gets compounded following the conventional scheme. Therefore, to control/readout multiple qubits with one transmission line cable from room temperature electronics to the deep cryogenic environment, a signal multiplexing technique is a necessity. In Chapter 4, an integration architecture with ultra-low-power superconducting electronics and superconducting multiplexing is presented.

The last challenge is to preserve delicate quantum information with thermalized environment and low insertion loss and crosstalk for a large-scale qubit array. The excitation energy in a qubit from the state “0” to the state “1” corresponds to a frequency of ~5 GHz, eqivalant to a temperature of 240 mK. To reduce any background excitation by thermal energy, the chamber

environment is cooled down to 20 mK, much lower than the excitation energy. To power on and simultaneously to avoid the increase of the chamber temperature caused by the wire induced of Joule heating ($P = I^2 \times R$), superconducting interconnects are used to decrease the on-chip DC resistance value by at least three orders of magnitude as compared to using Cu with the same dimension. For qubit control/readout within microwave superconducting transmission lines, although the conductor loss (AC resistance by the skin effect $\propto \sqrt{f}$) still exists even using superconductors, the value is at least 30X smaller (inverse propotional to the square root of conductivity) than using Cu as the interconnect material. As compared to another loss source by dielectric loss (dielectric conductance and loss tangent loss $\propto f$), the conductor loss is exceeded by the dielectric loss at a few hundreds of MHz to few GHz and becomes the dominating loss factor in insertion loss. The threshold for insertion loss (conductor loss + dielectric loss) on superconducting transmission lines between control/readout qubit should be less than 1 dB loss [14]. Crosstalk is another factor that degrades quantum information through electromagnetic energy sharing, which occurs between long coaxial cables and between on-chip transmission lines. Each coaxial cable spanning from room-temperature electronics to the deep cryogenic environment (few meter long) has capacitance (~ 100 pF/m for most material [15] with 50 Ω characteristic impedence) and inductance (~250 nH/m) available for coupling interface. To negate effective singal loops available for crosstalk pickup, this requires additional shielding on cables, which increases space and thermal load constrains inside a fridge chamber. For on-chip transmission lines, the threshold for crosstalk is that the value should not surpass the value of thermal noise (-23 dB with an on-chip pulse of 300 μV [16]) and not become the major noise source. There are on-chip methods to reduce crosstalk, such as shortened link length, via-fencing and increasing inter-signa-path distance, when the crosstalk is high. In Chapter 6, an low-loss

and low-crosstalk RF behavior is optimized and presented through the approach of shortened link length on the advanced packaging platform, Superconducting-IF.

1.3 Objective of this Dissertation

As mentioned previously, I/O management (no quantum cloning theory and no circuit fan-out), phase shift and latency (large system footprint), the heat and power dissipation per qubit (a distributed system), and transmission of delicate quantum information (lossy and high crosstalk environment) are the four major challenges to realize a large-scale superconducting quantum computing. To minimize these challenges, a heterogeneously integrated superconducting control/readout system on the same substrate with qubit chips is one of the promising solutions. In view of this, a heterogeneous integration Si platform called Superconducting-IF, adapted from the Silicon Interconnect Fabric (Si-IF), is investigated in this dissertation for furture quantum computing applications.

1.4 Organization of this Dissertation

This thesis is organized as follows: Chapter 2 introduces the state of the art of the superconducting quantum computing. Chapter 3 describes the packaging technology, especially in the superconducting field, and the Si-IF technology. Chapter 4 details the integration concept using Superconducting IF with superconducting electronics and comparison with existing state of the art quantum computer. Chapter 5 shows the realization of the Superconducting IF and how it is adapted from Si-IF. Chapter 6 presents the RF characterization of the superconducting interconnects on the Superconducting IF to build a low-loss and low-crosstalk integration. Finally, the conclusion and future outlook of this work are given in Chapter 7.

Chapter 2

State of the art in the architectures of superconducting quantum computing

In this chapter, the room-temperature-CMOS architecture is first introduced in the section 2.1, followed by the CryoCMOS architecture in the section 2.2 and the 3D-stacking architecture in the section 2.3. As another potential technology for quantum computing, Josephson-junction-based architecture featuring ultra-low-power-dissipation is briefed in the section 2.4. At the last of this chapter is the chapter summary.

2.1 Conventional room-temperature CMOS architecture

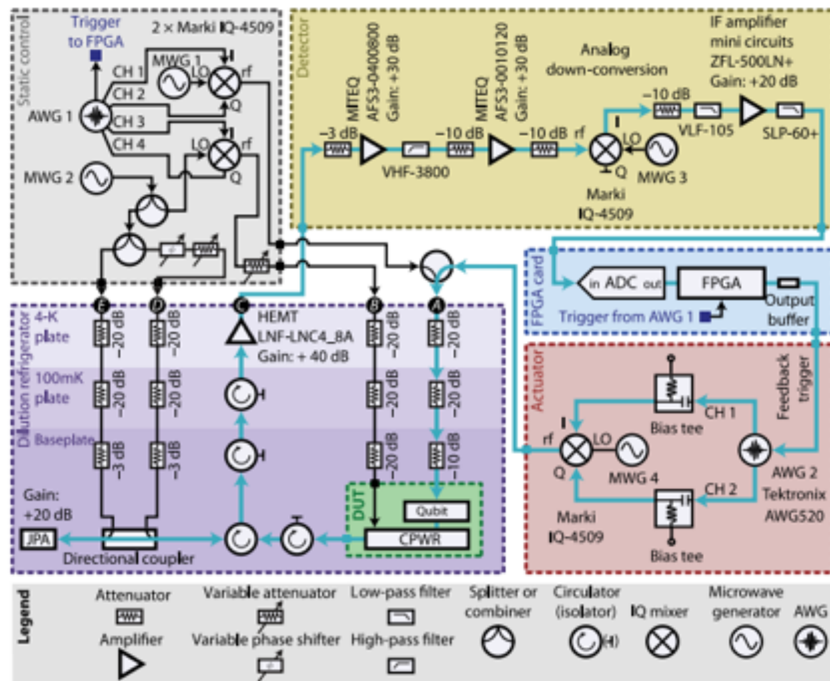


Figure 1: Existing control/readout scheme in a quantum computer. Reprinted with permission © 2018 American Physical Society [17]

The room-temperature-CMOS architecture is the first approach to realize quantum computing with distributed systems. For a commonly used rack-mounted quantum computer, there are four subsystems [17]: (1) a room-temperature FPGA instruction card, (2) room-temperature qubit control/readout electronics, (3) qubit chips with readout resonator at millikelvin stage, and (4) room-temperature qubit signal detector. Among all the subsections, only qubit chips sit inside a dilution fridge, other three subsections are outside a dilution fridge, as shown in Figure 1.

To manipulate the qubit states, a qubit needs to be in an operation region, i.e., excitation from the ground state "0" to its first excited state "1". This excitation is achieved through dedicated microwave signals. Each microwave coaxial cable that carries the signals is thermalized by three attenuators (-20 dB) that are installed at each cryostat (4K plate, 100mK plate, and base plate). The attenuators reduce both signals and noise from room-temperature electronics and can decrease the effective temperature of the microwave radiation from co-axial cables. Following the excitation pulses, the qubit control pulses are generated, triggered by a waveform generator and processed by a digital co-processor (FPGA in this case), followed by up-converted frequency through a I-Q mixer driven by a local oscillator signal to microwave frequencies.

To readout a qubit, a pulsed measurement of the microwave transmission through a coplanar waveguide resonator is used. The readout pulses are generated through a waveform generator and are shaped through I-Q mixing with a local oscillator signal for faster ring up and ringdown of the intracavity field [18], [19]. These signals are then transmitted into the coplanar waveguide resonator near the qubit with another three -20 dB attenuators. In the pulsed measurement, another pump signal is sent to the directional coupler at the bottom of the fridge with another three attenuators, an isolator, and a circulator. The isolator and circulator prevent the qubit chip from thermal noise and other types of noises. Combined with the the pump signal, the signal

from the resonator is sent back to room temperature electronics and is amplified by a 40-dB high-electron-mobility transistor (HEMT) amplifier before leaving the chamber of the dilution refrigerator.

To detect and analyze the signal exiting from the dilution refrigerator, the signal is first amplified by low-noise microwave amplifiers, then is de-noised by high pass filter, and is down-converted to an intermediate frequency. The process signal is then digitized by an ADC and forwarded to Digital Signal Processor inside the FPGA for feedback trigger. This completes one cycle of qubit control/readout and the time of each cycle can go as high as millisecond. The reported fidelity is around 80% (error rate 20%).

The aforementioned control/readout/detection procedure is designed for operation of a few qubits. When the qubit count increases to 50-100, each qubit on average requires at least one isolator, one-third amplifier, and two to three coaxial cables. All the hardware components and cables increase linearly with the number of qubits, meaning one million of isolators, three hundred thousand of amplifiers, and at least two millions of coaxial cables are require for the realization of one millions of qubits. This huge amount of hardware causes not only the footprint problem, but also the latency, phase matching, and scalability issues mentioned in the Chapter 1. Therefore it is challenging for large-scale quantum computing when the room-temperature-CMOS architecture is adopted.

2.2 Cryogenic CMOS (CryoCMOS) architecture

In view of the I/O, footprint, latency, and scalability, there is material in integrating either digital processors, readout circuits, or both inside a fridge and installing functional CMOS chips as close to qubit chip as possible to reduce any environment noises. The key advantage for this architecture is that it applies the most well-developed and manufacturable CMOS technology to realize qubit control/readout circuits; while the disadvantage is obvious: CMOS circuits generate significant amount of heat, requiring specific design to directly interface with millikelvin-required qubits. In the following paragraphs, Google [20]–[22], Microsoft [23], and Intel [24] examples regarding the CryoCMOS architecture are presented.

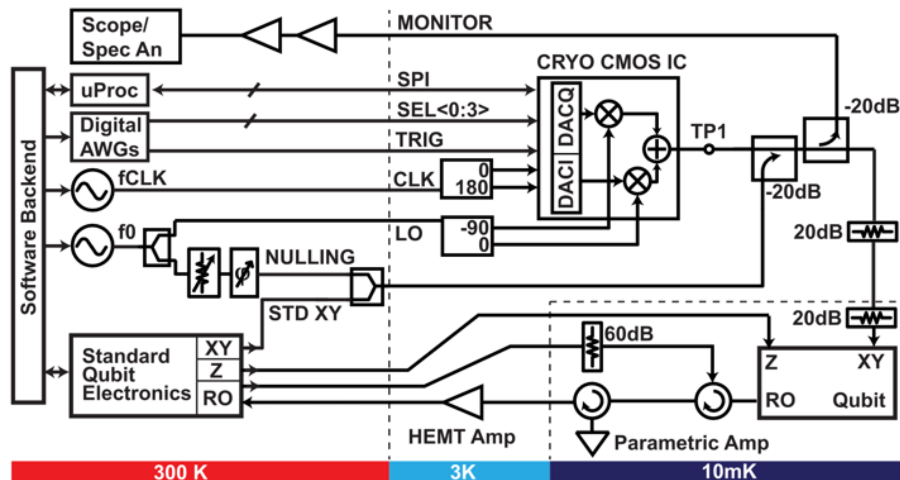


Figure 2: Google CryoCMOS architecture. Reprinted with permission © 2019 IEEE [20]–[22]

The CryoCMOS architecture of Google [20]–[22] is presented in Figure 2. The key improvement is that the CryoCMOS qubit controller is installed at 3K. The control/readout cycle is achieved by a room-temperature trigger, CryoCMOS qubit control (pulse shaping) at the 3K cryostat, and room-temperature readout. The CryoCMOS qubit controller is made with the 28-nm bulk CMOS technology having the function blocks for signal processing, including the in-

phase and quadrature-phase digital analog converter, digital controller for pulse generator, pulse controller, and mixer. The shaped pulses are attenuated by 40 dB (two 20 dB attenuators) to have an amplitude of 10-100 μV before entering the superconducting qubit chip. The qubit output signals are sent back to room-temperature detectors and analyzers for readout analysis. The clock frequency is maintained between 0.5 - 3 GHz and the overall DC+AC power is less than 2 mW at 3K. The instruction set is 2^4 codewords. As compared to conventional room-temperature rack-mounting controls, Google's CryoCMOS example provides similar energy relaxation time ~ 18 μs and specifications corresponding to an overall high-fidelity gate of 99.99% (error rate of 0.01%).

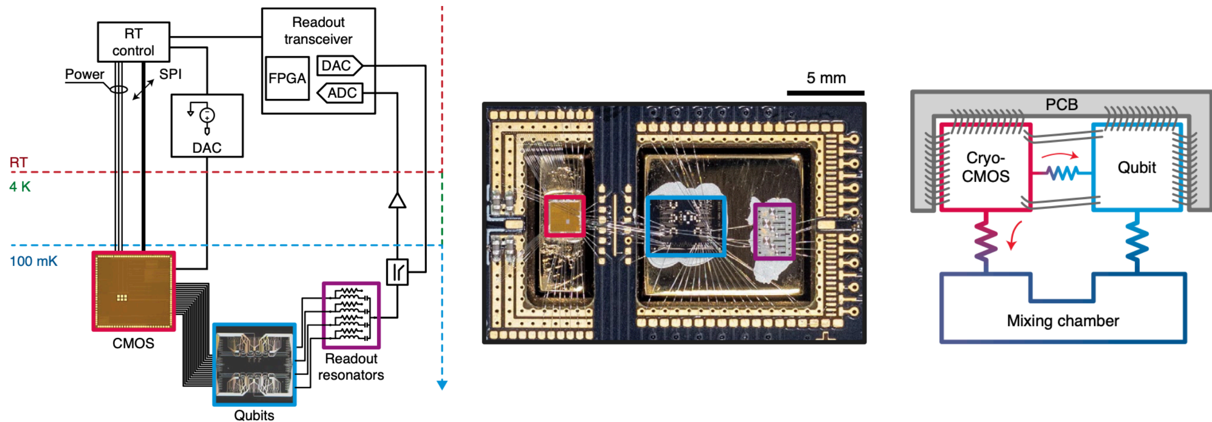


Figure 3 Microsoft CryoCMOS architecture. Reprinted with permission © 2021 Nature [23]

The CryoCMOS architecture of Microsoft [23] is presented in Figure 3. The key improvement is that a CryoCMOS control chip wire-bonds to a quantum dot qubit chip at 30 mK on a PCB. The control/readout cycle is achieved by a room-temperature FPGA trigger, 100-mK CryoCMOS qubit control (pulse shaping), and room-temperature readout. The CryoCMOS control chip is made using 28 nm FDSOI technology (control voltage ~ 0.8 V) with digital and analog function blocks, such as serial peripheral interface, waveform memory, autonomous

operation of the chip, ring oscillator, and D.C. charge locking; the corresponding qubit readout signal is between 10 to 50 μV. To compensate the high joule heating disadvantage of CryoCMOS chips, both the static and dynamic power dissipation are intentionally reduced with proper technologies and designs. For the static power dissipation (P = I x V), due to the structural nature of FDSOI, the leakage current can be decreased to ~3.2 nA for the whole chip, corresponding to a power of 3.2 nW with a voltage of 1 V. For the dynamic power dissipation (P = C x V^2 x f), the capacitance is reduced through reduction of the transistor count to 100,000 and the frequency is decreased to at most 10 MHz. The overall dissipated power is around 20 μW at 10 MHz, causing the chamber temperature to increase to around 100 mK. On average, each low-power CMOS transistor consumes ~0.2 nW at 20 mK. Since this Microsoft example emphasizes on the architecture and the power dissipation of CryoCMOS, there is no reported qubit gate fidelity in this work.

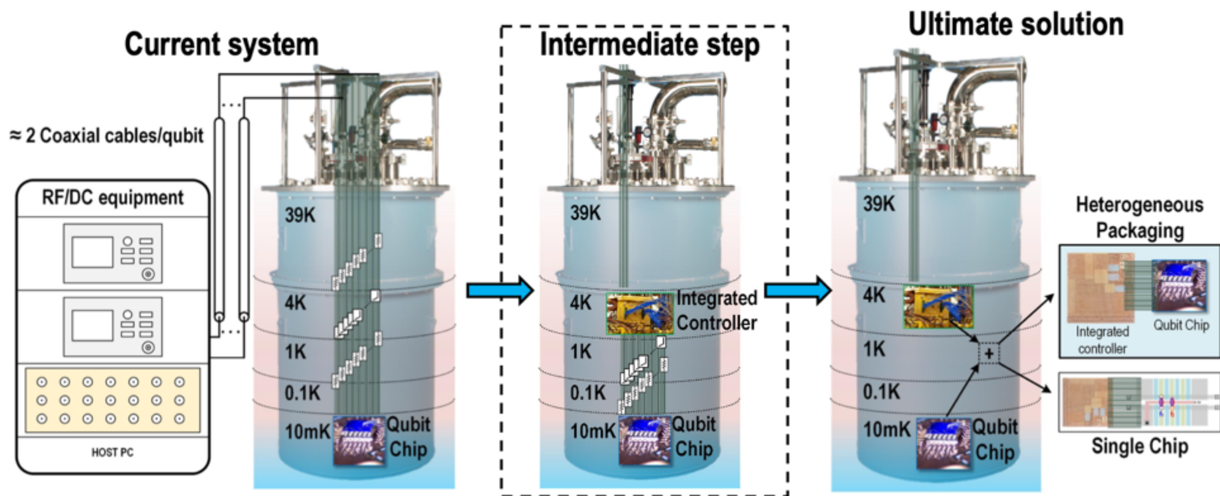


Figure 4 Intel CryoCMOS architecture. Reprinted with permission © 2021 IEEE [24]

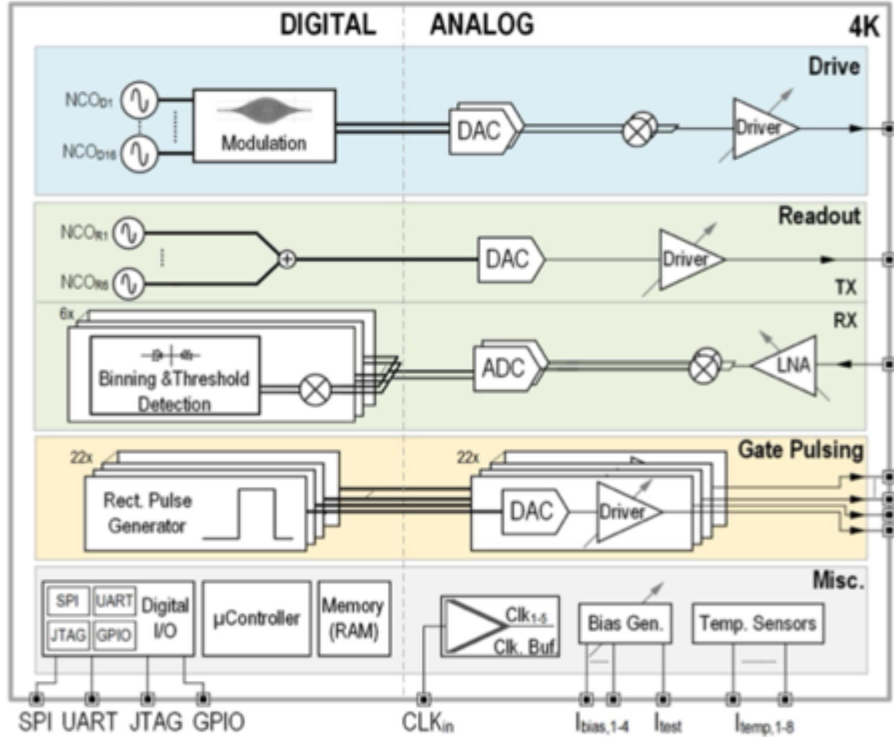


Figure 5 Intel CryoCMOS architecture (cont.). Reprinted with permission © 2021 IEEE [24]

The CryoCMOS architecture of Intel [24] is presented in Figure 4 and Figure 5. The key improvement is that the CryoCMOS qubit control and readout are integrated as a SoC at 4K. The control/readout cycle is achieved by room-temperature digital triggers, and 4K CryoCMOS qubit control and readout. The CryoCMOS microcontroller is made with the 22 nm FinFET technology having the SoC qubit control/readout (digital and analog function blocks, such as qubit drive, readout, gate pulsing, memory, microcontroller) and connects to spin qubit chip at a based temperature of 10 mK. The SoC chip operating at 1.6 GHz clock frequency can multiplexing control the 16 spin qubits and can send instruction set of 2^{19} codewords per qubit. This instruction set is 2^{15} more than the non SoC approach. The pulse amplitude is 300 µV. The consumed power including analog and digital is around 439.2 mW. This Intel example provides the specification for high-fidelity gate of 99.99% (error rate 10^{-4}). The comparison table between

the conventional scheme and the work of Google, Microsoft, and Intel is shown in Table 1.

Table 1 Comparison between the conventional scheme and the work of Google, Microsoft, and Intel

	Conventional	**Google [20-22] (2020)**	**Microsoft [23] (2021)**	**Intel [24] (2021)**
Technology	CMOS @ RT	CryoCMOS @ 4K	CryoCMOS @ mK	CryoCMOS @ 4K
Qubit control source	Analog @ 300K	Analog @ 4K	Mixed signal @ 100 mK	Mixed signal @ 4K
Qubit readout	Analog to RT	Analog to RT	Analog to RT	Analog to 4K
Qubit type	Superconducting	Superconducting	Topological (goal)	Spin
clock	0.3-2 GHz	0.5 GHz	Few MHz	1.6 GHz
Power	20mW	2.5mW	20 uW (10 MHz & 100k CMOS)	439.2 mW
Latency (control/ readout cycle)	Milliseconds	Milliseconds	Milliseconds	Microsecond (potential)

2.3 3D stacking architecture

MIT Lincoln Labs has leveraged their 3D superconducting TiN Through Silicon Via (TSV) stacking technology to isolate high-coherence superconducting qubits from surrounding noises and to alleviate interconnect crowding for a large-scale qubit array, as shown in Figure 6. The three-stack structure includes the top qubit chip, the middle interposer, and the bottom superconducting multi-chip module (SMCM) chip. The top chip contains high-coherence superconducting qubits that are fabricated and optimized on a separate silicon wafer. The middle interposer chip contains TSVs that line with superconducting TiN and connect the top qubit and the bottom SMCM. The three chips are assembled to form electrical connection through indium bumps, providing a low-resistance electrical path. The TSV-integrated interposer allows the multilayer signal routing in the bottom SMCM while retaining isolation of qubits from lossy dielectrics. The bottom SMCM has multiple planarized superconducting metal layers [25] that are connected using vias and uses processes of digital superconducting electronics such as single-flux quantum circuits or traveling-wave parametric amplifiers [21]. No significant impact on the qubit performance is reported in the three-chip-stacking method compared to the monolithic and the 2D flip-chip planar integration.

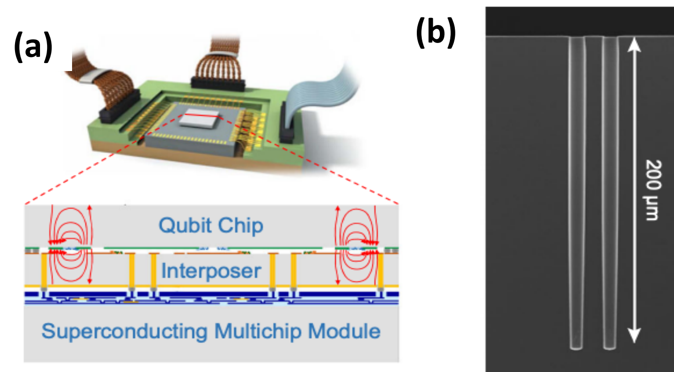


Figure 6 3D packaging on a qubit chip. Reprinted with permission [27]

2.4 Josephson-Junction-based (JJ-based) architecture

Besides the aforementioned CryoCMOS architectures, qubit control/readout through Josephson-Junction-based superconducting digital circuits is yet another option. The key attraction is its ultra-low heat dissipation, making superconducting qubit control/readout a promising way to be integrated proximally with qubits through flip-chip bonding technologies on the same substrate. The down side is its capability of vary large scale integration.

There are a family of JJ-based circuits: Adiabatic Quantum Flux Parametron (AQFP)[28] energy-efficient Single Flux Quantum (eSFQ)[29], LR-biased RSFQ logic[30], reciprocal quantum logic (RQL)[31], Energy-efficient Rapid Single Flux Quantum (ERSFQ) logic[32], and Rapid Single Flux Quantum (RSFQ)[33]. SFQ is a superconducting digital circuits that use Josephson Junctions to process quantized and digitized signals at 4K. Information is stored as the magnetic flux quanta in the form of SFQ voltage pulses. The amplitude for a magnetic quanta (Φ_0) is 2 mV x ps. In real applications, the voltage amplitude is 200 to 400 μV with a time domain spread of 5 to 10 ps. AQFP logic is another type of JJ-based circuits that uses the phases of AC current to determine logic “1” (positive current) or logic “0” (negative current) with a reset state (zero current). This can further reduce the power dissipation per JJ due to zero static power dissipation. Ayala *et al.* [34] in 2021 has used the AQFP technology to demonstrate ultra-low power superconducting digital processing and memory operation with a clock frequency of 5 GHz. Considering the dissipated power per JJ of 0.008 nW and a Josephson Junction count of 21460, the overall power dissipation is around 170 nW in the demonstrated chip.

In the superconducting circuits, the general power dissipation is categorized as the static dissipation ($P \propto I \times V$) and the dynamic dissipation ($P \propto L \times I^2 \times f$, or $P \propto \Phi_0 \times I \times f$). The

detailed power dissipation for each technology is listed in Table 2 (from low to high power dissipation). Since some of the reference papers only provide the overall energy dissipated per Josephson Junction without the switching speed, an averaged device on/off switching speed of 5 ps [35] is used during calculation to convert the energy dissipation per JJ to the power dissipation per JJ.

Table 2 Estimated power consumption of JJ-based family operating at 4 K with a clock frequency of around 5GHz

Device at 4K	AQFP [28]	eSFQ [29]	LR-biased RSFQ [30]	RQL [31]	ERSFQ [32]	RSFQ [33]
Power Dissipation (nW/JJ)	0.008	1.000	1.360	2.000	1.500	8.200

To integrate with qubits on the same substrate, the operating temperature of JJ-based family should be reduced to 20 mK. Based on Initiso S *et al.* [36], with a lower critical current process and a low voltage, the power dissipation of RSFQ can be reduced to 25 pW per Josephson Junction at 20 mK, which is 392 times smaller than the dissipation of RSFQ at 4K. Assuming the same scaling ratio for each type of the JJ-based device, the power dissipation at 20 mK is presented in Table 3.

Table 3. Estimated power consumption of JJ-based family operating at 20 mK with a clock frequency of around 5GHz

Device at 20 mK	AQFP	eSFQ	LR-biased RSFQ	RQL	ERSFQ	RSFQ
Power Dissipation (pW/JJ)	0.02	2.55	3.47	5.10	3.83	20.92

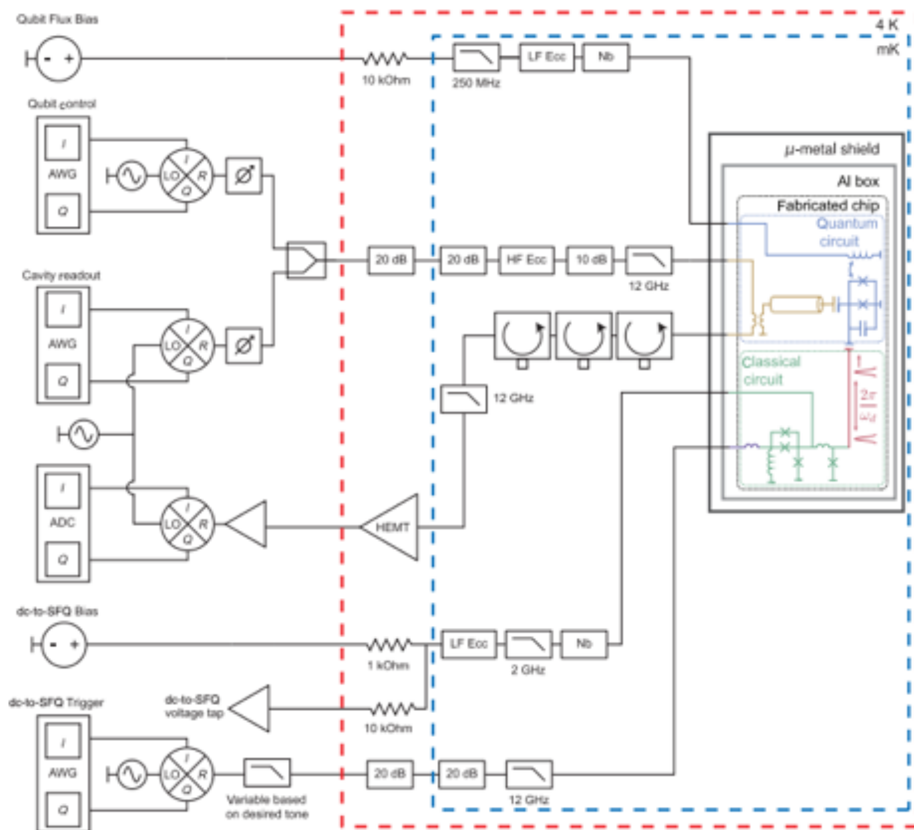


Figure 7. Architecture for qubit controlling with JJ-based devices. Reprinted with permission © 2019 American Physical Society [37]

Syracuse university in 2018 has demonstrated using previous mentioned JJ-based devices to control [38], [39] and readout [40] qubits on a single integrated Si chip without degrading the performance of qubits. Both demonstration of qubit control and readout are briefed in the following paragraphs. The architecture is shown in Figure 7 with the integrated chip of qubit control and qubit installed in the mK chamber. The detailed SEM image of the circuit is shown in Figure 8.

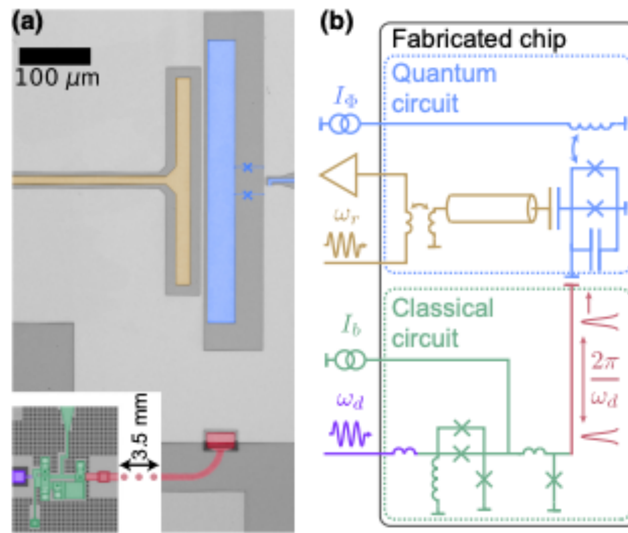


Figure 8 SEM image of integrated control with a qubit. Reprinted with permission © 2019 American Physical Society [37]

For the qubit control, the qubit is manipulated by quantized delta function-like pulses produced by JJ-based digital logics. Each pulse introduces a discrete rotation of the qubit state vector in the Bloch sphere, as shown in Figure 9. With a proper timing design, the pulses can be delivered to a qubit with a timing interval matching to the qubit resonant frequency, ensuring a coherent rotation and a coherent control. This approach is analogous to a swing receiving a sharp push in each cycle. Reported by E. Leonard *et al.*[37], the first trial of one qubit controlled by JJ-based circuits achieves gate fidelities of 95%. This fidelity can be increased through reduction in

the critical currents of the JJ-based digital logics and through multi-chip module packaging, reducing non-superconducting particles on the same chip flowing into qubits. Li *et al.*[41] reports with more complex pulse sequences involving irregular pulse interval, the gate fidelity in excess of 99.99% (error rate < 10^{-4}) can be achieved.

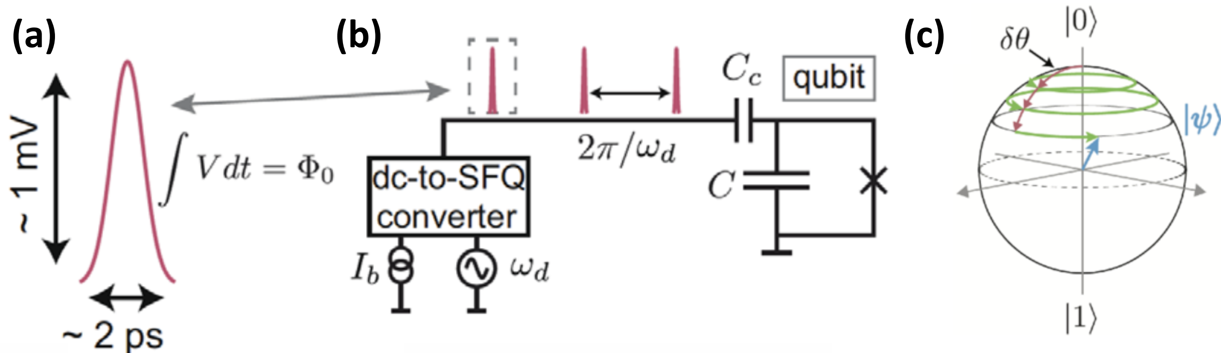


Figure 9 Qubit manipulation (a) JJ-based SFQ pulse (b) coherent control of qubits with time-domain pulse interval matching the qubit resonant frequency (c) discrete rotation of the qubit state vector in the Bloch sphere. Reprinted with permission © 2019 American Physical Society [37]

For the qubit readout, the use of propagating quantized flux as a direct probe of the state of a qubit is first investigated theoretically by Averin *et al* [42]. With the help of full circuit simulation [43], an experimental implementation with Josephson Photomultiplier (JPM) has been realized [40], [44] with the first trial fidelity of 92%. This is the first demonstration of high-fidelity readout of a superconducting qubit without microwave isolators protecting the qubit from amplifier noise. Reported by Howington *et al.*[31]&[37], to mitigate non-superconducting particles poisoning due to the SFQ circuitry, and to avoid the complexities of fabricated SFQ circuits and high-coherence qubits on the same chip, a multi-chip packaging module (MCM)

approach has been adopted to realize the high-fidelity readout of a superconducting qubit. Although no real number is reported in the literatures [31]&[37] describing how many precents of gate fidelity can be improved after application of MCM packaging, this passes an important message that heterogeneous integration with the dielet design by advanced packaging can play a significant role in quantum computing.

2.5 Summary

In this chapter, the room-temperature-CMOS architecture is first introduced in the section 2.1. The conventional distributed system architecture and the hardware setup are detailed. The conventional approach is challenging to realize a large-scale quantum computing. The CryoCMOS architectures in the section 2.2 include examples from Google, Microsoft, and Intel. A comparison table is organized between the conventional architecture and the CryoCMOS approaches. The 3D-stacking architecture through superconducting TSV from MIT is detailed in the section 2.3. As another promising technology for quantum computing, Josephson-junction-based architecture featuring ultra-low-power-dissipation is briefed in the section 2.4. Various types of JJ-based circuits and their power dissipation are presented. The literature of one qubit control and one qubit readout by superconducting circuits are introduced with the application of advanced packaging technologies.

Chapter 3

Overview of packaging technologies

Semiconductor packaging are leveraged to allow silicon dies with fine micrometer-scale wires to communicate with the outside centimeter-scale world and simultaneously to provide assemblies with mechanical supports and protection from outside temperature and humidity stimulus. In this chapter, packaging technologies related to semiconductor industries are introduced in the section 3.1. Next, the packaging technologies used in the quantum field are detailed in the section 3.2. With the past progress in packaging technologies, the advanced wafer-level integration – Silicon Interconnect Fabric (Si-IF) – is proposed and developed at UCLA CHIPS and is detailed in the section 3.3.

3.1 Packaging technology evolution

The development of packaging I/O interconnects is driven by the previously mentioned Rent's rule [9]: the more gates and functionalities in a die, the higher I/O count a die should have. In the past four decades, as compared to the 1000 X scaling in semiconductor device technologies, the packaging dimensions have only scaled by 4-5 X [46]. Recently, due to the demand of high-performance computing, innovations in the packaging field are blooming. In the subsection 3.1.1, a conventional integration approach is introduced and then in subsection 3.1.2, an advanced integration is described.

3.1.1 Conventional System Integration

Printed Circuit Board (PCB), as a mainstream packaging platform, plays an indispensable role

in the packaging industry and is adopted very widely and successfully over the past several decades. For a PCB assembly, Controlled Collapsed Chip Connection (C4) bumps at 100-150 μm pitch, flux, underfill, and molding compound are used first to secure bare dies on Redistribution-Layer-patterned (RDL) laminates as die carriers. The carriers are then assembled using wire-bonding and large-pitch (0.4 - 1 mm pitch) Ball Grid Arrays (BGA) [47] on a PCB having an inter-chip communication in the millimeter to centimeter scale.

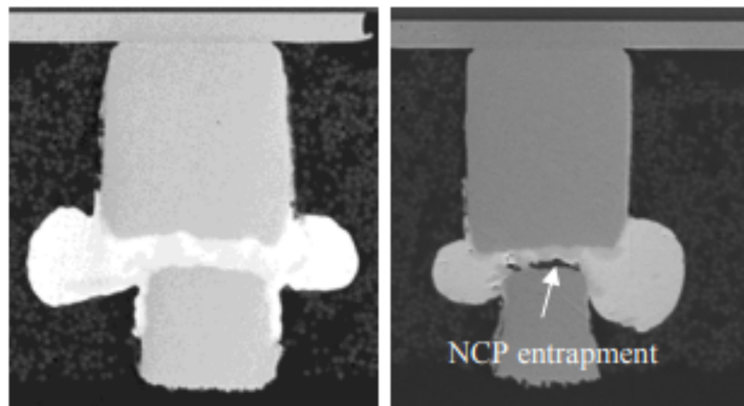


Figure 10 solder extrusion. Reprinted with permission © 2015 IEEE [48]

This type of assembly causes three major problems: (1) the millimeter-to-centimeter-scale communication causes impedance mismatch, signal latency and data bandwidth issues, (2) the high coefficient of thermal expansion (CTE) in polymer/organic material generates thermal stress during the solder reflow heating process (240 – 260 ℃) and leads to the PCB substrate warpage issue. Although the warpage of PCB can be minimized within 0.5% [49] through balanced distribution of dielectrics and Cu on the front and back sides, the warpage is not low enough for fine-pitch solders and limits the scaling of these solder-based interconnects, (3) solder extrusion, as shown in Figure 10, limits the fine-pitch scaling of flip-chip solders and is an uncontrollable reduction in the distance between bumps, requiring a minimum stand-off distance.

Although A. Manasson *et al.* in the reference [50] reveal a fabrication approach through evaporation to achieve very fine pitch (<10 μm) of In bumps, it only shows images after deposition and does not apply flip-chip bonding to demonstrate the fine-pitch feasibility of Indium bumps. Flip-chip bonding is the key process to determine whether fine-pitch joints can be attained without electrically shorting of nearby bumps. Solder materials can also cause reliability problems, such as intermetallic compound (IMC) formation, which is brittle and is similar to a ceramic material, and Kirkendall voids, which originate from diffusivity difference between bumps and under bump metallurgy (UBM) and can lead to electrical opening. Besides these issues, the height uniformity of bumps, which can vary over 15% and even up to 50% on 5×5 mm^2 dies [48], [49], is yet another major roadblock as it is very difficult to control during reflow process. The bump height is directly related to the underlying UBM area, the bumping process, and the surface tension of liquidus solder during reflow. The resultant tall bumps after reflow lead to high inductance in the whole interconnect system, eventually affecting data transmission.

3.1.2 Advanced System Integration

Due to slowing down of Moore's law, high performance computing by innovations in the packaging field is getting more attraction, such as monolithic approaches (System on Chip, and Wafer Scale Integration), and heterogeneous integration methods (Multichip Module, interposers, and 3D Integration).

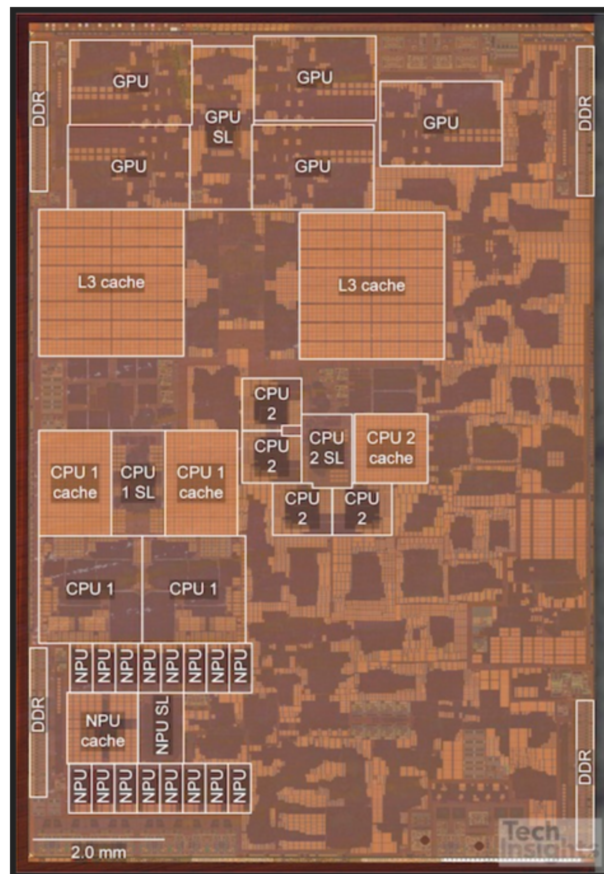


Figure 11. Apple A15 processor System on Chip (SoC) design with multiple function blocks.

(Picture source [51])

The monolithic method is realized through a high transistor density of ≥100 million transistors/mm^2 and the interconnect technology. The device technology is achieved by the advancement of Si fabrication techniques in the 5 nm node combined with Dennard's scaling

theory [52]. The interconnect technology is advanced by the increase of wiring layers and a wiring hierarchy of fine-pitch wires (<100 nm) connecting neighboring nodes and larger-pitch wires (few microns) connecting far nodes. Both evolvements merge more functional blocks in a single die. Figure 11 shows one of examples of System on Chip (SoC) using the monolithic method and several intellectual property (IP) and functional blocks necessary for a system are fabricated on one single Si die. Short inter-block spacings and fine-pitch wiring can be provided in a SoC, increasing energy efficiency and data bandwidth and reducing signal latency. Three limitations restrict the further development of SoC: one is the reticle limitation ($\sim 800\ mm^2$); another is the complex design in SoC; the other is homogeneity in technology. For the reticle restriction, even with advanced and denser transistors, the die size still increases to near reticle limits to accommodate complex functions, such as Central Processing Unit (CPU), Graphics Processing Unit (GPU), Neural Processing Unit (NPU), system cache, etc. The complex design in SoC includes IP hardening and Si validation for every tape-out, leading to a high non-recurring engineering cost and time to market. For homogeneity in technology, due to the fabrication tool limit, one SoC chip can only apply one technology node, unable to splice various optimized technology nodes in each function block. The idea of SoC is taken to the next level to Wafer Scale Integration (WSI) [53] by Gene Amdahl in 1980 to achieve low latency, high energy efficiency and high data bandwidth. The significant challenge is the wafer yield, making WSI extremely expensive. One of the immediate way is to design redundant cores and meshing network, compounding the key advantage of WSI by increasing the length of signal paths and reducing the system speed. Recently by Cerebras [54], a wafer scale chip is demonstrated using a 12-inch wafer for artificial intelligence and machine learning. Redundant 1.5% extra cores are included [55] in the design to reduce the impact of defects and it works during fabrication. The

wafer-scale chip presents a bright future to the chip industry but it is still a homogeneous system and the overall price for each wafer-scale chip exceeds two millions of US dollars.

The heterogeneous integration methods refer to assemble dielets or components from various technology nodes and materials on one common substate. Many innovations regarding heterogeneous integration have been demonstrated [56]–[58]. This method improves the latency and the bandwidth challenges on traditional PCBs [59] and decreases complex designs as well as overhead cost on SoCs by splitting a large SoC into small chiplets or dielets [60]. Several works [61]–[63] have discussed and presented the advantages of heterogeneous integration on different aspects of architectures, system performance and overall scaling. The earliest heterogeneous integration method can be traced back to 1980s as the Multi-chip-modules (MCMs), on which multiple dies are assembled laterally with finer I/O pitch (<100 μm) than PCBs [64]. With the concept of heterogeneous integration and IP reuse, design and manufacturing costs are significantly reduced up to 41% using MCMs [62]. The substrates for MCM can be ceramic (Aluminum oxide, Aluminum nitride, and Beryllium Oxide) and organic material (FR-4, and polyimides) in consideration of the coefficient of thermal expansion, the dielectric constant, and the thermal conductivity [65]. To further drive compactness and the high-interconnect-density redistribution layer (RDL) between dielets[57], [66], several materials, such as Si[67], glass [56] and organic [68], are proposed as the substrate material and are named as an interposer. The interposer adds an extra level in the packaging hierarchy with on-chip-like wiring (≤ 4 μm pitch) and the interposer assembly is packaged and assembled on PCBs. To connect to the underlying PCB, through-silicon vias (TSVs) are added within the Si interposer (200-300 μm thick [57]), minimizing the interposer warpage issue. One of the example adopting the interposer technology is the M1 Ultra chip assembling two M1 Max on one Si interposer, as shown in Figure 12.

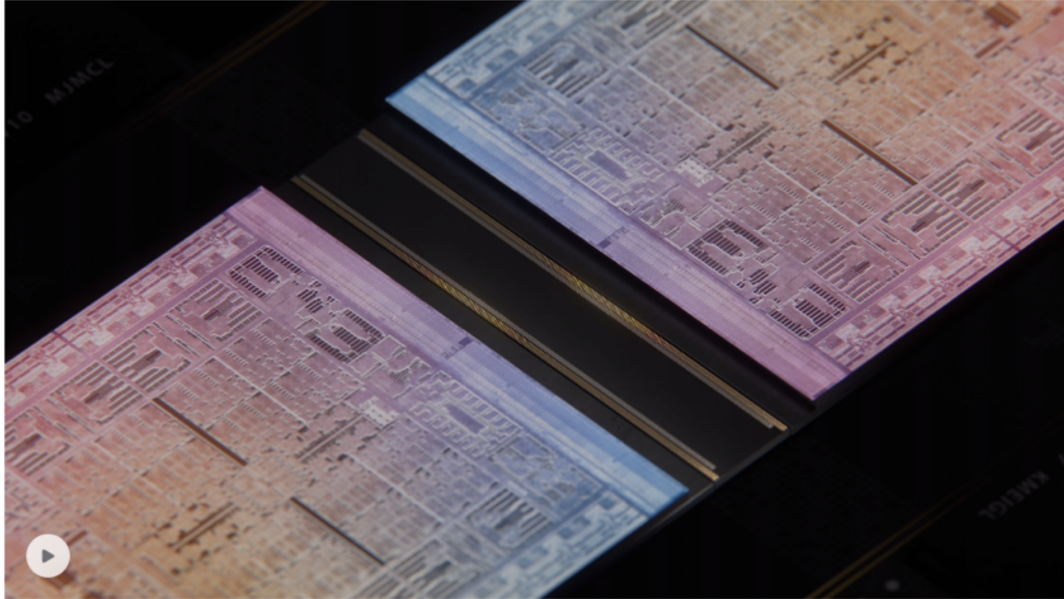


Figure 12. Apple M1 Ultra on Si interposer.(picture source [69])

The Si interposer that carries M1 Ultra chip has a area ≈ 2500 mm^2, closing to three times of the reticle size with the die stitching method [70]. Although the wiring on the interposer can be scaled down to ≤ 4 μm pitch, the die-to-substrate interconnect pitch (I/O pitch, bump pitch) is around 50 μm, which is limited by the solder extrusion and the warpage issues mentioned in the subsection 3.1.1. As reported by other works [71]–[75], high performance computing requires an interconnect pitch of ≤10 μm, which is far below 50 μm. In the subsection 3.3, an advanced packaging technology called the Silicon interconnect Fabric that can achieve sub 10 μm I/O pitch is introduced. Another heterogeneous integration method adopts vertical stacking of the dies on top of each other. This is known as 3D integration, reducing the chip area and providing vertical I/O interfaces between dies using TSVs. 3D integration can be classified as the front-end 3D (wafer-to-wafer bonding) and the back-end 3D (die-to-wafer bonding) [72]. The key advantage of wafer-to-wafer bonding is that it can provide a vacuum processing chamber and less mechanical vibration, offering a clean and flat surface for fine vertical interconnect pitch <1 μm

[76]. The downside includes homogeneity and yield (defective area on the first wafer bonding to function area on the second wafer). The key advantages of die-to-wafer bonding are heterogeneity and known-good-die for a high yield. 3D die-to-wafer bonding has been widely used for stacking up to 12 memory thinned dies, instead of logic-on-logic due to the thermal problem, the TSV I/O area, and keep out zone considerations.

3.2 Packaging in quantum application

As compared to the blooming semiconductor packaging, the packaging in the quantum field is still in the early development phase and is less discussed due to the small qubit count of 50 -100 and the sparse distribution of qubits. Because the number of qubit is so low, a single metal layer on a qubit process is sufficient and the most common and easiest way to have qubit dies communicate with the outside world is through Al wire-bonding. This provides flexibility and overcomes the chip carrier morphology and the distance between chip and carrier. The biggest challenge using wire-bonding, however, toward a future large-scale qubit array is the I/O count and density, as mentioned in the section 1.2. Wire-bonding can only approach the circuit through its periphery sides, limiting the I/O numbers. As an example, with a 5 X 5 mm^2 qubit chip, the I/O count is around 265 with a wire-bonding pitch of 75 µm. Although the I/O count can be increased through the stacking of multiple PCBs with wire-bonding on each layer, as shown in Figure 13, the I/O density is still far sparse as compared to the flip-chip MCM method.

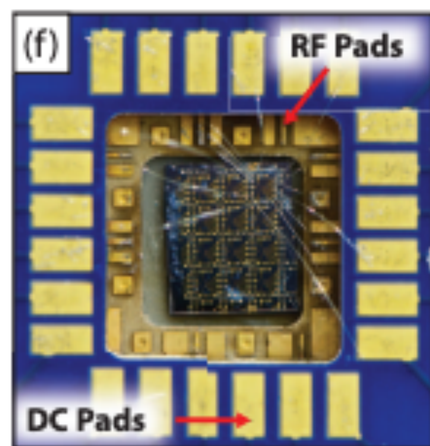


Figure 13. Wiring bonding to qubit chip. Reprinted with permission © 2012 AIP [77], [78]

Using the flip-chip approach, the I/Os can be made with bumps and are distributed across the

die area, instead only on the die sides. This significantly increases the I/O count and density and the most important thing is that the flip-chip method allows heterogeneous integration. In the subsection 3.1.2, the flip-chip interposer method, as shown in Figure 14, and the 3D integration method, as shown in Figure 6, have been introduced and are used in the integration of qubits [27] ,[37], [39] without significant degradation to the qubit characteristics.

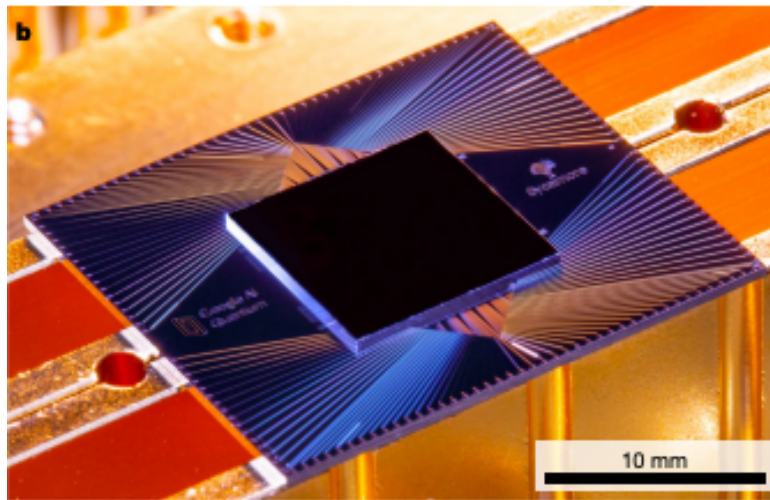


Figure 14. Flip-chip packaged qubit chip through the interposer technology. Reprinted with permission © 2019 nature [12]

To be compatible with Josephson junction in qubits, the processing temperature should be less than 150 ℃ since Josephson junctions made with a Nb/Al-oxide/Nb tri-layer inside superconducting dies are extremely sensitive to a high temperature. The electrical characterization of this tri-layer structure, such as the critical current, the normal state resistance, and the characteristic voltage, undergoes an irreversible degradation above 150 ℃ [79]–[81]. When the value of critical current within Josephson junctions varies, both the inductance of the junctions and the resonant frequency of the qubits can be away from their original designs. As such, low-temperature solder (In [12][27][82][83], In/Bi/Sn [84], In/Sn [85]) and adhesive

bonding [86] [87]are adopted in both methods and function as electrical links between dies and substrates. The comparison is presented in Table 4. Although the solder-based processes are compatible with quantum circuits, the interconnect pitches over 75 µm are still not dense enough. As an example, for two millions of I/O, at least a Si area of 11765 mm^2 is required, not including the carrier and the Si area for fan-out interconnects, connectors and inter-die spacing. With the volume constraint at the mK chamber, it is challenging to house such a sparse system, relevant huge count of coaxial cables and cryoelectronics. Therefore, in the Chapter 4, a new integrated and compact architecture applying the Superconducting Silicon interconnect Fabric with fine-pitch (≤ 10 µm) interconnect is introduced.

Table 4. Comparison between the existing flip-chip-demonstrated superconducting assembly methods

Assembly methods for SCE	**indium/bismuth/tin (In/Bi/Sn) solder reflow** [84]	**Indium/Tin (In/Sn)** [85]	**Adhesive bonding** [86]
Joints	Solder bumps	Solder bumps	Stud bumps
Temperature	Lowest to 70℃	140℃	< 180 ℃
Bonding Duration	-	-	-
Bonding pressure	-	-	-
Joint height	Over 20 µm	12 µm	10µm
Bonding strength	Below 10 MPa	-	-
Special treatment	140 ℃ Water soluble flux	Flux, two-step reflow	Adhesive underfill, reflow
Interconnect pitch	75 µm	> 100 µm	80 µm

3.3 Silicon-Interconnect Fabric (Si-IF) introduction

The Si-IF packaging platform adopts the ideas of wafer scale integration and heterogeneous integration, allowing heterogeneous unpackaged dielets to populate a Si wafer system through the flip-chip bonding. Si-IF technology, a board made with Si, aims to replace the classical PCB-based systems and enables integration of the system on a single packaging hierarchy. As compared to the organic substrate of PCB in the subsection 3.1.1, Si-IF uses Si as the substrate material, which is robust enough to reduce the warpage less than < 0.01% [88] in a 300 mm wafer. This allows I/O to scale down to a finer pitch less than 10 µm using Cu-Cu thermal compression bonding (TCB), as shown in Figure 15, and enables dielet to be placed in a proximal way with an inter-dielet spacing of 100 µm, as shown in Figure 16.

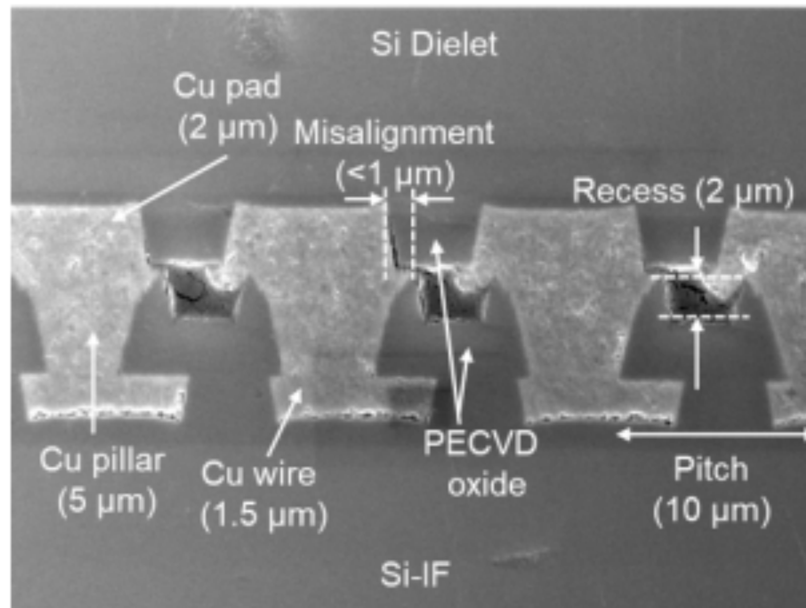


Figure 15. 10 um interconnect pitch on Si-IF. Reprinted with permission © 2021 IEEE [89]

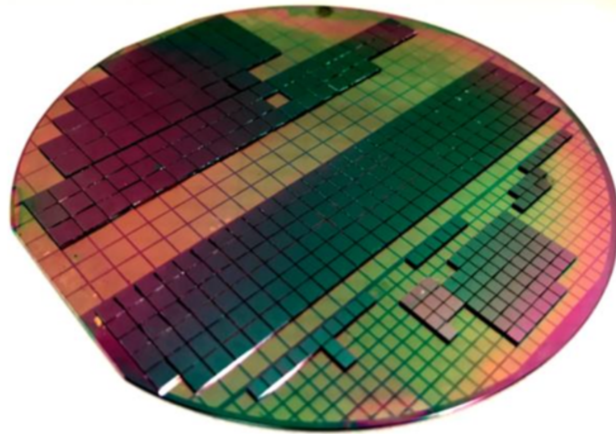


Figure 16 ~ 370 dies (1×1 mm^2 to 5×5 mm^2) at ≤ 100μm inter-dielet spacing on a 100 mm Si-IF. Reprinted with permission © 2018 IEEE [90]

The wiring dimensions can also be scaled down to ≤ 2μm. Due to a much compact system, the latency can be reduced to the pico-second level (through ≤ 100 μm inter-dielet spacing and ≤ 1 μm height of joints) and the bandwidth can be increased to over 3 Tbps per mm. Because of using dielet for integration (instead of packaged chips), the energy consumption per bit can be reduced to below 0.2 pJ/b. All of these scaling features result in a huge improvement in data bandwidth (120 to 300 X) and energy consumption per bit (100X) as compared to conventional PCB-based integration schemes [91], [92]. As compared to the interposer technology, Si-IF can still improve the data bandwidth (8 to 25 X) and energy consumption per bit (5-26 X). Si as a substrate material is an excellent heat conductor (149 W/m K), having a better thermal management as compared to conventional PCB packaging schemes. Using Si as the substrate material also provides another benefit: front-end-of-line high-tech toolings can be brought to the back-end-of-line process for advanced packaging development and fabrication. In a traditional PCB, there are different materials, such as Si, Cu, FR-4, solder, molding compound, underfill, etc., and multilayer packaging hierarchy. These materials and layers have mismatch thermo-

mechanical properties, causing Chip-Package-Interaction (CPI) related failures. For Si-IF, there are only materials same as that constitute dielets, such as Si, Cu, and SiO2, and this reduces the CPI failures. In Chapter 4, the powerful Si-IF is extended to the low-temperature region with four main changes as the superconducting Silicon Interconnect Fabric (Superconducting-IF) to be compatible with quantum devices.

3.4 Summary

In this chapter, packaging technologies, including traditional PCB assembly and advanced assembly through flip-chip heterogeneous integration, are introduced in the section 3.1. Next, in the section 3.2, the packaging technologies, including wire-bonding and flip-chip, are used in the quantum field. Among all the package technologies, it is common to apply large-pitch solder methods, which are not favorable for high-density I/O integration. Therefore, the advanced wafer-level integration Si-IF is developed having fine-pitch Cu-Cu TCB ($\leq$ 10 μm) and is detailed in the section 3.3. Using Si-IF, the latency can be reduced to the pico-second level (through $\leq$ 100 μm inter-dielet spacing and $\leq$ 1 μm height of joints); the bandwidth can be increased to over 3 Tbps per mm; the energy consumption per bit can be reduced to below 0.2 pJ/b. This powerful Si-IF technology is extended to the cryogenic field for quantum applications and is detailed in the Chapter 4.

Chapter 4

Concept of integrated quantum system on cryogenic Si-IF – Superconducting Silicon Interconnect Fabric (Superconducting-IF)

In this chapter, the cryogenic version of the Silicon Interconnect Fabric- Superconducting-IF is introduced briefly in the section 4.1, followed by the detail of the concept of quantum system integration on the Superconducting-IF from the aspects of thermal limits of qubit chips and the deep cryogenic chamber in the section 4.2. The comparison between the proposed concept in this thesis and the existing schemes shows in the section 4.3. The last part of this chapter is summary in the section 4.4.

4.1 Superconducting-IF introduction

As introduced in the section 3.3, Si-IF is a powerful system-on-wafer (SoW) advanced packaging technology that owns important characteristics, including picosecond latency, a large Si area, fine pitch interconnect (≤ 10 μm), short inter-dielet spacing (≤ 100 μm), metal-to-metal bonding, no solder and underfill, high bandwidth, and low energy consumption per bit. The mechanical die-to-substrate alignment accuracy can be achieved within ± 1 μm in the Si-IF technology. As compared to the interposer technology, Si-IF can improve the data bandwidth (8 to 25 X) and energy consumption per bit (5-26 X). This powerful Si-IF is extended to the low-temperature region as the Superconducting Silicon Interconnect Fabric (Superconducting-IF), as shown in Figure 17. Superconducting-IF will inherit all the features of Si-IF with four process

changes to accommodate superconducting control/readout chips and qubit chips. First, Nb is used as the superconducting interconnect material because it is mechanically hard, which indicates it can withstand high mechanical stress during thermal cycling tests. Besides this, Nb possesses the highest transition temperature among all existing metals. This allows it to have a larger critical magnetic field and a higher critical current, both of which are shown in the section 5.2. Both the maturity and already established techniques of superconducting digital processing make Nb a good candidate to apply on the Superconducting-IF platform. The interconnect material can be any superconducting material, such as Niobium Nitrite (NbN), Titanium Nitride (TiN), Niobium Titanium Nitride (NbTiN), and Yttrium barium copper oxide (YBCO), and is not only limited to Nb. Second, a thin layer ~5 nm of iridium (Ir) is passivated above Nb (in-situ deposition) to prevent Nb oxidation, which is detailed in the subsection 5.2.1, and to act as an Au diffusion barrier. Ir has the corrosion-resistant characteristic and it can prevent O_2 from penetration. The Ir film covers all the Nb interconnects, while the Au film is only deposited on the bonding pads. Third, Au is selected as the bonding medium since Au has oxidation-resistant, which is the most important feature for TCB, and soft characteristics and possesses sufficient diffusivity below 150 °C. One of the future possible applications of the Au bonding is to integrate with CMOS chips. With a proper system of Au diffusion barriers and sacrificial diffusion barriers, such as TiN or TaN or Ta, the used thin layer of Au is compatible with CMOS process. The integrity of these barriers is verified with the past experiments, in which Cu/Au capping is used for thermal compression bonding [90], [91] with the mentioned barriers on fully functional chips fabricated by two major foundries (in 16FF and 22FDX nodes). Last but not the least, the whole fabrication process is carried out below 150 °C, which is explained in the subsection 3.2, due to temperature-sensitive Josephson junctions. More detailed fabrication

process is presented in Chapter 5: realization of Superconducting-IF platform.

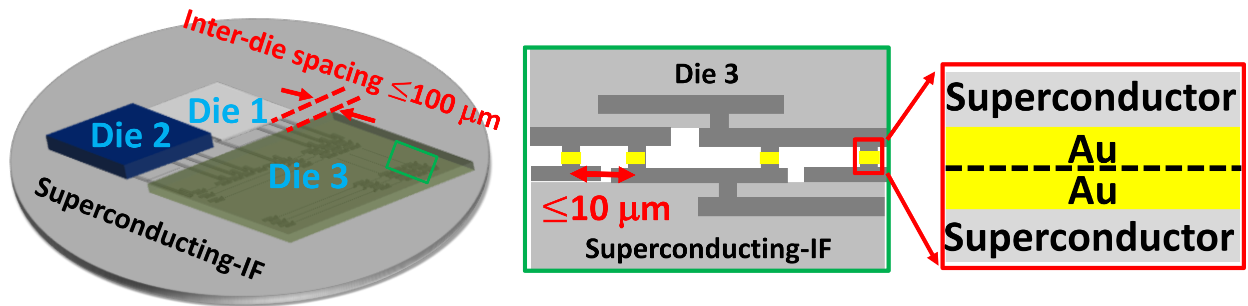


Figure 17. Superconducting-IF integration scheme with fine-pitch interconnects and proximal inter-die spacing. © IOP Publishing. Reproduced with permission. All rights reserved.

4.2 Proposed qubit integrated system on Superconducting-IF

As detailed in the Chapter 2 regarding the existing distributed system and the CryoCMOS approaches, besides the qubit processor for quantum operation, there exists a separate processor, in the form of either FPGA or CryoCMOS microcontroller, in charge of pulse generation/modulation and readout processing. This type of coprocessor operates either at room temperature (~ a few meters away from a qubit processor) or at 4K (~20 centimeters away from a qubit processor). In both situations, the coprocessors are far away from a qubit processor and rely on limited-number of coaxial cables carrying signals of 5-10 GHz for serial inter-processor communication. This is uncapable of minimizing the previously mentioned problems (I/O management, phase shift and latency, heat and power dissipation per qubit, and transmission of delicate quantum information) for a large-scale qubit array in the section 1.2. The best way is to combine the coprocessor and the qubit processor on the same substrate as an integrated system, analogous to a massive ENIAC computer invented in 1945 advancing to an integrated CMOS system [93]. The instant benefit of this integration is replacing serial coaxial cables with parallel interconnects for communication between the qubit processor and the coprocessor, reducing significantly the issues of phase shift and latency.

To integrate circuits in charge of pulse generation/modulation and readout processing with a qubit processor at a millikelvin chamber, the first issue to solve is that thermal phonons generated by the integrated circuits cannot affect the delicate qubit state. Based on this, Savin *et al.* [94] points out five thermal designs, as shown in Figure 18, related to the heat dissipated by the circuit of different complexity: (1) for the controlling circuits generating < 50 nW of heat, as shown in Figure 18 (a), controlling circuits can be designed nearby the qubit circuit on top of a 5 X 5 mm^2 chip and have direct on-chip interconnects, (2) for the controlling circuits generating <

500 nW of heat, as shown in Figure 18 (b), thermal insulation design is required to interrupt the spreading of thermal phonons between the controlling circuits and the qubit circuit on top of a 5 X 5 mm^2 monolithic chip for direct on-chip communication, (3) for the controlling circuits generating < 10 μW of heat, as shown in Figure 18 (c), the controlling circuits and the qubit circuit should sit on two separate Si chips and communicate through wire-bonding but can still on the same chip holder, (4) for the controlling circuits generating >10 μW of heat, as shown in Figure 18 (d), the controlling circuits and the qubit circuit should sit on two separate Si chips on two separate chip holders and communicate through wire-bonding, and (5) for the controlling circuits generating unlimited heat and the chamber temperature is kept at millikelvin, as shown in Figure 18 (e), controlling circuits and the qubit circuit should communicate through inductive or capacitive coupling without thermal conduction path between them.

There is a work by Microsoft [23], [95], mentioned in the section 2.2 following the thermal design (d). That work has a qubit CMOS controlling circuit using the low-power version of 28 nm FDSOI technology with 100,000 transistors operating at few MHz and at 100 mK on a separate Si chip wire-bonding to a few-qubit processor. On average, each low-power 28 nm FDSOI CMOS transistor consumes ~0.2 nW at 20 mK. With the advanced 7 nm FinFET, each transistor can reduce the power consumption to ~0.055 nW (55 pW) at 20 mK with a frequency of 5 MHz, which is extrapolated from the supercomputer Fugaku having a transistor count of 8.786 billion per node [96] and an overall node of 158,976 [97] and consuming overall power of a 29899 kW. For a large-scale qubit array, the thermal design (d) is still challenging even using the most power-efficient 7 nm FinFET technology since the CMOS count has to add up for circuits with much complex controlling functions. The overall heat of the increased number of CMOS will exceed the acceptable heat threshold in the thermal design (d) and will warm up the

substrate carrying a qubit processor.

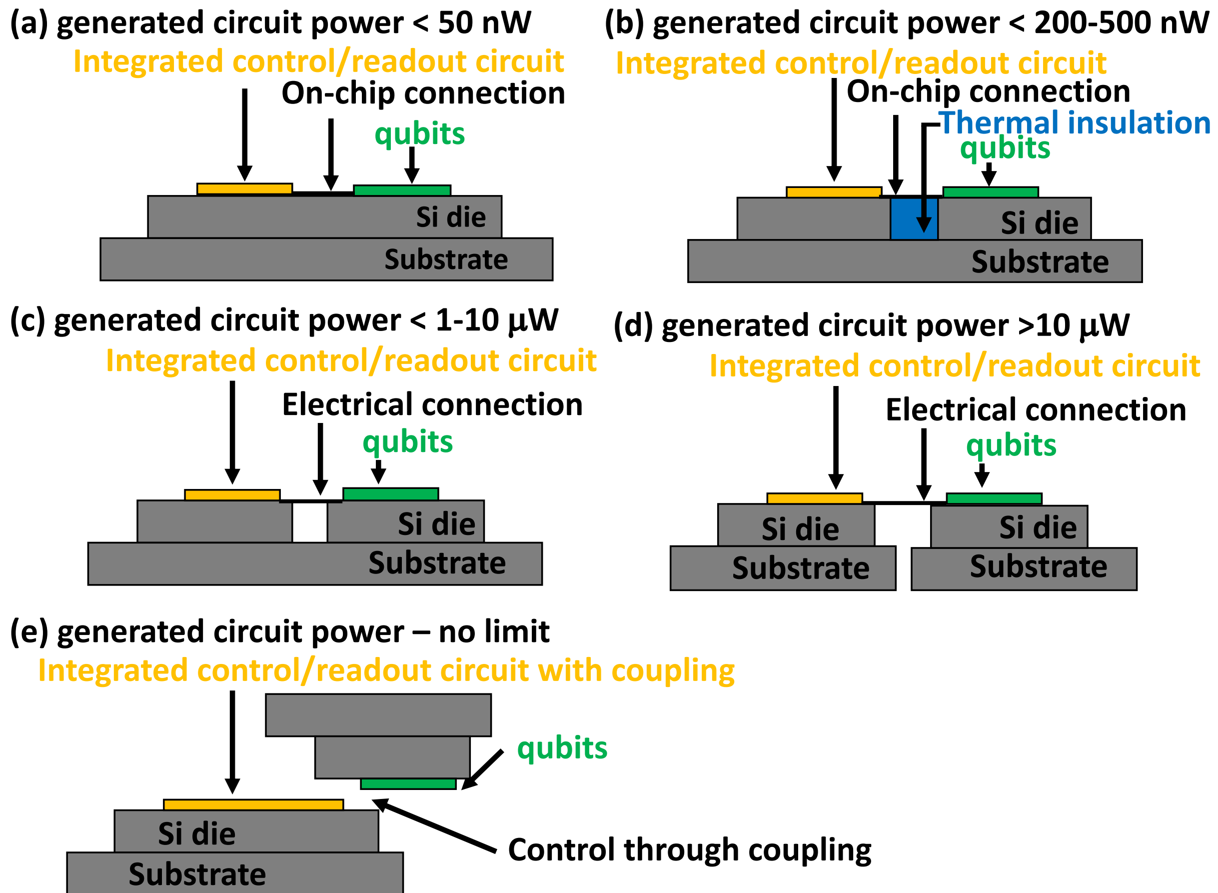


Figure 18. Thermal design with qubits based on the generated power of the control/readout circuits [94].

Besides the low-power CMOS technology, there is another branch of low-power technology capable of dissipating less power than the CMOS technology, which is the previously introduced JJ-based superconducting digital circuit. The estimated power dissipation of JJ-based family operating at 20 mK is presented earlier in Table 3 in the section 2.4. The power dissipated by the JJ-based superconducting digital circuit can be in the single digit of pW per JJ for the low-power SFQ technology and less than 0.1 pW/JJ for the AQFP technology, both of which consume 10 X

to 100 X less power than the advanced 7 nm FinFET does.

Combining the concept of thermal design (a) & (b), JJ-based superconducting digital circuit, and the flip-chip packaging technology, quantum system integration on the SoW Superconducting-IF is proposed. To integrate JJ-based control/readout systems with qubits through the thermal design (a) & (b) on the SoW Superconducting-IF, the maximum dissipated power should be less than an overall value of 500 nW and the system architecture is proposed in the following paragraphs. One thing to note is that the limiting power of 500 nW refers to the power dissipation in the monolithic way of system realization within the thermal design (a) & (b). Since there is no literature reporting relevant thermal budget for a system realization using the heterogeneous flip-chip integration, through which thermal phonons cannot spread through a common Si substrate on the dielet part [38], the limiting power of 500 nW is borrowed from the monolithic system realization [94] and is used for the calculation in the following paragraphs.

As introduced in the section 2.4, JJ-based technologies are demonstrated to be capable of qubit control [37]/readout [40] through multiple quantized pulses. To further expand this idea for a large-scale qubit array, separate dielets with the different functions are required, such as clock [98], pulse generation & modulation [99], memory& signal processing [34], multiplexer (MUX) [100], demultiplexer (DEMUX) [100], analog to digital converter (ADC), digital to analog converter (DAC) [101], amplifier (Josephson Parametric Amplifier, JPA) [102], and readout (Josephson Photo Multiplier, JPM) [40]. Since the concept of qubit control/readout through JJ-based circuits is very new, no literature realizes the control and readout on a single qubit on the same substrate at the time of writing this thesis and none of the work estimates how many JJ are needed to control each qubit with an integrated system for a large-scale qubit array. Based on the power dissipation, a rough calculation regarding the number of qubits that can be housed in an

existing dilution refrigerator is discussed as follows. To start the calculation, the thermal boundary conditions should be clarified and there are two competing constraints: one is the upper-bound power dissipation of 500 nW in the design (b) of an integrated qubit system; the other is the chamber cooling power of rough 20 μW at 20 mK stage [13]. Since the JJ count is not pointed out in each above-mentioned function block, the JJ count of 21460 in the function block of memory & signal processing [34] is assumed and extended to other eight basic blocks, leading to an overall JJ count of ~ 195,000. Indicated by the qubit control [37] and readout [45], one qubit control/readout needs at least 10 JJ (4 JJ for control and 6 JJ for readout and under the assumption of linear scaling), meaning a qubit number of 19,500 can be housed according to all the assumption. The overall power consumption of 195,000 JJ with the most power-efficient JJ based technology is ~ 4 nW, which is much smaller than the upper-bound power dissipation of 500 nW in the design (2). Regarding the required coaxial cables, assuming each wire-bonding connects to one coaxial cable, 28 cables are suggested by the figure in the reference [34], leading to an overall 252 cables for nine blocks. Each cable consumes on average 0.02 μW [13], indicating ~ 5 μW for 252 cables in the chamber. Assuming the overall cooling power at the mK chamber can all be consumed by the cables, it is achievable to have an overall cable number of 1008, equivalent to a JJ count of 780,000 and a qubit count of ~ 78000. Referring to the two competing constraints (the upper-bound power dissipation in the on-chip design (b) and the chamber cooling power at 20 mK stage), the number of qubit is actually limited by the chamber cooling power, not the limit of on-chip power. The number of qubit that can be integrated in this proposed approach on the Superconducting-IF is much higher than that reported by Krinner *et al.* [13] (~ 1000).

Integrating circuits with qubits on the same substrate generates noise to qubits and there are

methods to reduce noise-related errors. Since this thesis focuses on the system level integration, the methods are briefly introduced. For passive methods, there are many possible ways and two of the approaches are introduced in the following sentences. One is to use a periodic Al structure around qubits to capture phonons and prevent qubits from decaying [103], which can potentially increase the qubit lifetime by two-order of magnitude to over 100 μs. Another method is to deposit normal metal Cu trap to capture non-superconducting particles [104], through which qubit decay can be improved by 30X. For active methods, one of the ways is to adopt the interferometric Josephson isolator [105], which can provide more than 20 dB of protection against amplified noise. Besides the protecting structures around qubits, there is another way to reduce noises: using the noise-resistant qubit architecture with a gyrator [106]. This new qubit structure provides energy levels not dependent on noises and it is naturally protected against the common noise channels in superconducting circuits.

In the integration scheme in design (b), there exists a thermal insulation structure between the control/readout circuits and the qubits. One of the approaches to realize the thermal insulation structure is through the inlay demonstration in Si substrate reported by Dasgupta *et al*. [107], in which a quartz inlay (thermal conductivity: 0.78) is used. The process is as follows: (1) Si hole etching, (2) metal seed layer deposition inside the holes and on quartz dies, (3) quartz dies assembly inside the holes through metal-to-metal bonding, (4) metal electroplating to fill the gaps, and (5) lapping and polishing to get a flat surface for future processing. Besides the quartz inlay, other types of low thermal conductivity material, such as glass fiber (thermal conductivity: 0.043), can also be applied through the inlay approach.

The next step is to connect the wafer-level system to the outside world. Besides the bulk SMA connectors, there are miniaturization of connectors, such as Hg micro-connector [108], spring

probes[109], and fuzz buttons. The key for these connectors is that they can withstand mechanical cooling stress, or thermal shock, and make strong contacts between electrodes even at low temperature to continuously conduct electrical signals. For Hg micro-connector, due to the liquidus behavior of Hg, it can flow in the liquid form and fill any gap between electrodes within connectors. As the temperature drops below 33 K, liquid Hg freezes and becomes solid Hg, forming strong electrical connection. For spring probes and fuzz buttons, they have the feature of spring, capable of elastic deformation without fracturing at the same time remaining electrical connections. The Hg micro-connector, spring probes, and fuzz buttons are good candidates to fix external cables on the wafer-level system and provide miniaturization to ~ 100 μm.

After solving the miniature issue of connectors, the next issue is the mating cables for the miniaturized connectors. Besides the common bulk coaxial cables, Hamilton *et al.* have developed flexible cables [110] carrying multiple parallel microstrip lines/striplines on a flexible substrate with a dimension of few centimeter. On these flexible cables, HD4100 polyimide (PI) are used as the flexible substrate carrying microstrip lines/striplines made with high transition temperature (Tc) superconductor YBCO and low Tc superconductor Al and Nb with the overall thickness of flexible cables around 50 μm. The length normalized insertion loss is well less than 0.1 dB/cm and the cross-talk behavior is below ~ -60 dB because of low-loss PI. These flexible cables have passed five times of thermal cycle test from 4 K to room temperature with consistent RF performance after each cycle. These flexible superconducting cables are good candidates for the wafer level integration since they can realize meter-level signal transmission, can connect a low-temperature chamber to the next higher temperature chamber, and can be anchored by the miniaturized micro-connectors at the wafer edges. With miniature connectors at the edge of wafers and on-wafer superconducting multiplexing (1-to-8 multiplexing or 8 bit multiplexer

[100]), each microstrip line on a flexible cable can replace eight bulk coaxial cables, significantly increasing the number of qubits inside a dilution fridge. As an example, with 8-bit superconducting multiplexers, the qubit count in the system can be added up to 624,000 from the original count of 78,000 using bulk coaxial cables. Through continuous development of flexible cables and the 16 bits superconducting multiplexing, controlling over one million qubit array is possible and is not far-reach. The relevant key numbers to achieve a large-scale qubit array is summarized in Table 5.

Table 5. Potential way to achieve one-million qubits integrated on the superconducting IF

	C. Ayala [28]	Extrapolate to nine functional blocks at mK	Maximum JJ count limited by fridge cooling capability	Maximum JJ count by superconducting multiplexing	
				8 bits [100]	16 bits [100]
JJ count	21460	~ 195,000	~ 780,000 (78,000 qubits)	~ 6,240,000 (624,000 qubits)	~ 12,480,000 (1,248,000 qubits)
JJ power dissipation	0.45 nW	4 nW	16 nW	128 nW	256 nW
Cable count	28	252	1,008	1,008	1,008
Cable heat load	~ 0.56 mW	~5 mW	~20 mW	20 mW	20 mW

4.3 Comparison with existing scheme

Qubit integration with control/readout systems on the Superconducting-IF provides an alternative to solve the previously mentioned problems: I/O management, latency and phase shift, heat and power dissipation per qubit, and transmission of delicate quantum information. Using the fine pitch (≤ 10 μm) interconnects on the Superconducting-IF provides high-density I/O (≥ 10,000 per mm^2). Through proximal integration (100 μm) of control/readout dielet and qubit on the same wafer and ≤ 1 μm height of joints, control/readout signals can be transmitted on a wafer at the mK chamber and the latency can be reduced to the pico-second level. The latency of qubit control/readout cycle can be reduced to nanosecond. Since the pulses are generated on the Superconducting-IF and are close to qubits (much smaller than the wavelength), the phase shift is not severe as compared to a distributed system. For the issue of heat and power dissipation to drive one qubit, with superconducting multiplexing techniques and the integrated superconducting system, the heat from the cables, the cryogenic electronics, and the circuits is shared by multiple qubits, leading to an overall reduction in the heat and power dissipation. The architecture comparison between the concept proposed in this work and the convention method is shown in Figure 19. The FPGA coprocessor, and control/readout electronics sit at room temperature while only qubit processor sits at deep cryogenic environment. For one control/readout cycle, signals start with FPGA, then to control electronics, long distance to qubit processor, and eventually long distance back to readout electronics with quantum information. The concept proposed in this work is to move the coprocessor and the control/readout electronics to the deep cryogenic environment and to integrate them with the qubit processor on the Superconducting-IF platform.

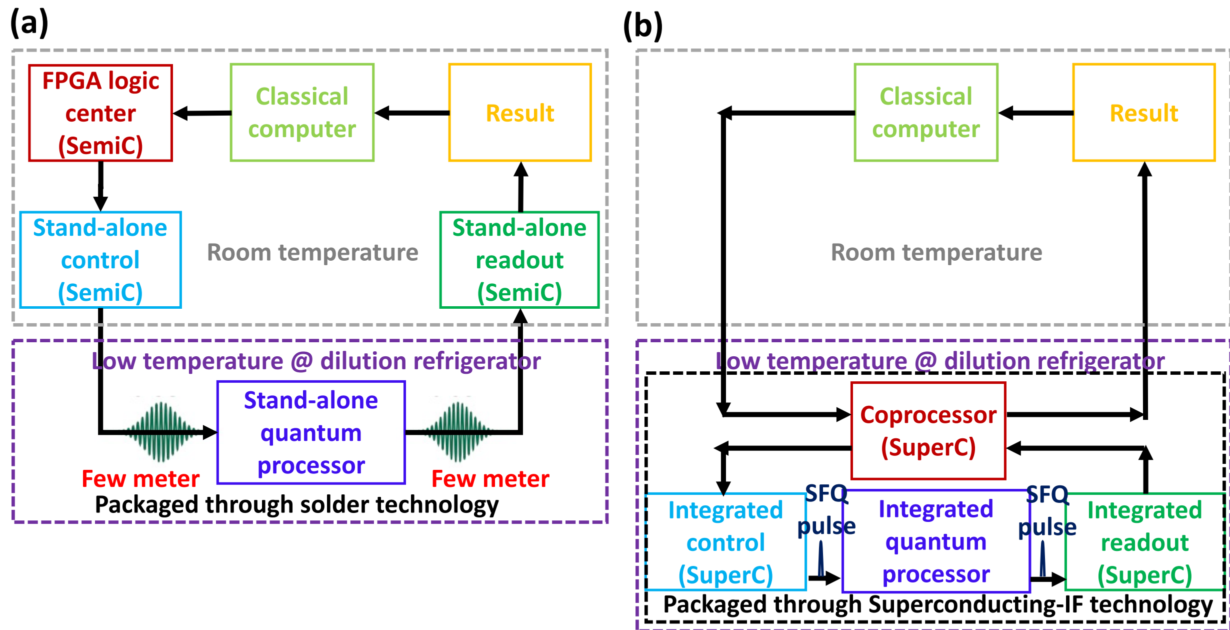


Figure 19 architecture comparison between this work and the convention method [111]

Table 6 comparison between the concept proposed in this work, the conventional scheme and the work of Google, Microsoft, and Intel

[110]	This work	Conventional	Google [20-22] (2020)	Microsoft [23] (2021)	Intel [24] (2021)
Technology	SFQ @ mK	CMOS @ RT	CryoCMOS @ 4K	CryoCMOS @ mK	CryoCMOS @ 4K
Qubit control source	Digital SFQ pulse @ mK	Analog @ 300K	Analog @ 4K	Mixed signal @ 100 mK	Mixed signal @ 4K
Qubit readout	Digital to mK	Analog to RT	Analog to RT	Analog to RT	Analog to 4K
Qubit type	Superconducting	Superconducting	Superconducting	Topological (goal)	Spin
clock	20-40GHz (goal)	0.3-2 GHz	0.5 GHz	Few MHz	1.6 GHz
Power	0.2 uW/qubit	20mW	2.5mW	20 uW (10 MHz & 100k CMOS)	439.2 mW
Latency (control/ readout cycle)	Nanoseconds	Milliseconds	Milliseconds	Milliseconds	Microsecond (potential)

For one control/readout cycle, instead of going back and forth between the room temperature environment and the deep cryogenic environment, the qubit control and readout are implemented in-situ at the deep cryogenic environment, shortening control/readout cycle latency & phase shift

and providing much faster turnaround speed. As compared to the conventional work, the detailed comparision between the concept proposed in this work, the conventional scheme and the work of Google, Microsoft, and Intel in Chapter 2 is presented in Table 6. The significant difference is the latency per control/readout cycle and the average power required to drive a qubit. As mentioned in Chapter 1 and the section 4.2, due to proximal integration of control/readout coprocessor near qubit chips on the same substrate at the mK chamber, the signal transmission distance is greatly reduced from the meter scale (between a qubit chip and room temperature electronics) and the centimeter scale (between a qubit chip and 4 K cryostat chips) to the millimeter or even micrometer scale (within the same cryostat). The relevant latency and phase shift can therefore be decreased to nanoseconds. For the average power required to drive a qubit, since superconducting multiplexing methods are applied and the relevant control/readout parts are closer, the number of I/O cables can be significantly reduced and each cable can control multiple qubits with a much better signal power efficiency.

4.4 Summary

In this chapter, the cryogenic version of the Silicon Interconnect Fabric - Superconducting-IF is introduced briefly in the section 4.1. Superconducting-IF will inherit all the features of Si-IF, including picosecond latency, a large Si area, fine pitch interconnect ($\leq 10\ \mu m$), short inter-dielet spacing ($\leq 100\ \mu m$), metal-to-metal bonding, no solder and underfill, high bandwidth, and low energy consumption per bit. As compared to the interposer technology, Si-IF can improve the data bandwidth (8 to 25 X) and energy consumption per bit (5-26 X). The Superconducting-IF uses Nb/Ir as the interconnect material for its excellent property and stability during process. The Au interlayer bonding is selected for its capability for low-temperature bonding at a JJ-compatible temperature of below 150 ℃. In the section 4.2, the detail of the concept of quantum system integration on the Superconducting-IF from the aspects of thermal limits of qubit chips (< 500 nW) and the deep cryogenic chamber (< 20 μW) is described and the maximum count of qubits under the thermal constraint is calculated to be 78,000. Through the development of flexible cables and the superconducting multiplexing, controlling over one million qubit array is possible using the quantum system integration on the Superconducting-IF. In the section 4.3, the comparison between the proposed concept in this thesis and the existing schemes is presented. Due to proximal integration of control/readout coprocessor near qubit chips on the same substrate at the mK chamber and superconducting multiplexing methods, the latency per control/readout cycle and the average power required to drive one qubit can potentially reduced to nanoseconds and 0.2 μW.

Chapter 5

Wafer-level Superconducting-IF realization

For the future wafer-level integration of qubit arrays with proximal control/readout systems on Superconducting-IF, the wafer-level Superconducting-IF packaging platform should be realized as the first step. The step-by-step realization is detailed in this chapter. On the Superconducting-IF, it consists of two major parts: one is the Nb interconnects; the other is the Au interlayer fine-pitch bonding technology. The deposition and fabrication of Nb interconnects is presented in the section 5.1, including the issue of Nb oxidation and the passivation method to reduce the Nb oxidation for further processing (subsection 5.1.1) and the logistics to select bonding medium (subsection 5.1.2). The Au interlayer fine-pitch bonding technology is introduced in the section 5.2 covering pre-bonding check (subsection 5.2.1), Au-Au bonding optimization for low-temperature applications in the form of design of experiment (subsection 5.2.2), and relevant electrical test (subsection 5.2.3). Since part of this project is in collaboration with MIT Lincoln Lab, samples with Au interlayer are shipped to UCLA and flip-chip bonded with the developed Au interlayer technology. The bonding results is presented in the section 5.3. After all preliminary experiments showed excellent results, a wafer-scale flip-chip assembly was demonstrated, as shown in the section 5.4, with the Au interlayer technology on Nb interconnects with various die sizes from 2 X 2 mm^2 to 5 X 5 mm^2, emulating an integrated qubit control/readout system with fine-pitch interconnects and a proximal die spacing (≤ 100 μm). On this wafer-scale assembly, there is a large Si area and space for a large-scale qubit accommodation, die-to-die spacing, fine-wiring fan in and fan out, qubit noise cancellation methods and connections to the outside world. To populate the wafer-scale assembly, an

automated bonding procedure is introduced with an average bonding duration of 8.5 s for each dielet. The summary of this chapter is presented in the last section 5.5.

5.1 Deposition and fabrication of Nb interconnects

As mentioned in the section 4.1, Nb is selected as the interconnect material for its maturity and easy processing. One major issue is the oxidation problem after processing, making the superconductivity disappear. In the subsection 5.1.1, the Nb oxidation behavior is detailed, and in the same subsection, it is explained the Nb surface is passivated with Ir to protection its superconductivity. In the subsection 5.1.2, since the Nb-Nb direct bonding was tested to be challenging for a low temperature (<150℃) implementation, various methods were verified and the Au interlayer technology standed out due to its fine-pitch capability and high bonding strength at a low bonding temperature.

5.1.1 Nb oxidation issue and top passivation layer

As reported by literatures [112], [113], when pure Nb metal is exposed to air, the surface Nb layer interacts with O_2 gas and becomes Nb oxide (Nb_2O_5, NbO, $NbO_{0.2}$) with a thickness of around 2 nm. As part of the surface Nb oxide grows with a larger grain size, the lattice-mismatch-induced stress between the Nb oxide and beneath Nb (Nb: 3.3 A; Nb_2O_5: 3.9 A) drives O_2 to diffuse deeply into bulk Nb below. This is not the same self-limiting process as the oxidation behavior of Al, forming a dense Al oxide to prevent O_2 diffusion further, but is similar to the oxidation behavior of Cu. This indicates the whole Nb film can get oxidized and causes the superconductivity of Nb to disappear. In addition to the lattice-mismatch stress, temperature is another factor causing O_2 to penetrate and reacts with Nb. When Nb is heated over 150 ℃ for a long time, the O inside Nb diffuses faster based on the Arrhenius equation. Through the Fick's

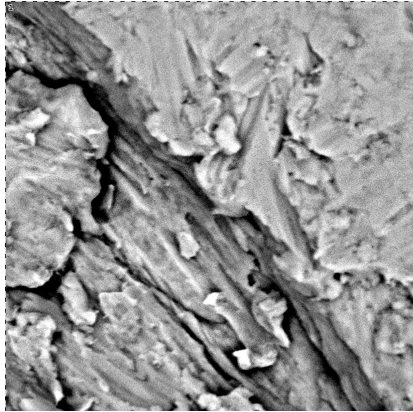
law, the final diffusion profile is a Gaussian distribution from the surface. The O concentration at the surface is 8 at%, which is the maximum solubility of O in Nb. Within the first diffusion length, the concentration is 4.88 at%. After three diffusion lengths, the O concentration is below 1 at%.[114], [115]. The transition temperature of Nb drops 1K when 1 atomic percent (at%) of O exists inside Nb.

Besides the intrinsic oxidation behavior of Nb pure metal, extrinsic causes, such as the deposition tool and the source problem, can lead to non-superconducting Nb film. At UCLA, evaporation tools (SLOAN and CHA mark 40-2) were tested for high-quality Nb films but due to their low-vacuum processing chamber (10^{-6}-10^{-7} torr) and the periodic shift of metal pellets inside the chamber, no transition temperature was measured. To breakdown the source of Nb oxidation, the EDX material analysis was tested on source pellets and a blanket Nb film on a Si wafer, as shown in Figure 20, with proper vacuum storage process before analysis. Both the source pellets and the blanket film were oxidized and the oxygen content in the film was twice of that in the pellets, meaning extra oxygen contaminated the film during deposition and the tools did not isolate oxygen properly. Another batch of samples with Nb film on Si were deposited using fresh Nb pellets in both evaporation tools, no Nb superconductivity was observed. Sputtering tools (Ulvac Jsp-8000 and DC magnetron) were used for testing. With the two-chamber design and a high-vacuum processing chamber (10^{-7}-10^{-8} torr), the transition temperature of Nb was successfully measured at 9K, as shown in Figure 21, close to the intrinsic Tc value of Nb. To further maintain the Tc of Nb and to prevent the Nb intrinsic oxidation behavior during later processing (DI water cleaning, dehydration baking, PR coating/baking/patterning/stripping), a thin layer of protective Ir was deposited on top of Nb sequentially in the same chamber without breaking the vacuum. The deposition parameters were

as follows: the power and pressure for Nb was 200 W and 5 mTorr, and growth rate was 0.57 A per second; the power and pressure for Ir was 100 W 3 mTorr, and growth rate was 0.83 A per second.

(a) Nb pellets for evaporation

Element Number	Element Symbol	Element Name	Atomic Conc.	Weight Conc.
41	Nb	Niobium	46.00	85.27
6	C	Carbon	31.49	7.55
8	O	Oxygen	22.51	7.19



(b) Blanket structure

Element Number	Element Symbol	Element Name	Atomic Conc.	Weight Conc.
8	O	Oxygen	44.52	16.38
41	Nb	Niobium	32.05	68.48
14	Si	Silicon	23.43	15.14

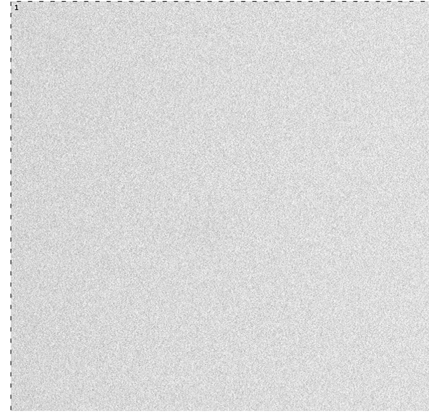


Figure 20. EDX elemental composition analysis on (a) a source pellet and (b) blanket Nb structure on a Si wafer. The oxygen content in the Nb film is twice of the oxygen content in the pellet.

Based on the phase diagram of Ir-Nb [116] under thermal equilibrium, the maximum solubility of Ir in Nb is 22 wt%. In our case, the wt% of Ir is 12%, which falls in the Nb rich region. The measured Tc, as shown in the subsection 5.2.3, was close to the intrinsic Tc value of Nb. This indicates even though there is some alloy formation, the amount of the alloy is negligible since the Tc does not get affected. The Ir layer is a protective layer for the Nb interconnects to prevent oxidation but does not play a role in the die assembly process, which is introduced in subsection 5.2. Ir is not the only capping material that can protect Nb from oxidation; there are another

materials, such as molybdenum rhenium (MoRe) [117] and Al_2O_3 [118], reported to be used on top of Nb. Au and Al are also possible candidates to prevent Nb from oxidation, but to adopt the process of record in the magnetron sputtering tool and to avoid possible contamination between existing and newly bought targets, Ir is the final solution of the passivation film on top of Nb in this thesis.

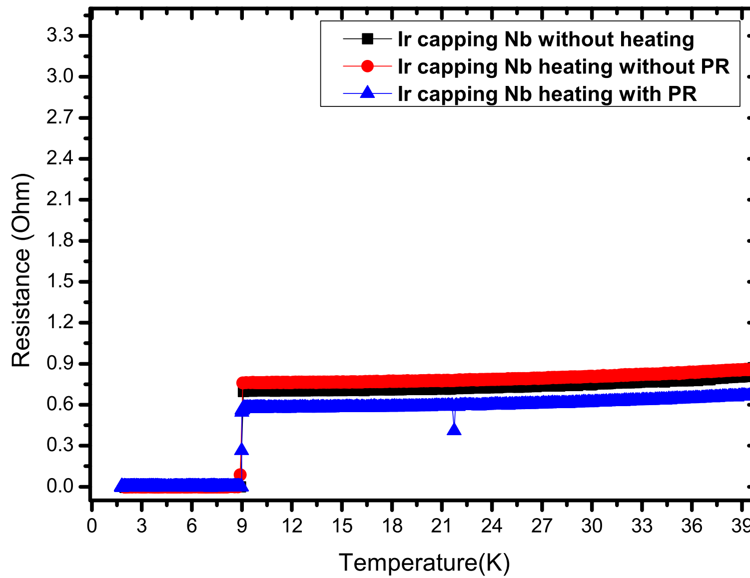


Figure 21. Preliminary Tc test of Nb/Ir samples with the protection of Ir against Nb oxidation during heating up. The Tc was maintained at 9K without significant degradation.

5.1.2 Selection of proper flip-chip bonding medium

The adequate bonding medium for the large-scale superconducting application requires a low temperature assembly, high bonding strength, and the fine-pitch capability. The challenges for Nb-to-Nb direct bonding are its low diffusivity at <150℃ (diffusivity $\ll 10^{-30}$ cm^2/sec within polycrystal structure [119]) and robust surface oxide. Nb is a refractory metal having a high melting point (>2400℃), leading to a strong and high energy inter-atomic metal bonding. Robust

Nb oxide is hard to break through using a high pressure and the byproduct, niobium fluoride, after the treatment of formic acid requires a high temperature (200-300℃ [120]) to dissociate. A trial experiment was implemented on a Nb sample with a 10-second in-situ treatment of formic acid. Both the testing and the control sample de-bonded very easily without bonding imprint and reaction at the interface, as shown in Figure 22. Besides the Nb-Nb direct bonding, solder materials were explored to adhere superconducting die. Although dies were assembled successfully through In-Sn solders (eutectic bonding) and In-Sn-Au (solid-liquid interdiffusion bonding, as shown Figure 23), the bonding force was much less than the military spec of 50 N on 2 X 2 mm^2 dies and the solders could not achieve fine-pitch applications. For direct Cu-Cu bonding, although the resistivity of Cu can be lower than that of Au, most reliable Cu-Cu bonding approaches require the temperature much higher than 150℃ for a long bonding duration. This can degrade the performance of Josephson junctions that constitute superconducting qubits. There exist a few Cu bonding approaches that can be done below 150℃ but these methods are difficult to redo at UCLA because of limitations from tools and materials. Therefore, a low-temperature Au-to-Au interlayer bonding technology was developed, as detailed in the subsection 5.2.

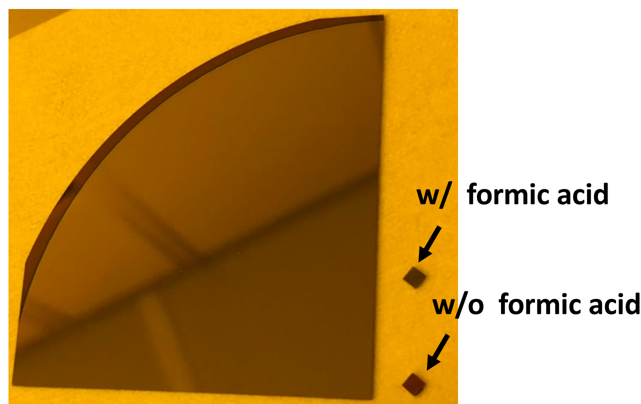


Figure 22. Nb-Nb bonding test with the control and experiment set. No bonding happens in both

conditions.

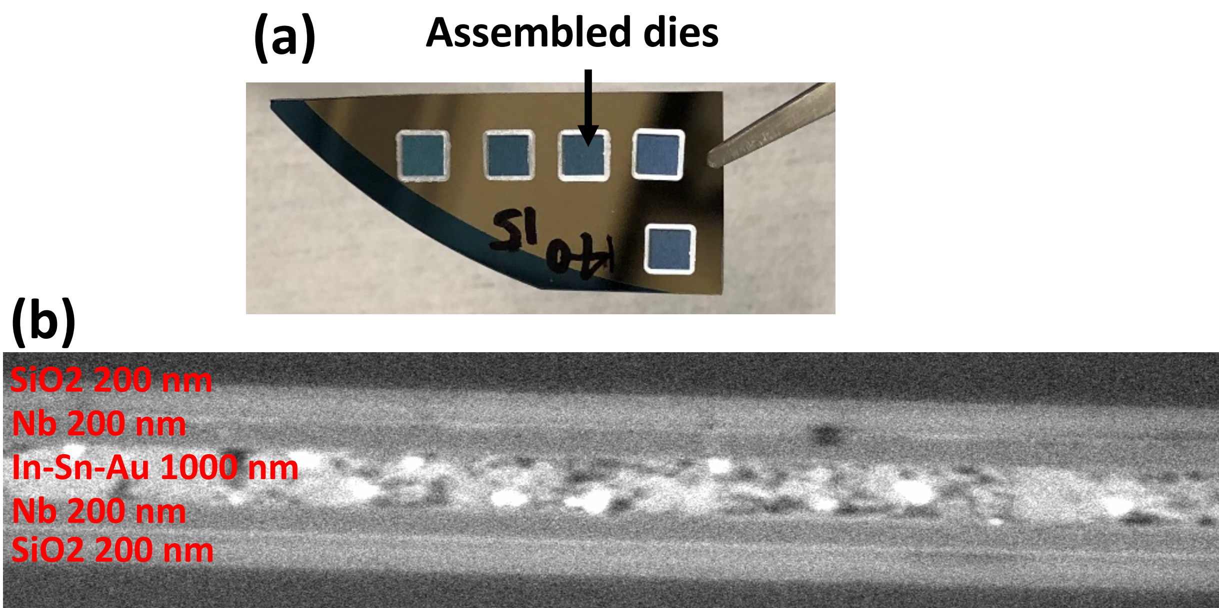


Figure 23. Successful In-Sn-Au solid-liquid interdiffusion bonding of superconducting dielets: (a) top view of the samples, (b) SEM image of the cross section of In-Sn-Au solid-liquid interdiffusion bonding at bonding interface.

5.2 Au interlayer technology on Nb interconnects

The Au-to-Au bonding, also know as the Au interlayer, was developed for low-temperature and fine-pitch applications. Before applying Au-to-Au bonding on Nb interconnects, a checklist was examined, as presented in the subsection 5.2.1, including the superconductivity of Nb after each process, the bulk adhesion between Nb and the underlying film, and the surface roughness for a good bonding. In the subsection 5.2.2, the bonding design-of-experiment (DOE) was implemented following Taguchi's method and the Au interlayer was optimized through the shear force using a blanket structure regarding the bonding parameters (bonding pressure, bonding interface temperature, and bonding duration). After the Au interlayer bonding was optimized, a large-pitch Kelvin testing structure was fabricated, as shown in the subsection 5.2.3, for the optimization of fabrication process, electrical behavior, such as Tc, contact resistance at the bonding site, critical current, and temperature cycling. Next, a fine-pitch daisy chain structure, as shown in the subsection 5.2.4, was fabricated for the demonstration of I/O pitch ≤ 10 μm. To test the reliability of the Au interlayer bonding, temperature cycling of the flip-chip dielets with multiple bonding area was conducted. After the Au interlayer was tweaked to be reliable, the Au interlayer was then applied on superconducting Nb interconnects for the measurements of critical current and temperature cycling.

5.2.1 Pre-bonding check

First, a theoretical calculation of the diffusion length of Au into Nb was investigated below the temperature limit of Josephson Junction at 150°C. This calculation is an important metric to conceptually study the depth of Au diffusion into Nb under the processing condition after Au deposition on Nb/Ir and after TCB at a temperature below 150°C. This theoretical investigation

shows that the superconductivity of Nb will not be altered after processing and is later verified through the electrical characterization as shown in the subsection 5.2.3. The calculation of the diffusion of Au into Nb/Ir starts with the estimation of diffusion length, which is related to the square root of atomic diffusivity and diffusion duration. It is reported by Esser and Christou [121] that below the temperature at 350 °C, faster grain boundary diffusion dominates, instead of slower bulk diffusion. The diffusion formula is an Arrhenius relation, including the pre-exponential term and the activation energy. Based on the information indicated by Esser and Christou [121], the pre-exponential term and the activation energy are 7.02×10^{-15} cm^2/s and 0.266 eV for Au diffusion into Nb. The calculated diffusivity at 150°C is in the order of 10^{-18} cm^2/s and the corresponding diffusion length (σ) is around 3.23Å with an overall processing time of 2 min. According to the model of metal interdiffusion and the solution to the error function with an Au interfacial concentration of 50 at%, the Au concentration inside Nb in the third σ, which is around 1 nm, is much less than 1 at%, as illustrated in Figure 24, (with 1 atomic % of Au in Nb, Tc decreases 0.3 K) while the overall Nb layer is 100 nm, meaning the contaminated Nb depth is only 1% of its total thickness. This worst-case theoretical scenario indicates the Au diffusion into Nb is negligible. In real samples, the amount of Au diffusion into Nb is much less than the worst-case scenario since the bonding temperature is lower than 150°C (exponential decay of diffusivity with temperature) and in the bonding area, Au does not contact Nb directly since there is a layer of Ir in between, which acts as an effective Au diffusion barrier. Literature by Liehr *et al.*[122] also indirectly supports the statement that Au diffusion into Nb is negligible below 150°C: after prolonged annealing at 450 °C for two hours, only ~ 3% of the Au film diffusion into Nb is observed. As compared to Liehr *et al.*[122], our case reduces the driving force of Au diffusion in two aspects: the processing temperature is lower than 150°C, at which

the Au diffusivity is reduced exponentially compared to that at 450 °C; the bonding time duration is much shorter than 2 hours - about 1.5 seconds. Thus, Au diffusion into Nb should be insignificant and the superconductivity of Nb should be unaltered after Ir and Au deposition and after Au-Au TCB, which is verified in our experiments.

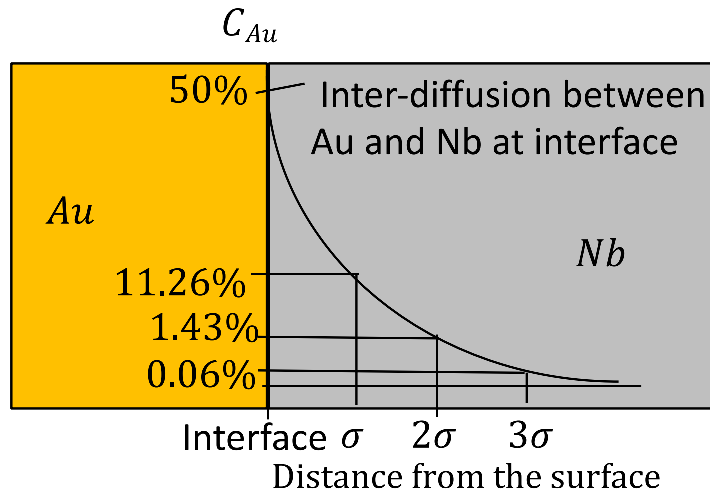


Figure 24. The worst-case scenario of Au atom diffusion into Nb film below 150 ℃ without Ir film in detween. © 2020 IEEE

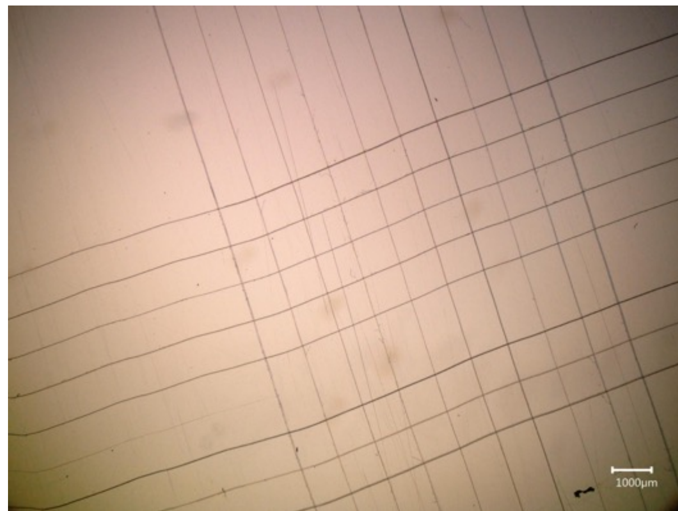


Figure 25. Adhesion test between the Au and the Nb/Ir film through tape-adhesive test. © 2020 IEEE

The adhesion between Au and Nb/Ir was first evaluated through a tape test. In the standard tape test, a specialized tool with ten blades in the front head was used twice to scratch the surface of the samples. After the first scratch, the sample was then turned 90° clockwise and the second scratch was implemented. Finally, a strip of scotch tape (3M) was adhered on the scratched sites and was peeled off. No flakes were observed on the surface of the samples, meaning the adhesion between Au and Nb/Ir was good throughout a large area, as shown in Figure 25. While somewhat primitive, this "tape-adhesive test" is used widely and is necessary but not sufficient condition for adhesion.

The film surface flatness is the key to achieve excellent bonding quality and strength (in general the surface root mean square roughness should be less than 10 nm on bonding pads). The surface roughness was examined through the Bruker Dimension FastScan Atomic Force Microscope (AFM) at least five points on each processed sample and each point spans 1 X 1 μm^2. The sample roughness of a bare Si wafer, after Nb/Ir deposition, and after titanium (Ti)/Au deposition with various thickness and deposition rate were investigated and the measured five values on each processed sample were averaged out, as shown in Figure 26. In the fabrication process of fine-pitch daisy chain samples, the Au deposition rate was finalized to 2 Å/s from original 1 Å/s. This enables easier and cleaner Au lift-off process, as a higher rate of e-beam deposition reduces the pattern step coverage and the heating duration of lift-off photoresist. This Au deposition rate of 2 Å/s leads to an increased roughness of the film (1 nm rougher) but is still good for thermal compression bonding. In Figure 26(a), it is also found that after 200 nm Au deposition on Nb/Ir, the surface roughness, instead of adding up by intuition, is 15% lower, indicating the Au deposition process can nano-planarize the Nb/Ir surface film roughness by around 0.5 nm.

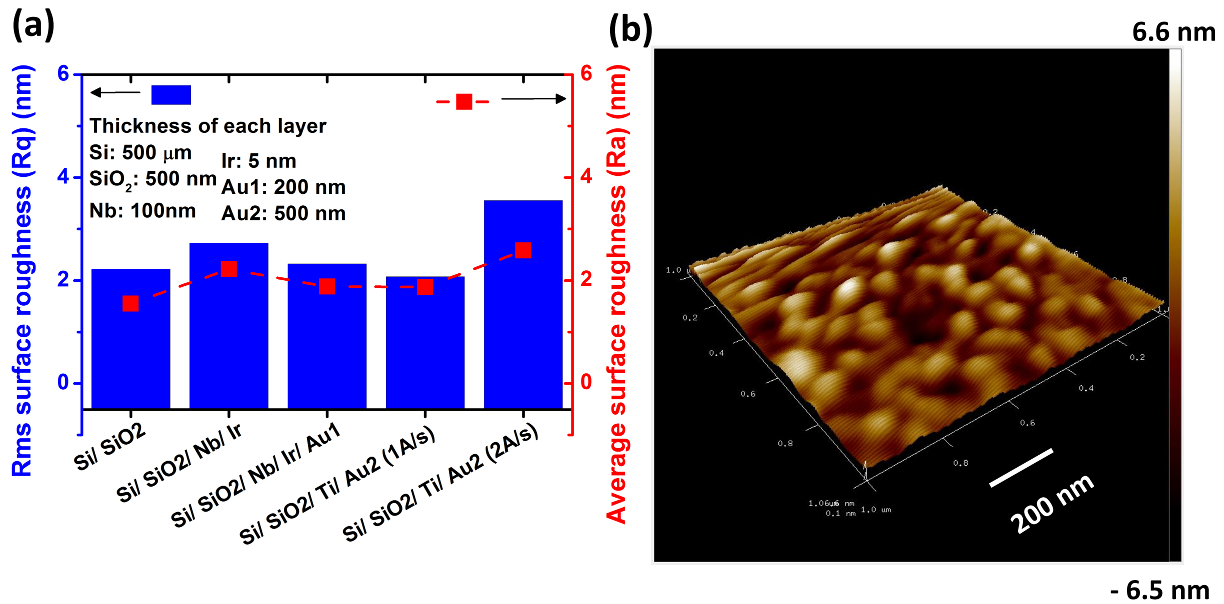


Figure 26. Surface roughness study: (a) surface roughness measurement on the samples after each film deposition, (b) Surface scanning image on a sample with 50 nm Ti/500 nm Au (a deposition rate of 2 Å/s). The grain size of Au is around 50-100 nm wide. © IOP Publishing. Reproduced with permission. All rights reserved.

5.2.2 Au-Au bonding Design of Experiment (DOE)

To optimize the bonding quality through the shear test, blanket samples were made as follows: (1) blanket wafers with 500 nm thermal oxide were cleaned using acetone, isopropanol (IPA), and DI water, (2) 100 nm Nb sputter deposition with 5 nm Ir above as a passivation layer. The Nb and Ir sputter rate were 0.57 Å /s and 0.83 Å /s, respectively, (3) 200 nm Au film was then deposited on top of Nb/Ir using e-beam evaporation with a rate of 1 Å/s, and (4) the whole wafer was diced into 2 × 2 mm^2 as dies and the other wafer is diced into 6 × 20 mm^2 as substrates for better fitting into the Dage shearing machine. Before the flip-chip bonding, both the dies and substrates were cleaned through acetone with ultrasonication, IPA, DI water rinse, and last were

activated by a 3-min 30-sccm Ar plasma with a power of 40 W, which sputtered off surface impurities and organic contaminations and enables a better Au-Au bonding integrity. The bonded dies were sheared using a Dage 4000plus bond tester, through which bonding parameters were optimized with the highest average mechanical bonding strength and the smallest standard deviation. The starting Au-Au bonding parameters began at 50% of the parameters in the reference [90] (bonding-interface temperature of 200 °C, a bonding duration of 3s, a bonding pressure of 100MPa, and a 3-min Ar clean) and added up to 75%, 100%, 150%, and 200% of that. Images related to the substrates after shear tests during DOE experiments are presented in Figure 27 and Figure 28. As expected, with higher bonding temperature and loner duration, more Au atoms diffuse to the opposite side and more bonding happens across the die area, causing more area under die to be sheared off.

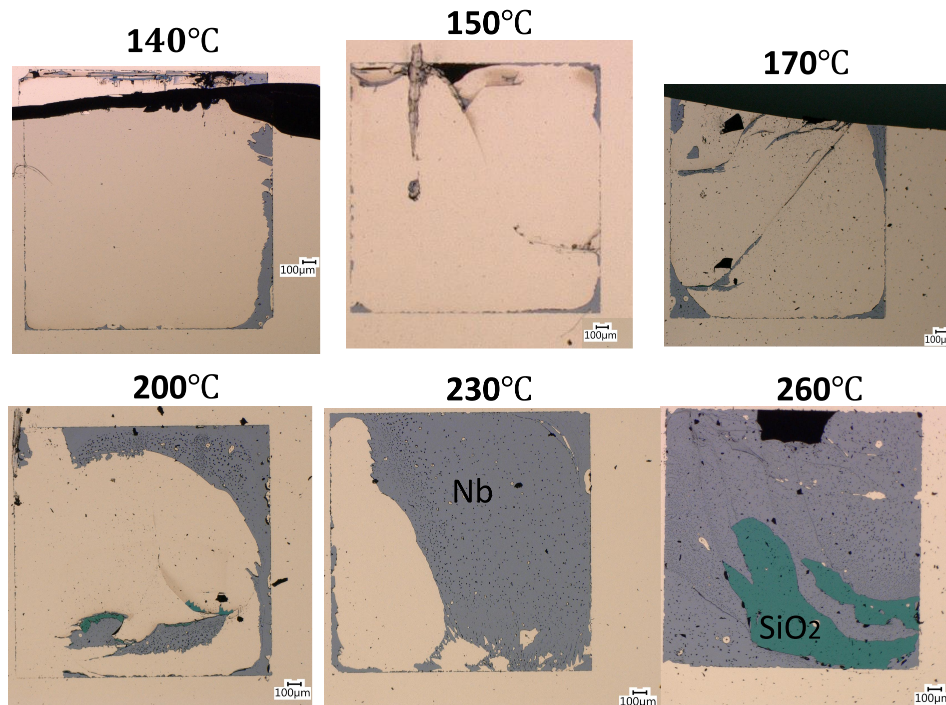


Figure 27. Images on the substate side after shear test with various bonding temperature in DOE experiments

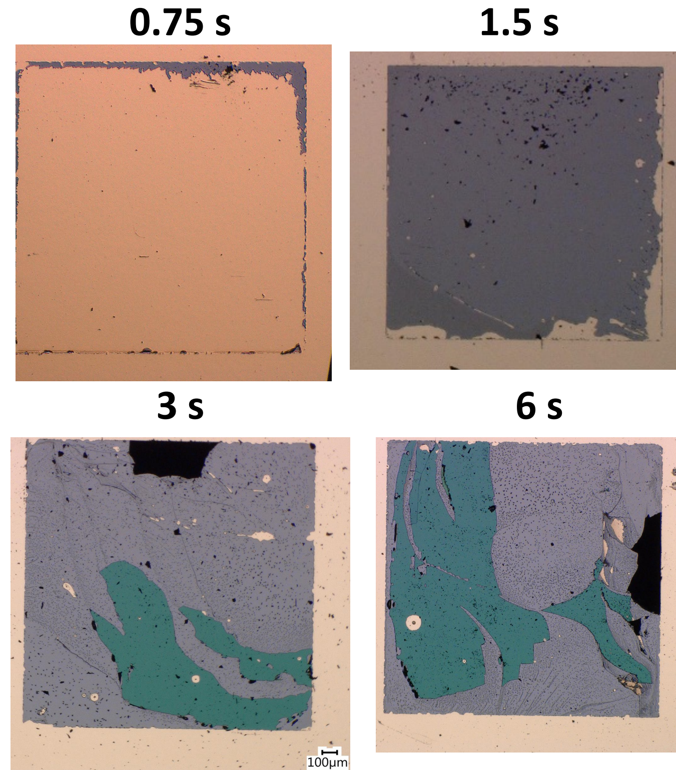
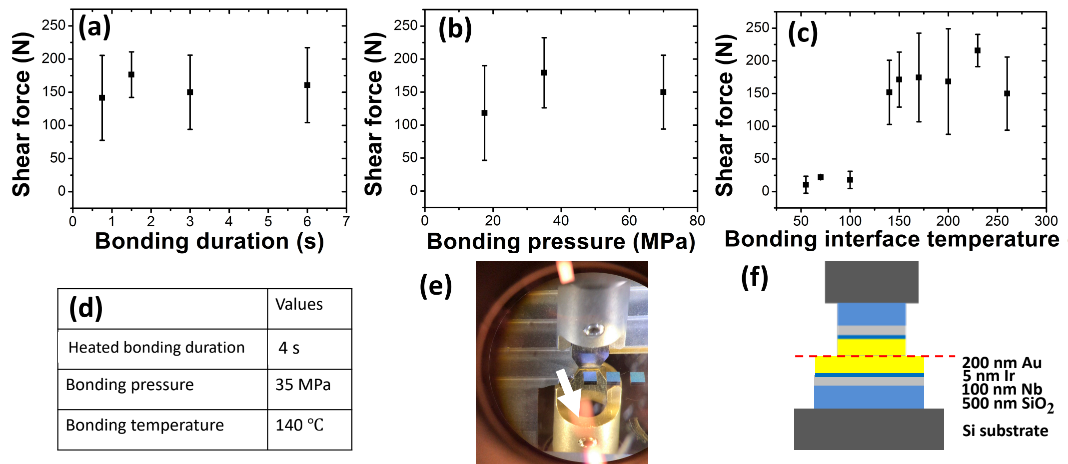


Figure 28. Images on the substate side after shear test with various bonding duration in DOE experiments

As compared to the previous lab work [90], this work focuses on low-temperature bonding applications due to the thermal budget of Josephson Junctions so the Au-Au bonding is optimized to a temperature lower than 150 °C. Figure 29 shows the DOE during bonding optimization. Annealing was implemented in first few trial runs but no significant difference regarding the shear force was observed. As such, the annealing process was ignored in later real runs. The final optimized parameters are the bonding temperature at 140 °C, the bonding duration of 1.5 s, the bonding pressure of 35 MPa, and the 3-min Ar plasma surface pre-bonding activation without annealing, as shown in the Figure 29 (d). Under all the chosen conditions, the bonding force was measured over 100 N on over 50 dies of 2 by 2 mm^2. As shown in Figure 30, the Au interlayer technology is compared with existing low-temperature integration technologies

using solders, including In [82], In/Bi/Tin [84], and In/Sn [85]. The shear strength of the Au-Au low-temperature TCB in this work is over 30 MPa, while the other three types of solder are all below 10 MPa because of brittle ceramic-like Intermetallic Compound (IMC) formation at joints. The red dash line (reading through the right Y axis) corresponds to the military standard 883G for the die shear strength in the stringent 2 X condition, indicating the shear force should exceed 50 N for a die size of 4 mm^2. The shear strength of the other three types of solder bumps is all below the threshold, while that of the Au-Au low-temperature TCB in this work can attain three times of the threshold, which is 150 N, meaning Au-Au TCB is mechanically stronger than solder bumps. In the subsection 5.2.3, since this Au-Au bonding is required to withstand low-temperature applications, the Au-Au TCB is verified to be mechanically robust through 10 runs of temperature cycling from 300 K to 4 K.



(d)	Values
Heated bonding duration	4 s
Bonding pressure	35 MPa
Bonding temperature	140 °C

Figure 29. Optimization of the Au-Au bonding conditions with respect to shear strength tests (2 x 2 mm^2 dies): (a) bonding duration (b) bonding pressure (c) bonding-interface temperature (d) optimized parameters for low-temperature Au-to-Au bonding DOE results (e) image during shear test (f) the schematic blanket structure for shear tests with a die area of 2 × 2 mm^2. © 2020 IEEE

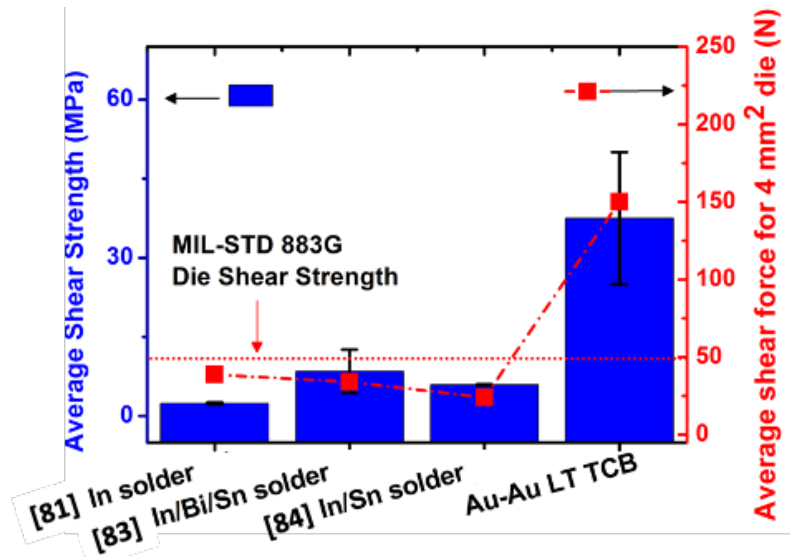


Figure 30. Mechanical strength comparison between different low temperature integration technologies and Au-Au LT TCB in this work. © IOP Publishing. Reproduced with permission. All rights reserved.

Table 7. Comparison between this work and the existing flip-chip-demonstrated superconducting assembly methods. © IOP Publishing. Reproduced with permission. All rights reserved.

Assembly methods for SCE	**Au-Au TCB (This work)**	**Indium (In) solder compression bonding** [82]	**indium/bismuth/tin (In/Bi/Sn) solder reflow** [83]	**Indium/Tin(In/Sn)** [85]	**Adhesive bonding** [86]
Joints	**Au thin-film (<500 nm)**	Solder bumps	Solder bumps	Solder bumps	Stud bumps
Temperature	140°C	Lowest to 25 °C	Lowest to 70°C	140°C	Below 180 °C
Bonding Duration	1.5 s	-	-	-	-
Bonding pressure	35 MPa	20 MPa	-	-	-
Joint height	**< 1 *μm***	12 μm	Over 20 μm	12 μm	10μm
Bonding strength	**Over 30 MPa**	1 MPa	Below 10 MPa	-	-
Special treatment	3 min Ar plasma	In-situ ion milling, Stringent plasma clean	140 °C Water soluble flux	Flux, two-step reflow	Adhesive underfill, reflow
Interconnect pitch	**≤10 *μm***	100-150 μm	75 μm	Over 100 μm	80 μm

5.2.3 Electrical tests of large-pitch Kelvin structure

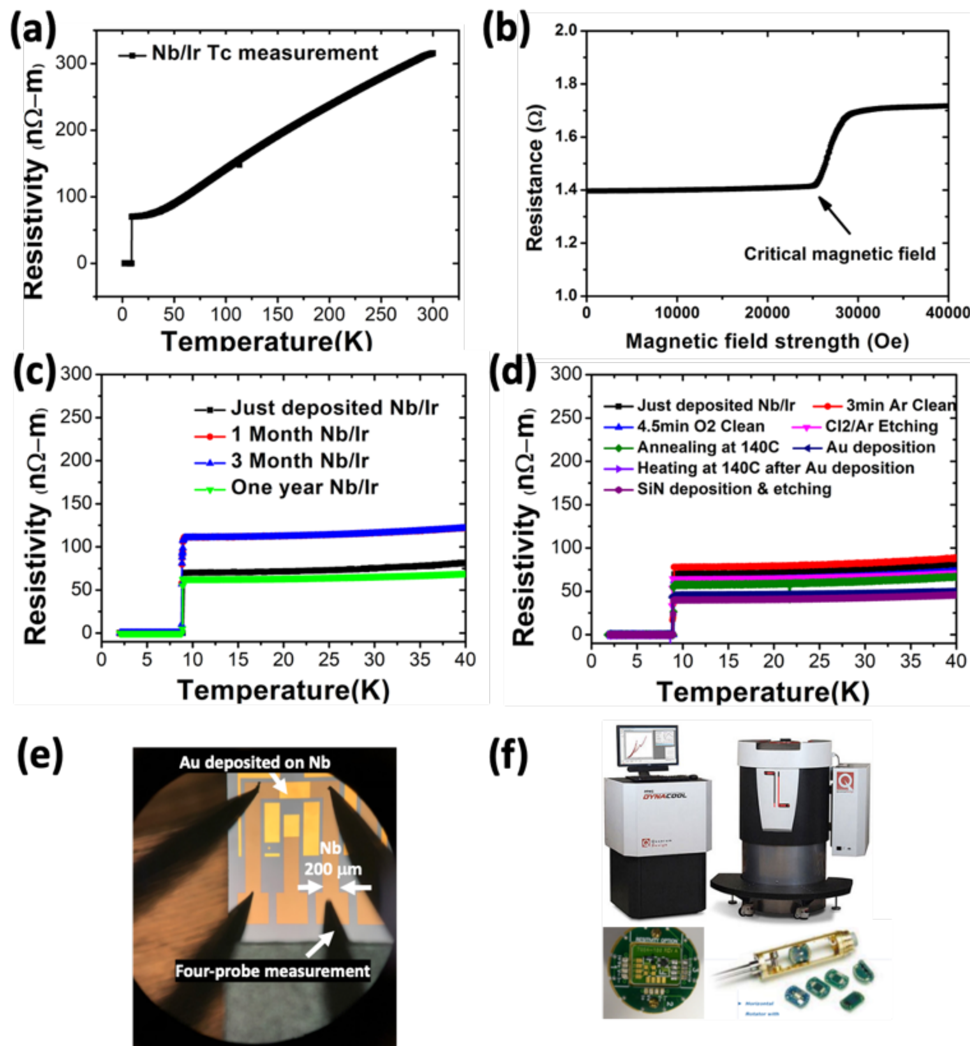


Figure 31. Low-temperature measurement results: (a) the transition temperature of Nb before any process, (b) Critical field of Nb before any process, (c) Tc check on Nb/Ir samples after one month, three month, and one year storage, (d) Nb Tc check after numerous processes using blanket samples, (e) Room-temperature-measurement image in a manual probe station under microscope before sending samples to PPMS, (f) The image of Physical Property Measurement System (PPMS) [123]. © 2020 IEEE

Before any processing, the Tc and the critical magnetic field strength of Nb/Ir was measured at 9 K and with 25000 Oersted (Oe), as shown in Figure 31(a) and (b). The Tc and the critical magnetic field strength [124] are both close to theoretical values, meaning the quality of the superconducting Nb film is excellent without significant oxidation issue after Ir capping. The capping capability of Ir on top of Nb was examined through the Tc comparison before and after one month, three month, and one year exposure in air. To confirm the stability of the Tc of Nb after each fabrication step, the Tc was measured right after each processing step, such as the Ar and Oxygen surface clean, annealing emulating multiple processes, the Cl_2/Ar dry etching, the Au deposition, SiN deposition, and the TCB. Results indicate no changes in Tc under all mentioned testing conditions, as shown in Figure 31 (c)&(d). Figure 31(f) shows the puck as the sample holder used in the PPMS. To connect samples to the puck, one end of the silver (Ag) wire was manually glued on the sample pads through Ag paste and the other end was soldered to the pads on the puck.

The resistivity variation in each measurement above the transition temperature is attributed to two reasons in the measurement stage, not related to the processing stage. First, different levels of current-crowding effect on two types of samples, which are blanket and patterned samples, lead to this inconsistency. In the measurements after Ar and O_2 treatment and annealing emulating multiple processes, around 1×1 cm^2 blanket Nb samples are utilized, while 200-um-wide Nb wires are used in the rest measurements. Second, when Ag paste, a diameter of around 300 um, and Ag wires are placed on the same type of samples, they are not aligned exactly on the same position at two edges in each sample. Although the resistivity of each measurement in numerous testing conditions varies above the transition temperature, the values of the transition temperature, which is the most important characteristic in superconducting electronics, do not

change. An important point to note here is that in Figure 31(c), the Tc of fresh Au deposition on Nb/Ir and after heating at 140 °C do not change, meaning the quality of Nb does not get degraded and the amount of Au diffusion into Nb is negligible. This experimental result validates the conclusion of the theoretical calculation discussed in the subsection 5.2.1.

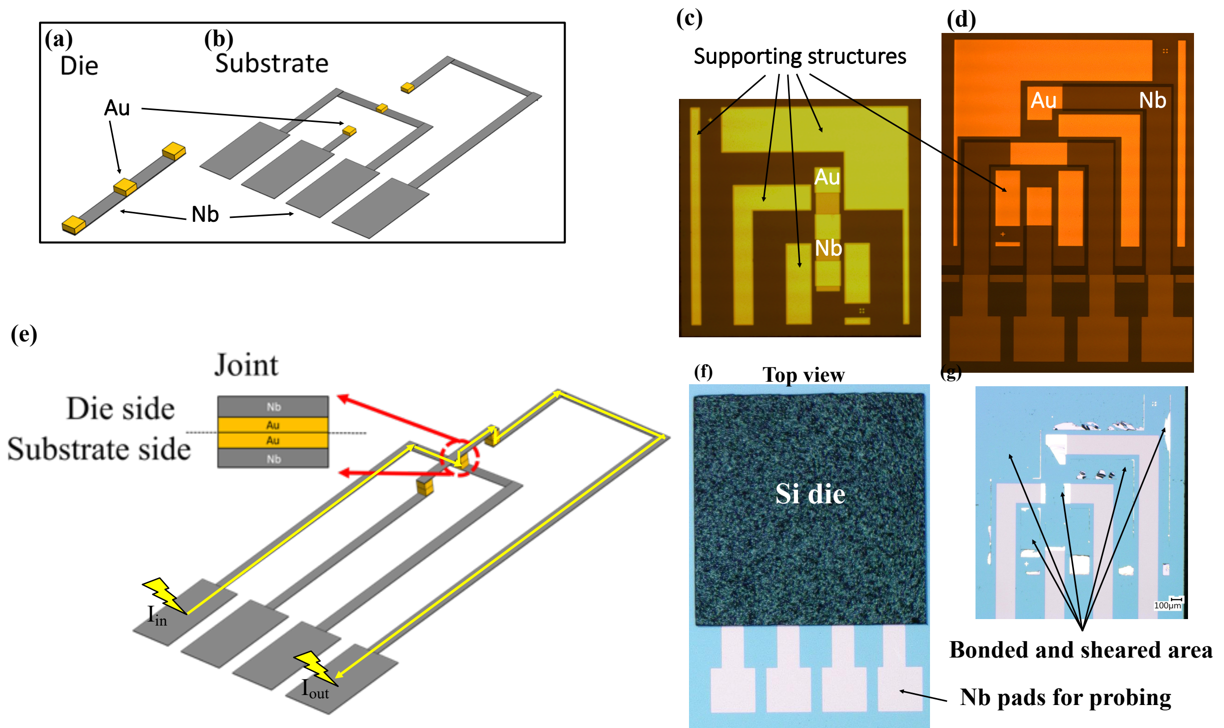


Figure 32. Details of the Kelvin structure: schematic view of the Nb/Ir interconnects and Au bonding sites on the (a) die side and (b) substrate side. Microscopic images of the structures fabricated on (c) die and (d) substrate with the interconnect, Au bonding sites, and supporting structures after dicing. (e) The schematics of the continuous electrical path after flip-chip bonding (f) Microscopic top-view image of the complete Kelvin structure with a die flip-chip bonded to a substrate with four probing pads for four-probe measurement. (g) The image after the bonded die is sheared by the test bonder. © 2020 IEEE

After ensuring that the superconductivity of Nb does not degrade during any processing step,

the next step is to fabricate the test structure and to characterize the flip-chip bonded samples electrically. The Kelvin test structure is introduced as the test structure and there are two parts in the Kelvin structure: the substrate side and the die side. In both sides, there are Nb/Ir interconnects with a width of 200 μm, Au bonding sites, and dummy structures for mechanical supporting, as shown in Figure 32 (a)&(b). The substrate site also includes fan-out wiring and probing pads for the standard four-probe measurement. After fabrication and flip-chip assembly, the whole Kelvin structure was then electrically connected and was adhered to the puck. The sample with puck was cooled down to 4 Kelvin for superconducting measurement using Physical Property Measurement System (PPMS).

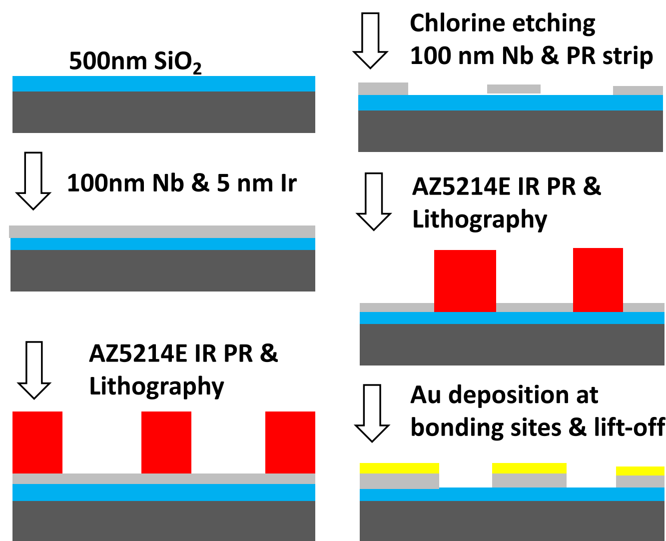


Figure 33. The fabrication process of the Kelvin structure in both the die and the substrate sides for electrical characterization: 500 nm thermal oxide is first grown, followed by Nb/Ir deposition in an ultra-high vacuum system without breaking the vacuum. Patterns are defined through lithography using AZ 5214 photoresist. The whole wafers are etched using Cl dry etching. Au thin film is deposited and patterned at the bonding sites through the AZ 5214 EIR lift-off process. © 2020 IEEE

The fabrication process of Kelvin structure, as shown in Figure 33, starts with a 100 nm Nb and a 5 nm Ir sputtering on a four-inch cleaned Si wafer. Next, patterns are defined through lithography using positive tone photoresist (PR) AZ 5214 with a thickness of 2 μm (the rotation speed of major spin coating is 2000 rpm). The PR is hard-baked at 120℃ for two minutes for the later dry eting. Chlorine (Cl) dry etching is then applied to the samples with a 10-sccm Ar and a 30-sccm Cl_2 gas flow and a Reactive Ion Etching (RIE) power of 100 watts. After dry etching, the AZ 5214 PR is stripped off and the next AZ 5214 EIR negative tone process is utilized for the following Au lift-off process at bonding sites. The die and the substrate part are diced into an area of 2 X 2 mm^2 and 10 X 10 mm^2, as shown in Figure 32 (c)&(d). At last, the dies and the substrates are cleaned through acetone with ultrasonication, IPA, and DI water, followed by exposure to Ar plasma with a 30 sccm Ar gas flow and flip-chip bonding with the optimized bonding parameters. The flip-chip bonded Kelvin structure is presented in Figure 32 (e)&(f). Over 10 bonded samples were measured. Shear test is done with expected force with the sheared image shown in Figure 32 (g).

Figure 32 (e) shows a conceptually current flow in a bonded Kelvin structure. The current begins at the current probe named I_{in}, flows in the substrate side, moves to the die side through the bonded Au-Au contact, flows back to the substrate side through another Au contact, and eventually ends in the other current probe named I_{out}. The two voltage probes can be placed separately around the two current probes to measure the voltage drop of the whole structure, through which the overall resistance can be measured. The Kelvin structure can measure the contact resistance of Au-Au contact by placing the two voltage probes in the inner two pads instead of the outer two pads.

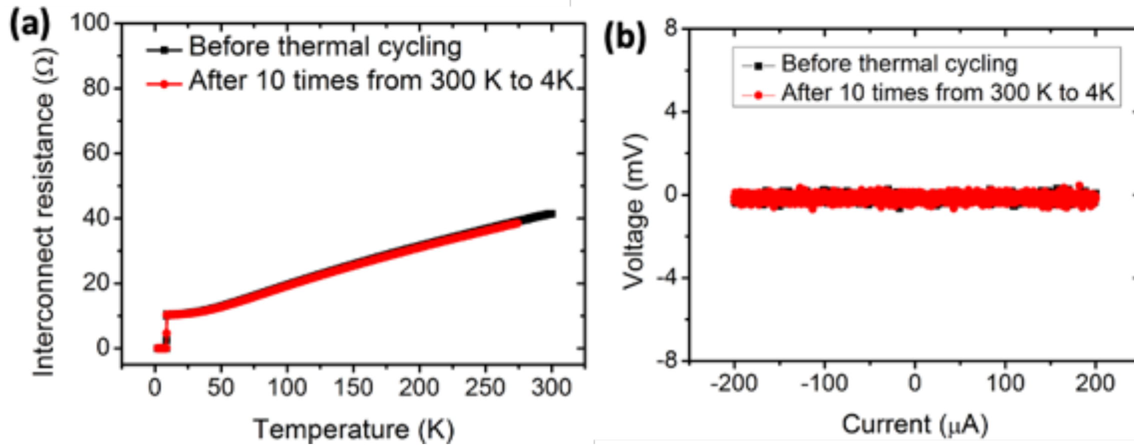


Figure 34. (a) Resistance-temperature measurement of the Kelvin structure (200-μm-width Nb interconnects with Au joints) before and after 10 times of temperature cycling between 300 K and 2 K. The Tc of the whole structure is at 9K and the resistance below Tc is around 0.02 Ω. © IOP Publishing. Reproduced with permission. All rights reserved, (b) Voltage-Current measurements of the Kelvin structure before and after temperature cycling between 300 K and 2 K.

Figure 34 shows the resistivity-temperature (R-T curve) of Nb interconnects with Au joints before and after 10 times of temperature cycling between 300 K and 2 K. The measured Nb length is around 4.2 mm in this structure with a width of 200 μm and a resistivity of 215 nΩ -m, which was measured in the preliminary test at 300 K, and the resultant theoretical resistance is 45.1Ω. The measured resistance at 300 K is 46 Ω, which agrees with the calculation. The transition temperature is at 9 K and the resistivity shows a smooth curve above Tc, drops abruptly at Tc with a transition width of 0.2 K, and remains 0.02 Ω at 2 K, which comes from the resistance of Au and Ir since both of them are not superconductors at this temperature. The Au-Au contact resistance was measured to be 0.002 Ω at 2 K with a contact area of 200 × 200 μm^2 and the resultant specific contact resistance at 2 K was calculated to be $8 \times 10^{-8}\ \Omega - \mathrm{cm}^2$.

Reliability test using temperature cycling (repeating cooling and warming processes between 300 K and 2 K) was performed with a cooling and warming rate of $5°C/s$. After 10 temperature cyles, the Tc and overall interconnect resistance remain the same, indicating Au joints are connected even after 10 times of cylcing. This also reveals that the strength of the low-temperature Au-Au TCB is robust enough to withstand the structural stress originating from the coefficient of thermal expansion (CTE) mismatch between each material during the thermal contraction and thermal expansion. The values for each each material are as follows: Au (14 ppm/K), Nb (7 ppm/K), Ir (6 ppm/K), and Si (3 ppm/K).

On the die side, besides the Au-island structure, as shown in Figure 35(b), another structure called the continuous-Au-film structure, as shown in Figure 35(c), was also made. The latter one is the standard Si-IF process and allows more alignment margin compared to the Au island structure for the fine-pitch daisy chain application in the subsection 5.2.4. Figure 36 shows the interconnect resistance of the two die structures versus temperature. The R-T curve shows that the Tc and the resistance below Tc are both not affected and the resistance of the continuous Au structure is less than the resistance of the Au island structure above Tc, which is an expected result of a basic current flow in a parallel electrcial path. Therefore, the electrical characterization of dies with a continuous Au film is the same as that of dies with a Au island film. Dies with the continuous Au film are applied to the fine pitch daisy chain structure, enabling a larger margin of alignment.

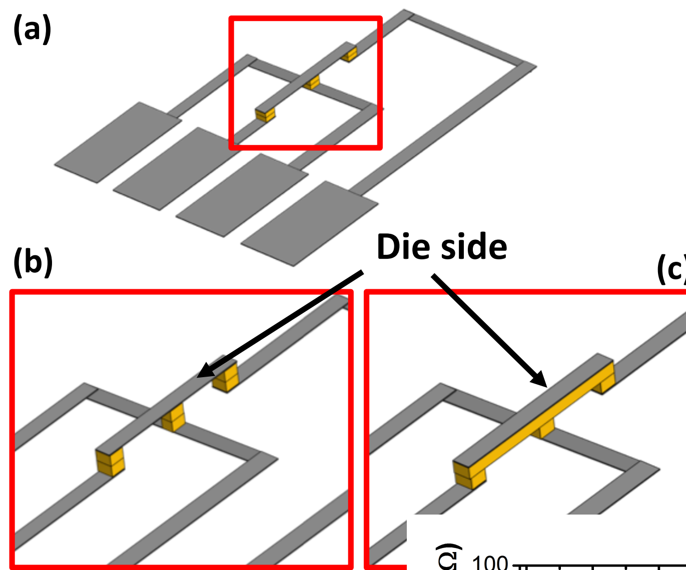
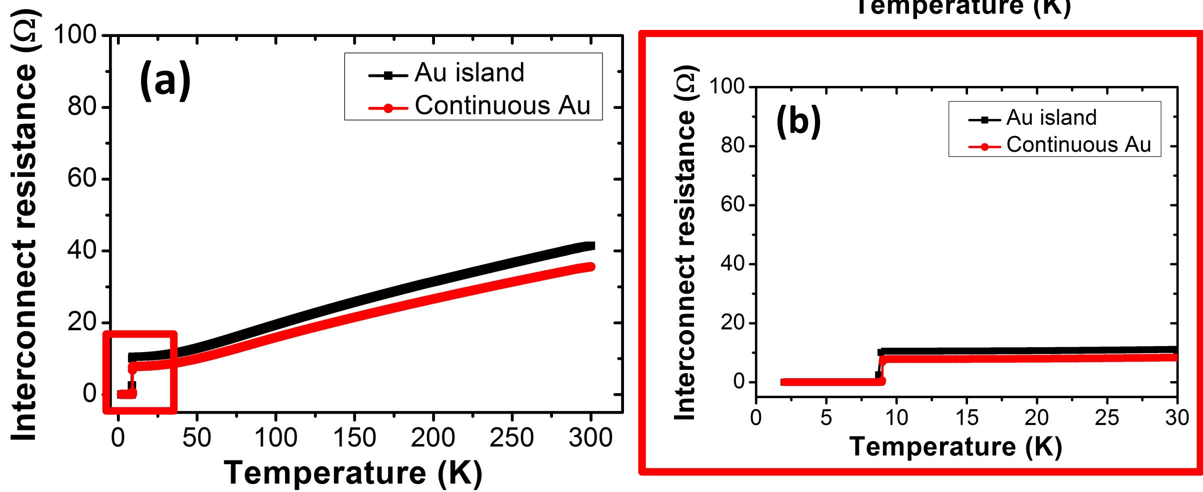


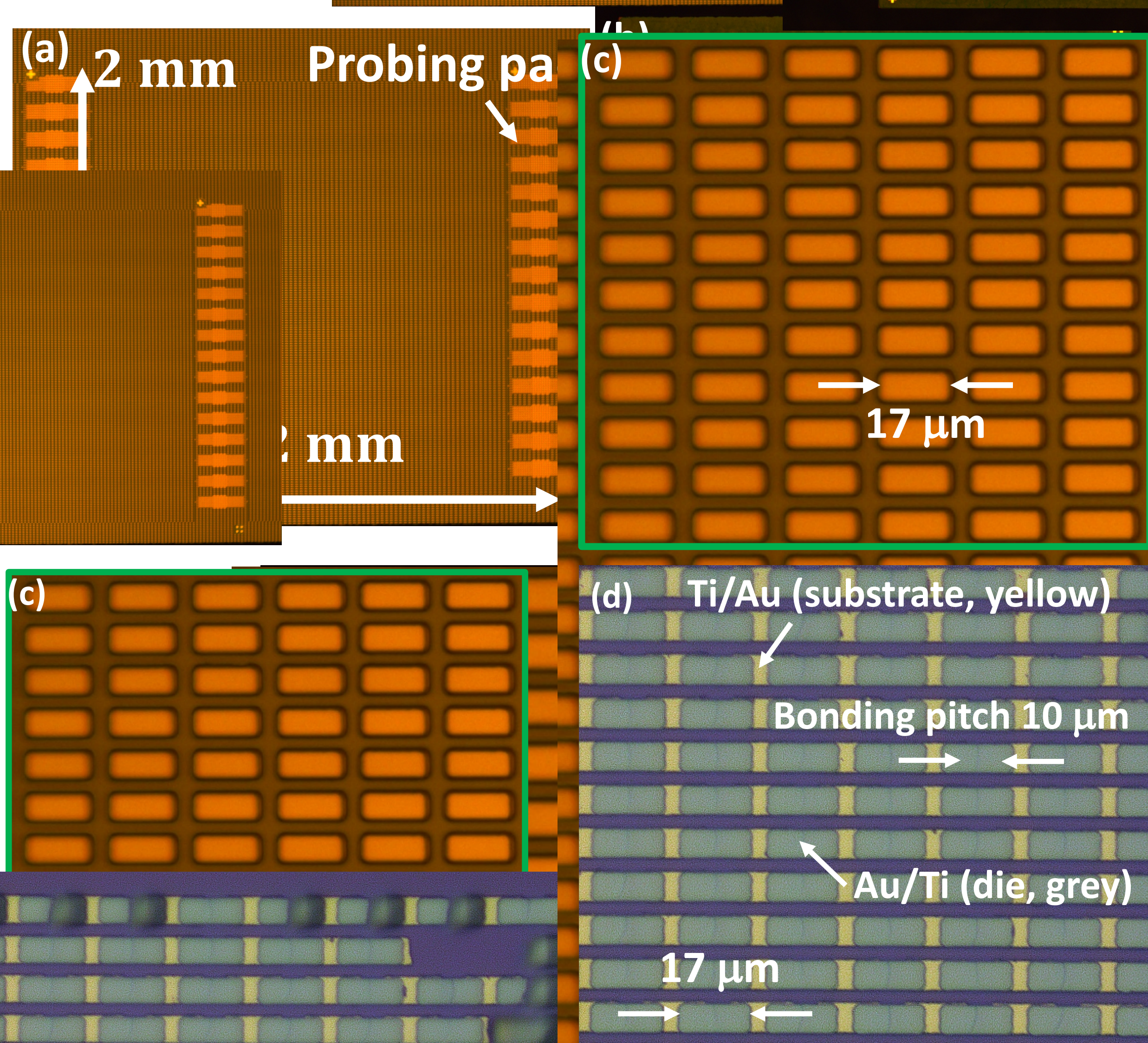
Figure 35. (a) Schematic view of flip-chip bonded
bonded structure with an Au island structure in the
structure with a continuous Au film structure in the d
permission. All rights reserved.



F Au island structure and the continuous
A (a). More detail information is in the
t rights reserved.

5.2.4 Electrical tests of fine-pitch daisy chain structure

After using the Kelvin structure for the development of the fabrication process and for characterization of the superconductivity of flip-chip bonded samples, the next step is to make fine-pitch daisy chain structure using the same process. Since the high-quality Nb/Ir wafers are precious, the fine-pitch low-temperature Au-Au TCB was tested first before applying on Nb interconnects. For small patterns, a thicker Au film was experimentally verified to provide a stronger bonding strength, overcoming the die warpage issue and achieving a reliable Au-Au bonding across the whole die area. Au film thickness of 500 nm were deposited and patterned through lift-off process to form pads of 17×7 μm^2 on the both die and substrate sides having a I/O pitch of 10 μm, as shown in Figure 37 (a)-(c). For the die-to-substrate bonding, dies were taught to align to substrates with an overlay accuracy of $\pm$ 1 μm. The alignment was determined by shearing off the flip-chip bonded dies and observing the imprint on the mating pads across the bonding area under a microscope, as shown in Figure 37(d). The yellow area is pads on Superconducting-IF. The grey Ti pads originally are adhered to the die side but after shearing off, they adhere to the pads in the substrate side since the Au-Au bonding strength is stronger than the adhesion strength between pads and die silicon.


(a)
2 mm
Probing pa
2 mm
(c)
17 µm
(d)
Ti/Au (substrate, yellow)
Bonding pitch 10 µm
Au/Ti (die, grey)
17 µm
(c)
50 µm

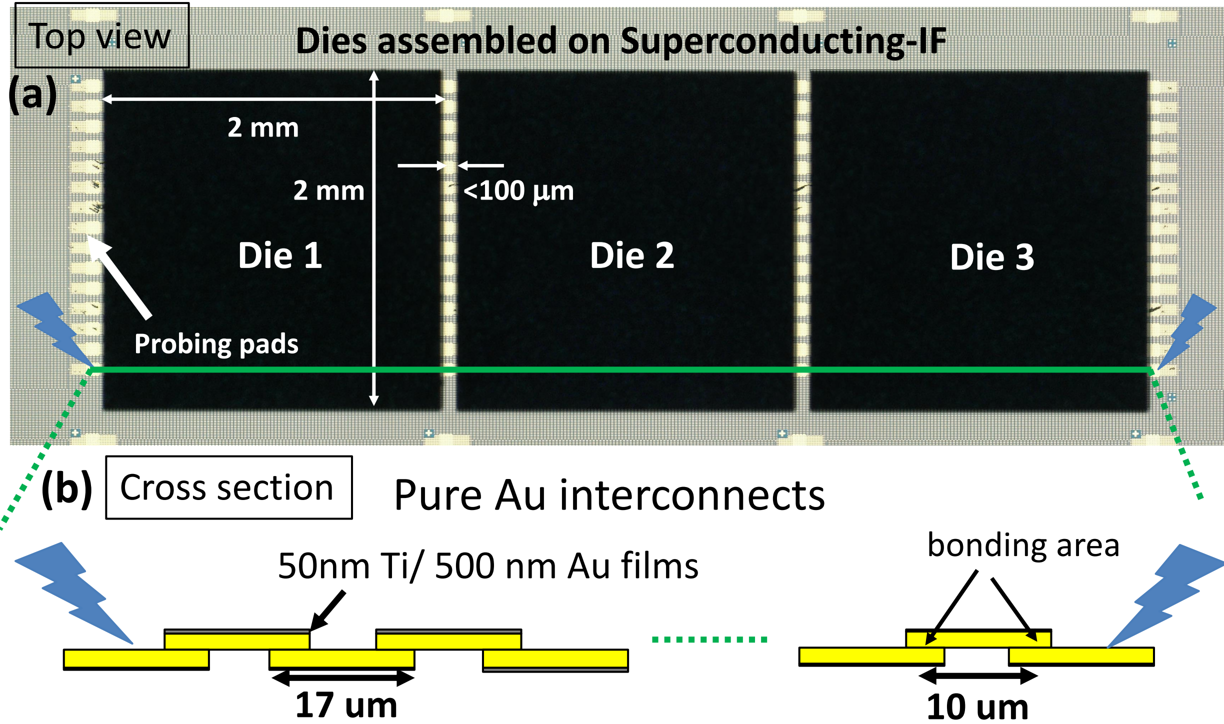


Figure 38. (a) Image of three dies assembled on Superconducting-IF and measured in a row. (b) Schematic view of the cross-section view. The pad size is 17 μm and the bonding pitch is 10 μm. © IOP Publishing. Reproduced with permission. All rights reserved

After flip-chip TCB process, a continuous daisy chain structure is completed and there are three dies connected in a row, as shown in Figure 38(a). In the substrate side, there are probing pads, which can test multiple dies at a time. Figure 38(b) shows the cross-section scheme of the measured path. There are 15 rows of probing pads and in each row there is a number of 570 Au-Au bonded joints, making overall 8550 joints. For each die, the length of the 50 nm Ti/ 500 nm Au electrical path is 2.5 mm and the expected resistance is ~ 17 Ω (the ideal resistivity of Au and Ti are 24.4 and 420 nΩ-m), making the overall resistance of three dies ~ 50 Ω. Over 10 bonded samples are measured and all of them show the same curve. Figure 39(a) shows one of the R-T curves of the Ti/Au fine-pitch (≤ 10 μm) daisy chain structure with a resistance value at 300 K

approaching the theoretical value. When the temperature drops, the resistance shows a smooth curve down to 2 K and the value at 2 K is around 10 Ω, which is 5 times less compared to the resistance at 300K.

A reliability test was implemented, as shown in Figure 39 (c), using temperature cycling (repeating cooling and warming processes between 300 K and 2 K) with a cooling and warming rate of 5 °C/s. As there is no existing standard to know how many runs of temperature cycling are sufficient, it is reported by Das *et al.* [1] that at least 10 runs of cyling are used. After 5 times and 10 times of the temperature cycling, the overall R-T curve remained the same down to 2 K, representing Au joints were connected after 10 times of cycling. This shows that even though in each bonding pad the contacting area (17 μm X 7 μm) is much smaller than the contacting area (200 μm X 200 μm) in the Kelvin structure, the mechanical strength of the low-temperature Au-Au TCB is robust enough to withstand the structural stress during the thermal contraction and thermal expansion. The fine-pitch (≤ 10 μm) demonstration is presented in the SEM cross section image, as shown in Figure 40.

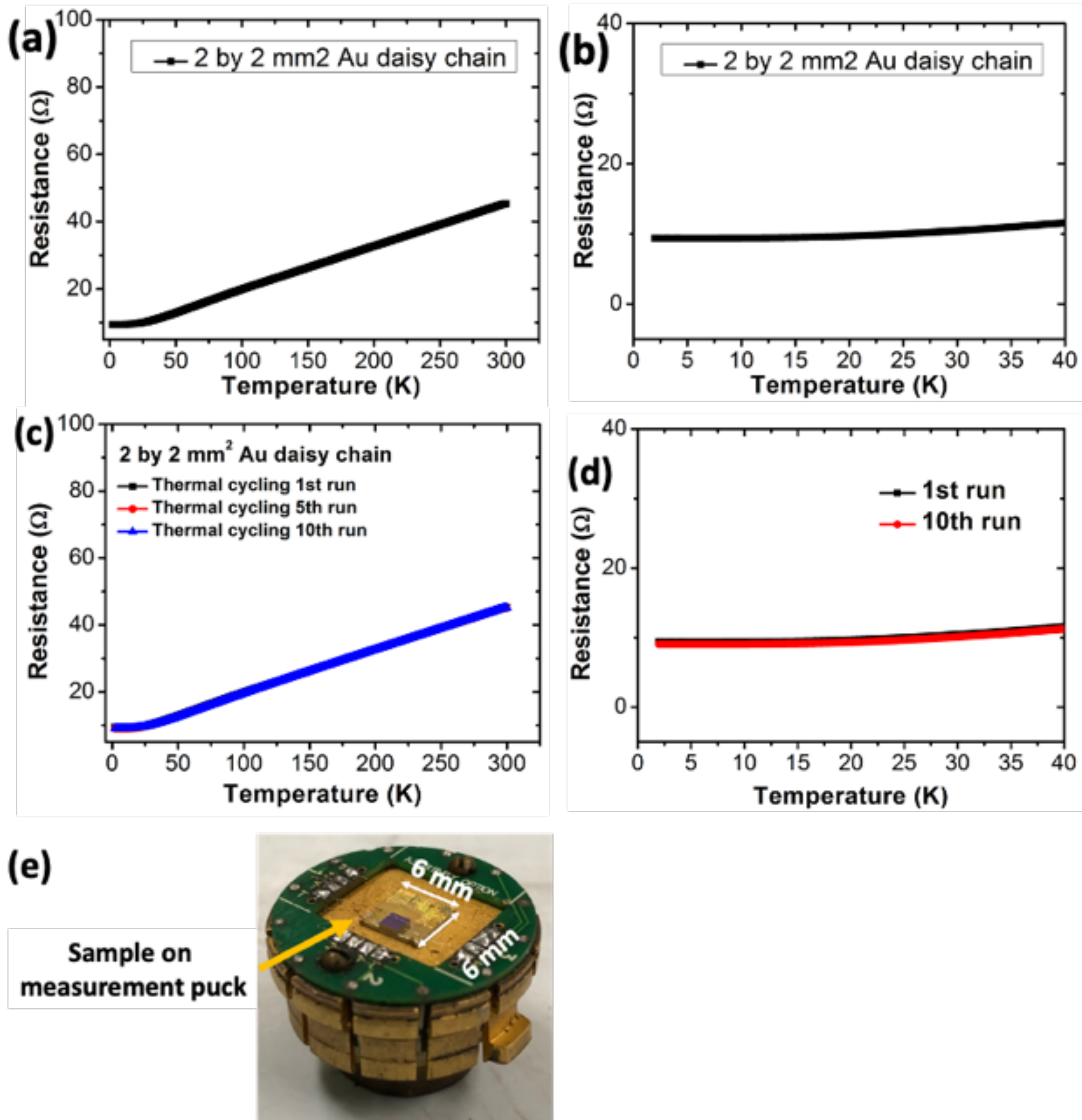


Figure 39. (a) R-T measurement of Ti/Au interconnects and joints at low temperature. (b) zoom-in of (a) between 2 K to 40 K. (c) R-T curve after 10 times of temperature cycling between 300 K and 2 K. (d) zoom-in of (c) between 2 K to 40 K. (e) samples on measurement a puck. © IOP Publishing. Reproduced with permission. All rights reserved.

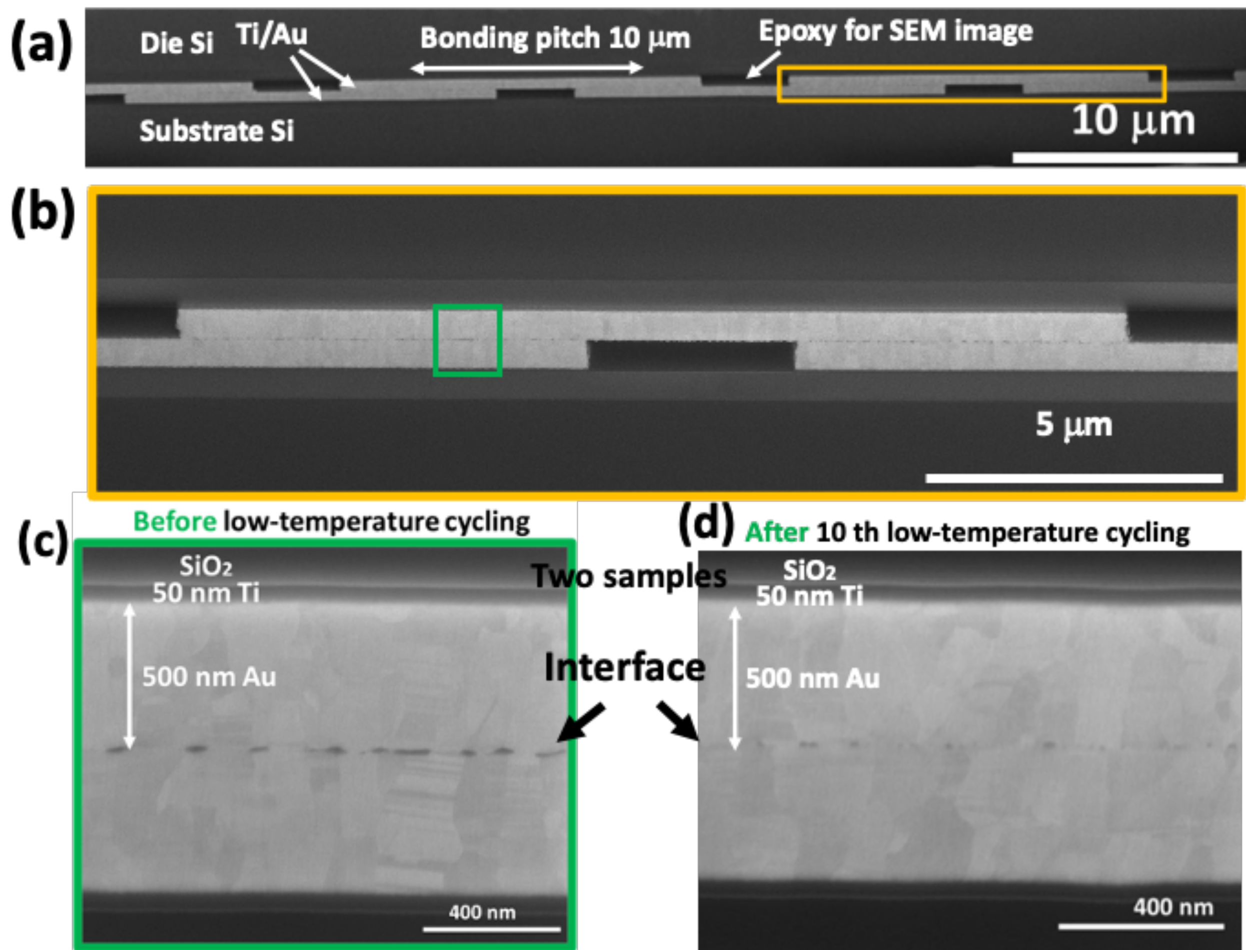


Figure 40. (a) Cross section image on the Au-to-Au bonding sample with fine-pitch demonstration. (b) Zoom-in of (a). © 2020 IEEE. (c) Cross-section image on the bonding interface before temperature cycling. © 2020 IEEE. (d) Cross-section image on the bonding interface after temperature cycling on another sample.

The next step is to electrically characterize the Au interlayer on different bonding areas. The fabrication process is the same as previous samples. The difference is in the design, including length of the daisy chains, the number of joints, and the bonding pressure. For dies with bonding area of 2×2 mm^2, 3×3 mm^2, 4×4 mm^2, and 5×5 mm^2, the bonding pressure were calculated through 300 N bonding force divided by the contacting area and the values are 714 MPa,

250MPa, 126 MPa, and 80 MPa, respectively. Following the lift-off process, various sizes of dies and substrates were fabricated and presented in Figure 41 & Figure 42. After flip-chip bonding, the assembly process was completed. Figure 43 shows the images for aligned samples and alignment check.

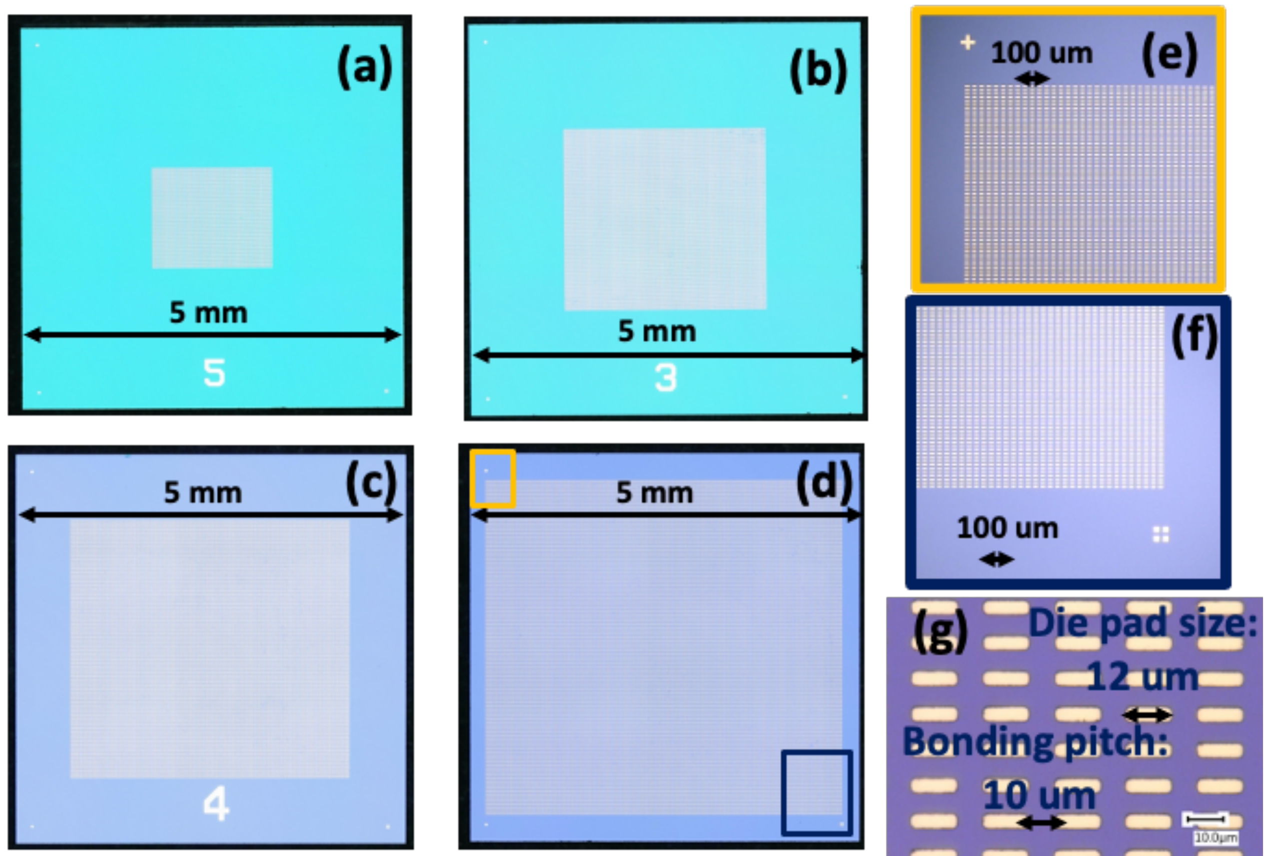


Figure 41. Image of different bonding areas of Au daisy chain samples on the die side with a die size of 5×5 mm^2: (a) 2×2 mm^2 (b) 3×3 mm^2 (c) 4 × 4 mm^2 (d) 5 × 5 mm^2 (e)&(f) zoom- in image of alignment keys on (d). (g) zoom in image of bonding pads. The die pad size is 12 × 5 μm^2 and the bonding pitch is 10 µm.

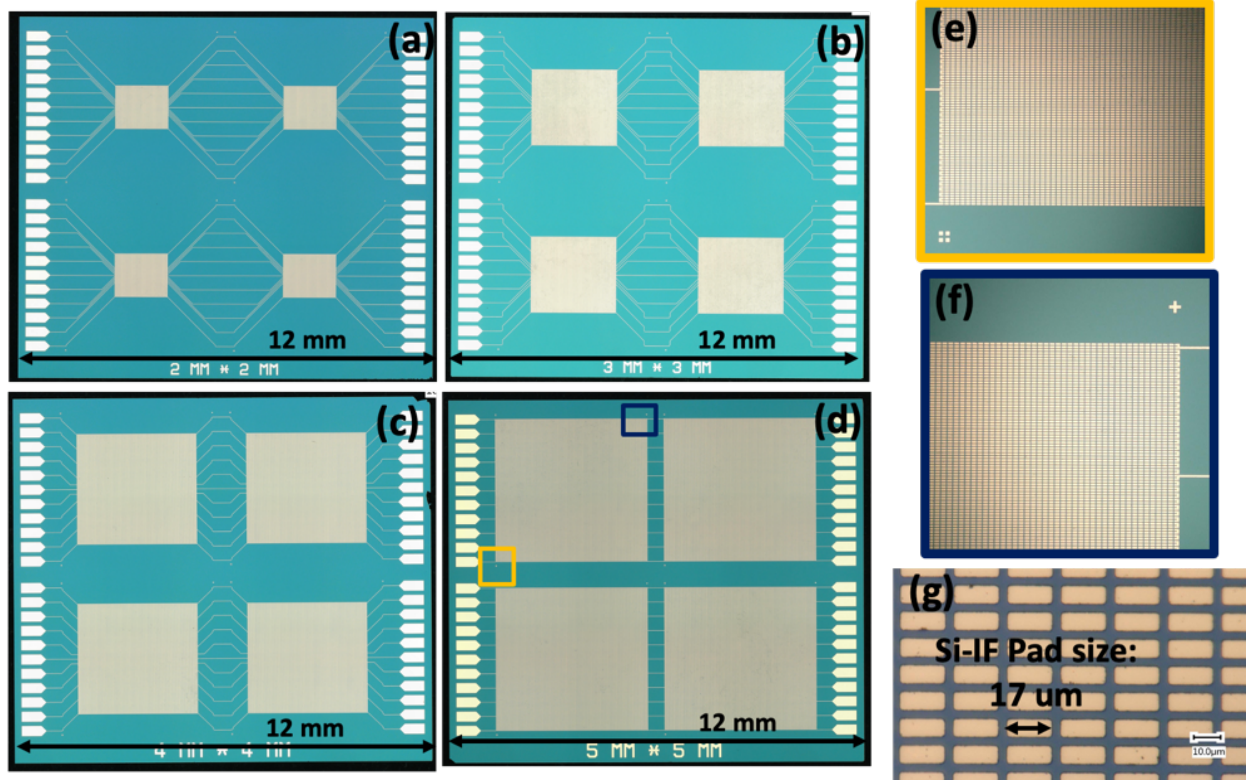


Figure 42. Image of different bonding areas of Au daisy chain samples on the substrate side with a die size of 5×5 mm^2: (a) 2×2 mm^2 (b) 3×3 mm^2 (c) 4 × 4 mm^2 (d) 5 × 5 mm^2 (e)&(f) zoom- in image of alignment keys on (d). (g) zoom in image of bonding pads. The pad size is 17 × 7 μm^2 and the bonding pitch is 10 µm.

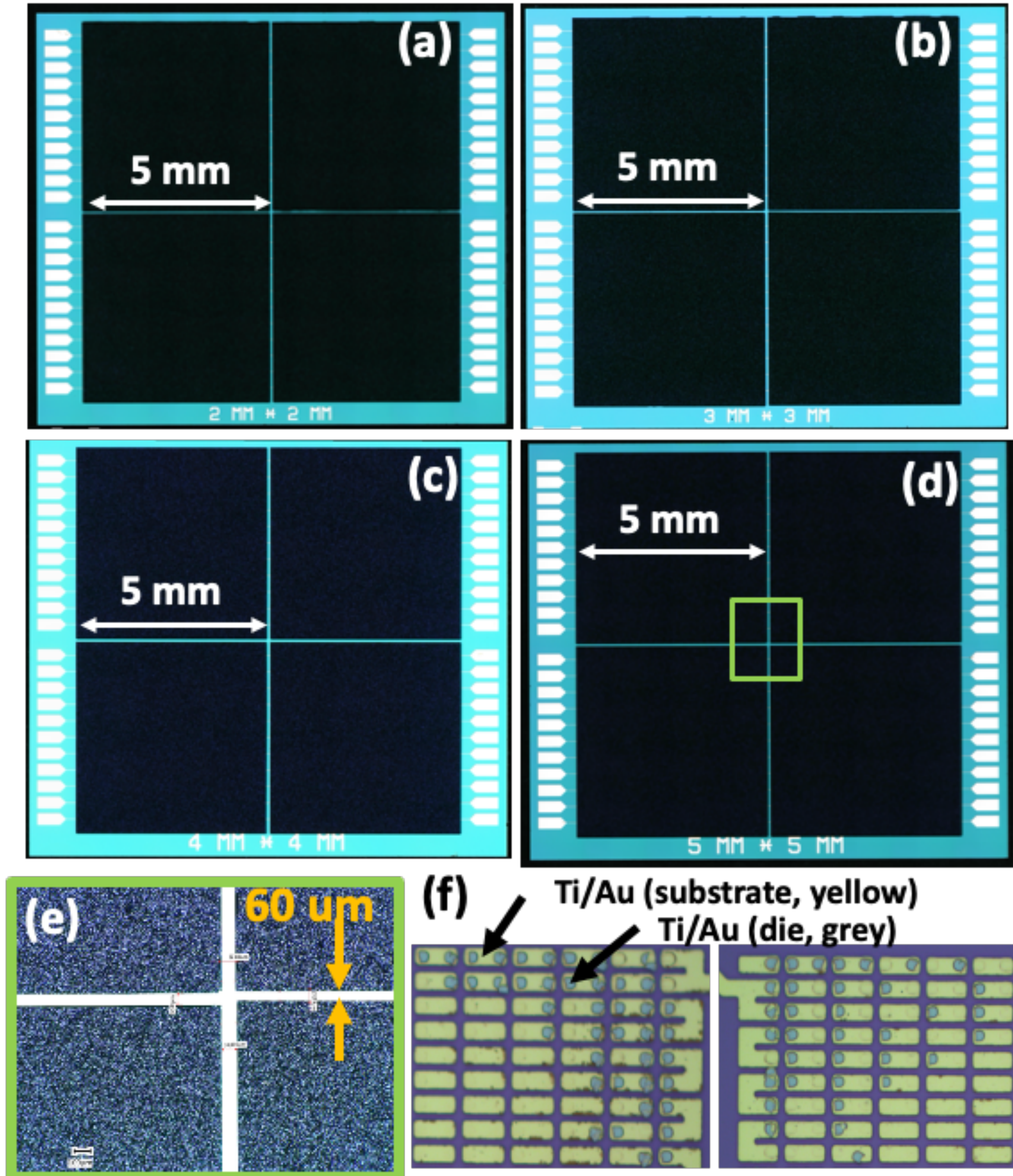


Figure 43. Image of different bonding areas of Au daisy chain samples after flip-chip bonding with a die size of 5×5 mm^2: (a) 2×2 mm^2 (b) 3×3 mm^2 (c) 4 × 4 mm^2 (d) 5 × 5 mm^2 (e) zoom- in image of alignment keys on (d). (f) Image for a good alignment. © 2020 IEEE

In the dies with the largest around 5×5 mm^2 bonding area, there are two hundred thousand joints with an overall Au contacting area 3.8 mm^2 (the bonding pressure is around 80MPa).

Table 8 shows detailed parameters for dies with bonding area of 2 × 2 mm^2, 3 × 3 mm^2, 4 × 4 mm^2, and 5 × 5 mm^2. For each die with around 5×5 mm^2 bonding area, the daisy chain length of 50 nm Ti/ 500 nm Au is around 3000 mm and the corresponding ideal resistance (the theoretical resistivity of Ti and Au are 420 and 24.4 nΩ-m) is 20 kΩ. Since two dies are designed to be connected in shunt, the room temperature resistance should be around 10 kΩ. For each dies with bonding area of 2 × 2 mm^2, 3 × 3 mm^2, 4 × 4 mm^2, and 5 × 5 mm^2, the expected resistance is 1.1, 3, and 6 kΩ. The corresponding measured value of resistance for dies with bonding area of 2 × 2 mm^2, 3 × 3 mm^2, 4 × 4 mm^2, and 5 × 5 mm^2 at 300K (1, 4 , 6, and 11 kΩ) and down to 2 K (0.1, 0.7 , 0.9, and 2 kΩ), as shown in Figure 44. The room temperature resistance is close to their calculated values. Temperature cycling test (300 K to 2 K) was implemented on overall 16 various die sizes, as shown in Figure 45, and all the cycling results showed the same trend (smoothly decreased resistance without electrically open in the repeated cooling down and heating up phases). This means the Au interlayer is reliable and uniform across various die sizes.

Table 8 Parameters for dies with bonding area of 2 × 2 mm^2, 3 × 3 mm^2, 4 × 4 mm^2, and 5 × 5 mm^2

Die size (mm ×mm)	Bonding size (mm ×mm)	No. of joints	Contacting area (mm^2)	Bonding pressure (300N/contacting area, MPa)	Au thickness (nm)	Total one-die length (mm)	Ideal one-die resistance (kΩ)
5×5	2 × 2	21k	0.42	714	500	314	2.2
5×5	3 × 3	60k	1.20	250	500	901	6
5×5	4 × 4	120k	2.38	126	500	1787	12
5×5	5 × 5	200k	3.80	80	500	3000	20

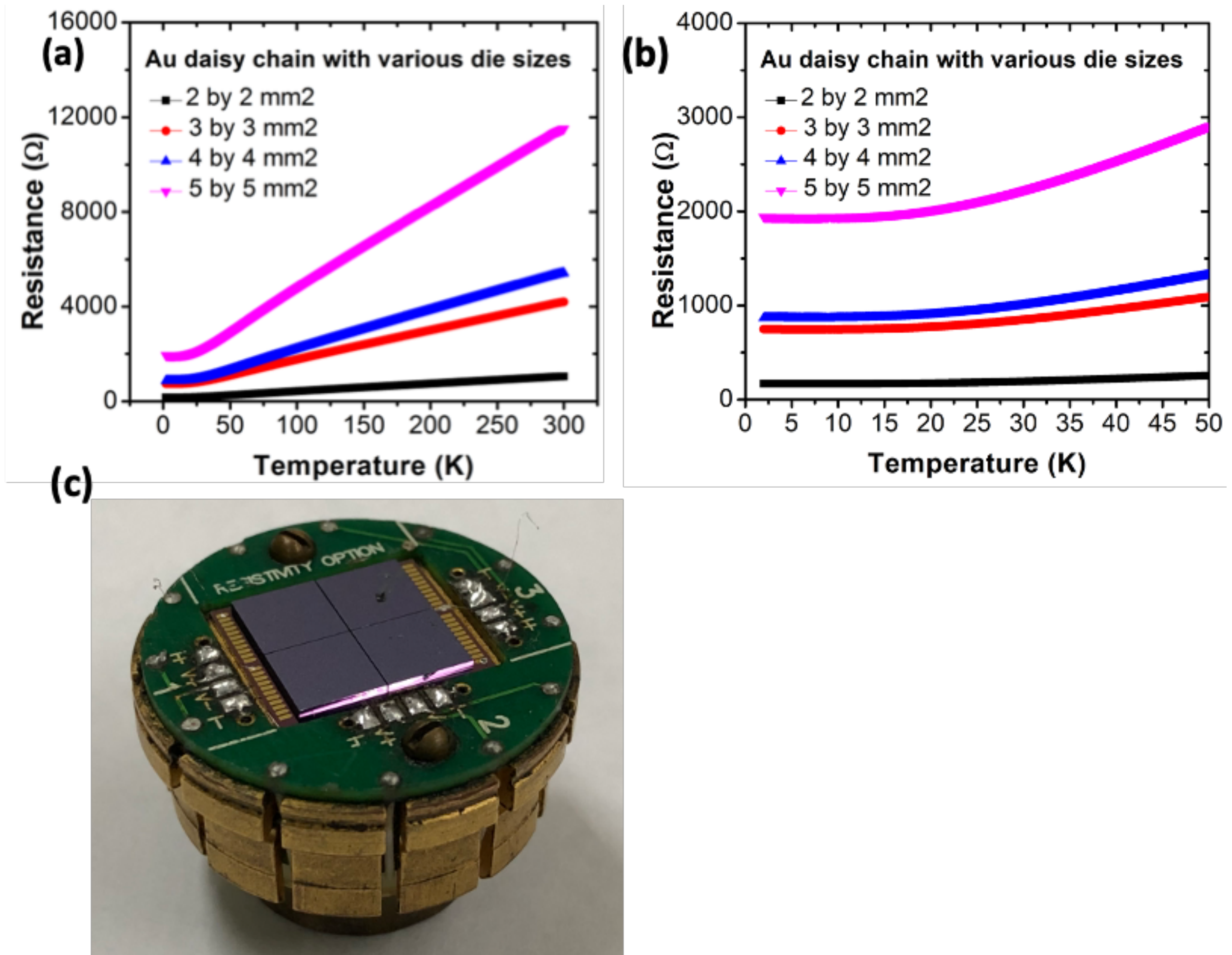


Figure 44. (a) Resistance-temperature electrical measurement on different die bonding area (2×2 mm^2, 3×3 mm^2, 4×4 mm^2, and 5×5 mm^2) down to 2 K, (b) Zoom-in (a) between 2 K to 50 K. (c) Bonded samples on a measurement puck. © 2021 IEEE.

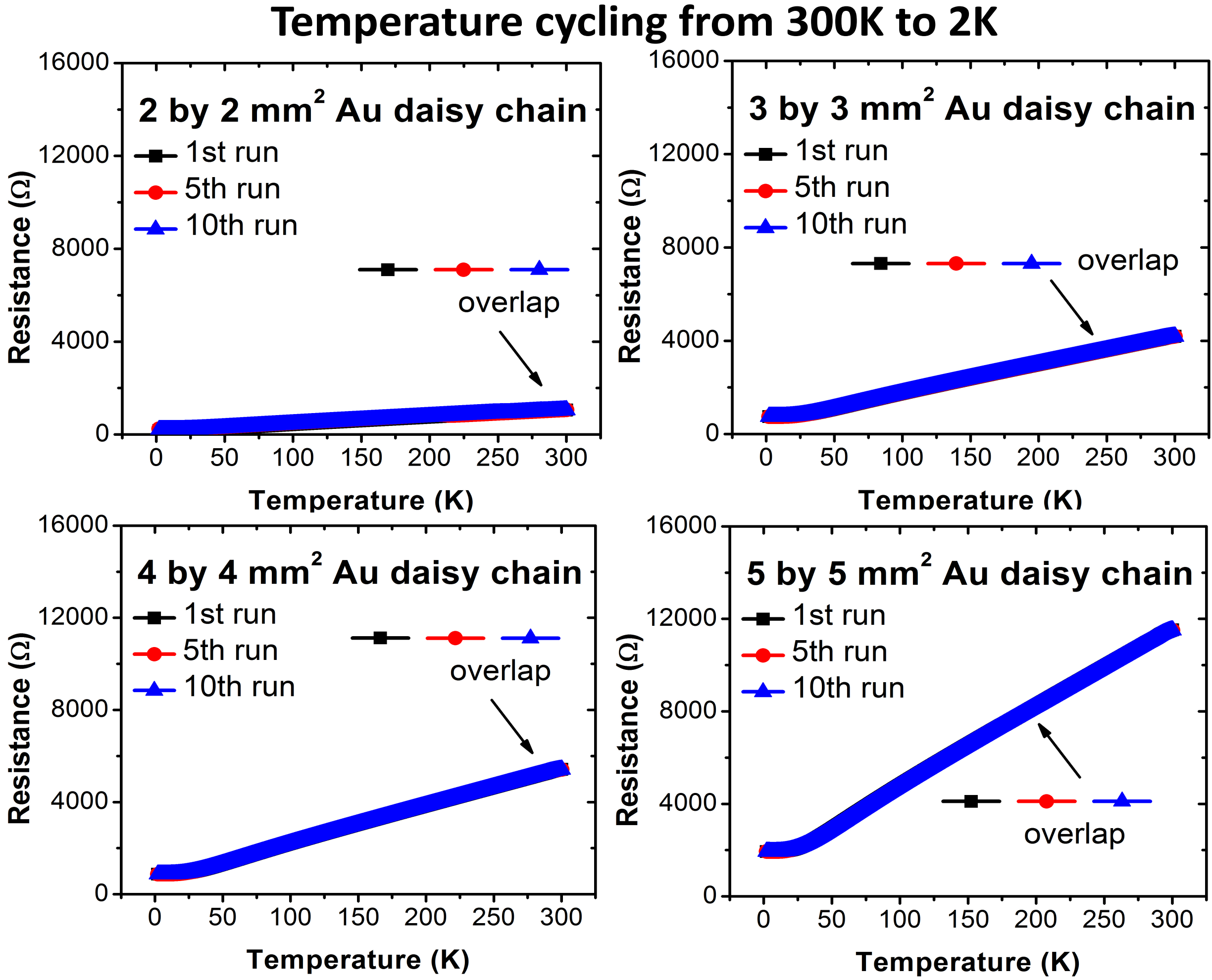


Figure 45. Electrical measurements of the Au interlayer (temperature cycling from 300 K to 2 K on various sizes of dies: 2×2 mm^2, 3×3 mm^2, 4×4 mm^2, 5×5 mm^2.

After confirming the Au interlayer bonding is reliable and uniform on different die sizes, the next step is to apply the Au interlayer on Nb/Ir superconducting interconnects. The fabrication process was the same as that for the Kelvin structure and is stated as follows: (1) 100 nm Nb/5 nm Ir sputter deposition without breaking the chamber on oxidized wafers, (2) patterning 100 nm Nb/5 nm Ir through a subtractive Cl_2/Ar dry etching, (3) Au interlayer deposition and patterning through the previously mentioned Au lift-off process, (4) 2 × 2 mm^2 dies and 10 × 10 mm^2 dies substrates singulation, (5) ultrasonicating cleaning and Ar plasms surface activation, and (6) flip-

chip bonding over 10 surrogate superconducting dies on the superconducting-IF. The bonded samples were cooled to 2 K to characterize the electrical connectivity and reliability. As shown in Figure 46. The Tc is around 8 K (close to the intrinsic value of Nb) with an overall smooth resistance around 1.4 Ω at 2 K. The length of the superconducting path with Au interlayer is 2.5 mm. The reliability test (300 K to 2 K) shows the three resistance curves overlapping (1st, 5th, and 10th run) without any joints open and significant Tc degradation, indicating the Au interlayer technology is reliably compatible with Nb/Ir superconducting interconnects. Another key parameter to evaluate the wafer-level system for quantum computing is its current carrying capability to stay within the superconducting region. The current carrying capability is known as the critical current (Ic) or the critical current density (Jc), over which the superconducting state disappear. To calculate the limit, the Ginzburg-Landau theory with correction is used and the theoretical value in our dimension is 27.2 MA/cm^2 [125]–[127]. The control Nb/Ir wire with a width of 7 um and without the Au interlayer bonding was fabricated and was measured to have an Ic of 100 mA. This value is equivalent to a Jc (Ic divided by the interconnect cross section) of 13.2 MA/cm^2, which is near half value of the ideal one with correction. The ratio difference between the ideal calculation and the measurements of the control set is comparable to the value reported by K. Ilin *et al*. [125], validating both the calculation and the control set up in this thesis. K. Ilin *et al*. [125] report that the ratio difference comes intrinsically from non-superconducting metal and the defects inside superconducting film. The experiment set─ Nb/Ir wire with a width of 7 um and with the Au interlayer bonding─ was fabricated and was measured to have an Ic of 30 mA, eqivalent to a Jc of 4.3 MA/cm^2 (14% of the theoretical value), as shown in Figure 47. The FIB-SEM cross section image of the samples with Nb/Ir interconnects and the Au interlayer shows the fine-pitch (≤ 10 μm) demonstration in Figure 48.

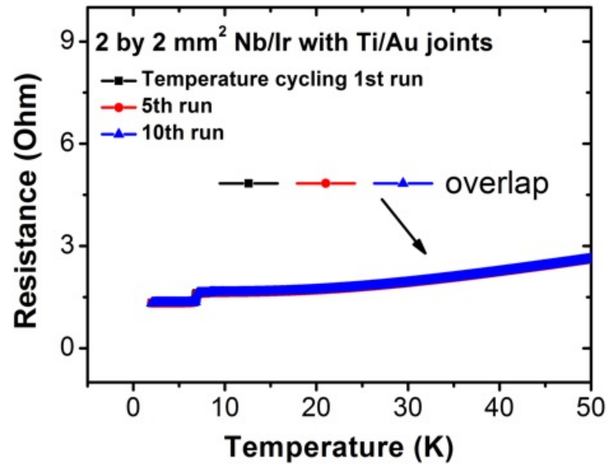


Figure 46. Electrical measurements of superconducting Nb interconnects with the Au interlayer (temperature cycling from 300 K to 2 K). Tc is around 8 K. © 2021 IEEE.

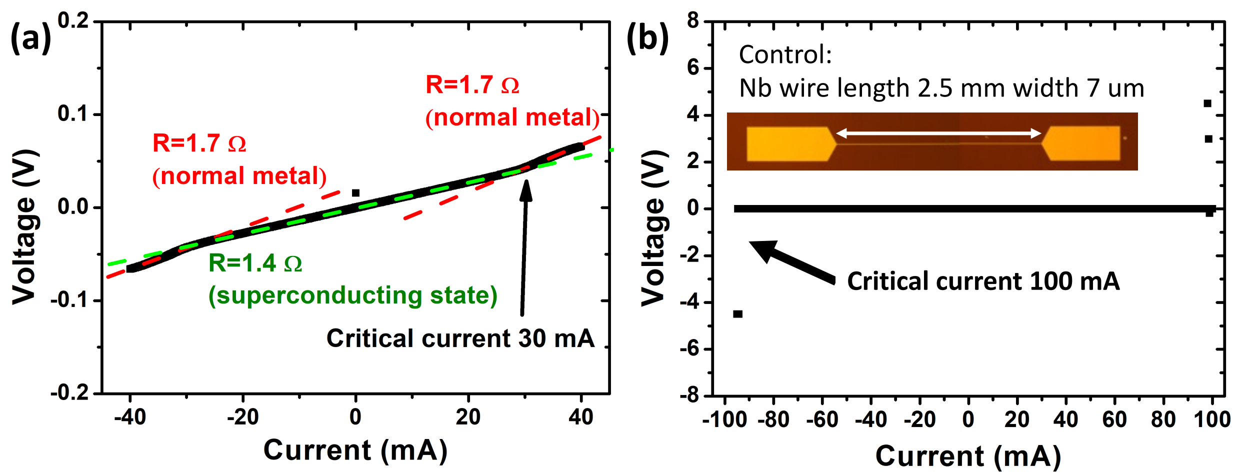


Figure 47. Critical current measurements at 4.2K without any applied magnetic field. (a) In the superconducting interconnects with the Au interlayer, the critical current is 30 mA (critical current density: 4MA/ cm^2, 14% of the theoretical value) (b) Nb interconnects control experiment without the Au interlayer flip-chip bonding. The critical current is 100 mA (critical current density: 13.2MA/ cm^2, 48% of the theoretical value). © 2021 IEEE.

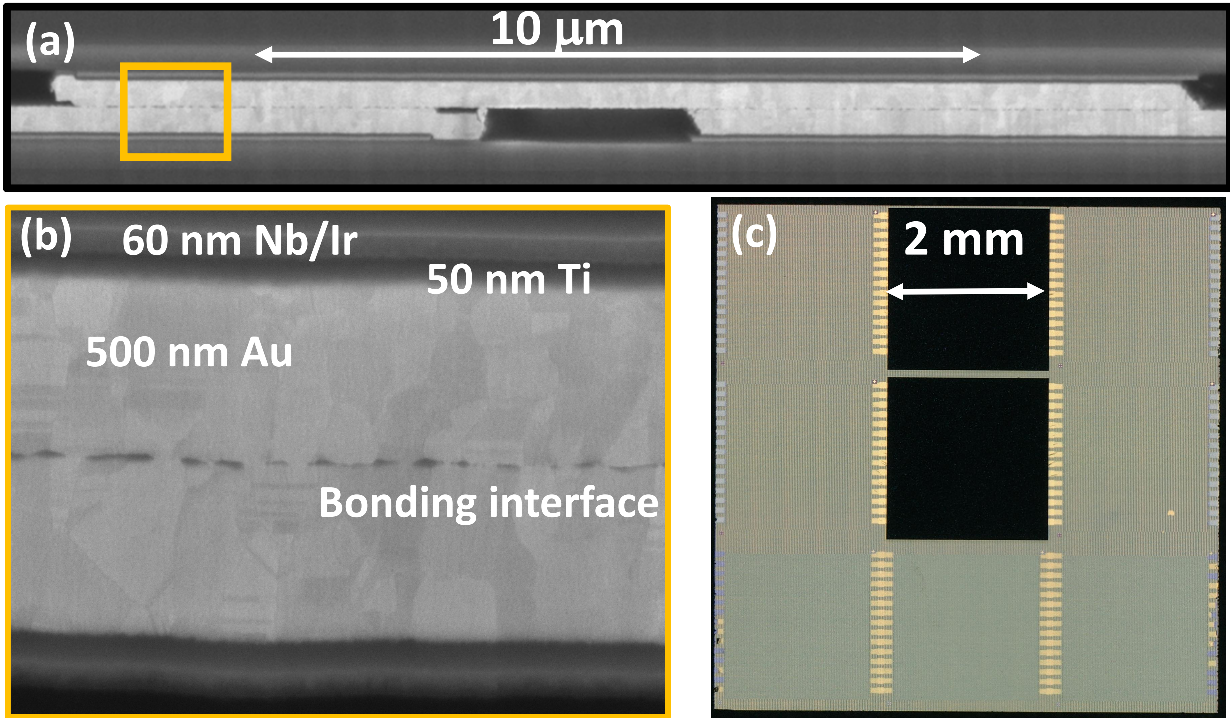


Figure 48. Cross-section image of bonded superconducting dies: (a) 10-µm-pitch Au interlayer on superconducting interconnects, (b) Zoom-in at the bonding interface of (a), (c) Bonded dies on the Superconducting-IF. © 2021 IEEE.

Although the Jc drops due to the non-superconducting Au interlayer, the superconductivity of the Nb/Ir single wire with the Au interlayer is still preserved. The most important part is that a single Nb/Ir wire with the Au interlayer can not only provide a critical current that is 7-8 X higher than the required current to drive the functional superconducting active chip with over 20,000 JJ for both processing and memory operations [34], but also find a path to fulfill the I/O requirement for quantum computing, which is listed as the most important problem to solve in Chapter 1. As compared to the Nb interconnects with In bumps [128], although an equivalent Jc of 6.6 MA/cm^2 is reported, the theoretical Jc value of 2-µm-wide Nb interconnects is 102 MA/cm^2. The ratio between the reported value and the theoretical value is only 6%, meaning the

demonstrated Au interlayer can produce less defect to the underlying Nb superconducting interconnects and can have a Jc closer to the theoretical value.

For superconducting qubits, the operating temperature is required to be in the sub-K region. This manuscript is a proof-of-concept to test whether the developed Au interlayer technology can survive as the temperature reduces to 2 K (our PPMS limit). Based on results in the subsection 5.2.1, the Au interlayer bonding is shown to be mechanically strong enough to withstand the stress of thermal contraction from 300 K to 2 K (ΔT: 300K→2K = 298 K). Since the thermal stress is linearly dependent on the temperature [121], when the temperature further decreases from 2 K to a few mK (ΔT < 2 K, which is less than 1% of ΔT: 300K→2K) in future applications, the Au interlayer should be able to overcome the extra small thermal stress. In the subsection 6.2.4, it is verified that the Au interlayer can survive experimentally down to 0.25 K using another PPMS dilution fridge module.

5.3 Collaboration with MIT Lincoln Lab

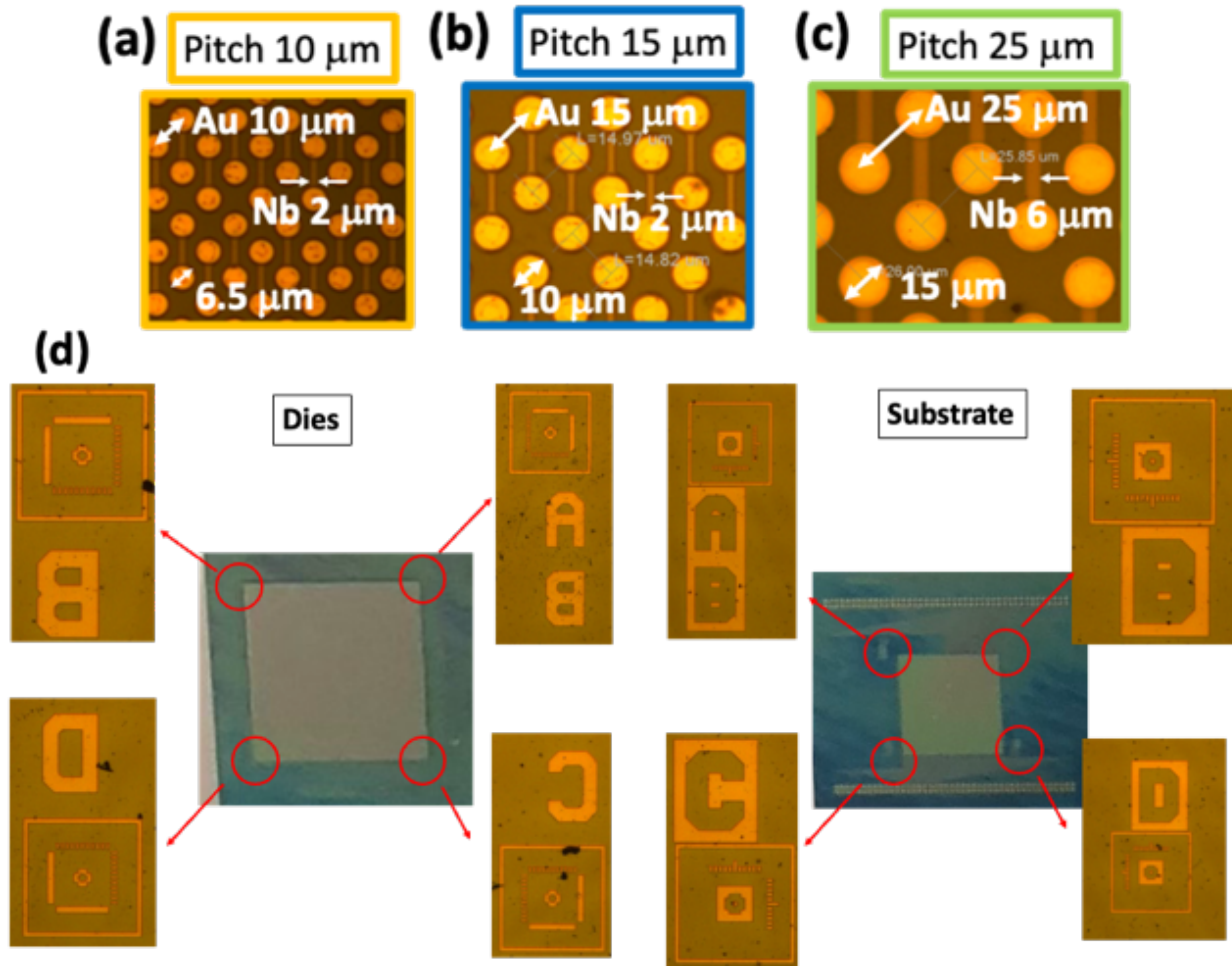


Figure 49. MIT Lincoln lab samples: (a) 10-µm pitch of Au bonding pads with a diameter of 6.5 µm and 2 µm wide Nb interconnects, (b) 15-µm pitch of Au bonding pads with a diameter of 10 µm and 2 µm wide Nb interconnects, (c) 25-µm pitch of Au bonding pads with a diameter of 15 µm and 6 µm wide Nb interconnects, (d) Top view of the dies and the substrates with alignment marks for die-to-IF flip-chip bonding.

Since part of this project is a collaboration with MIT Lincoln lab, samples made in the MIT Lincoln lab are sent to UCLA for flip-chip bonding with the developed Au interlayer technology.

The images and information of the dies with various pitches of Au pads (10, 15, and 25 μm) and the underlying Nb wires (width: 2 and 6 μm) are presented in Figure 49 and Table 9.

Table 9. Comparison between dielets with 10-μm/15-μm/25-μm pitch of Au bonding pads

Diagonal pitch (um)	Pad diameter (um)	No. of Bumps	Contacting area (mm^2)	Bonding pressure (300N/contacting area, MPa)	Au thickness in each side (nm)	Interconnect length (mm), width (um)	Ideal resistance (kΩ)
10	6.5	100k	3.3	91	200	23.88, 2	8.75
15	10	69k	5.4	55	200	16.42, 2	4.3
25	15	32k	5.6	53	200	13.13, 6	1.2

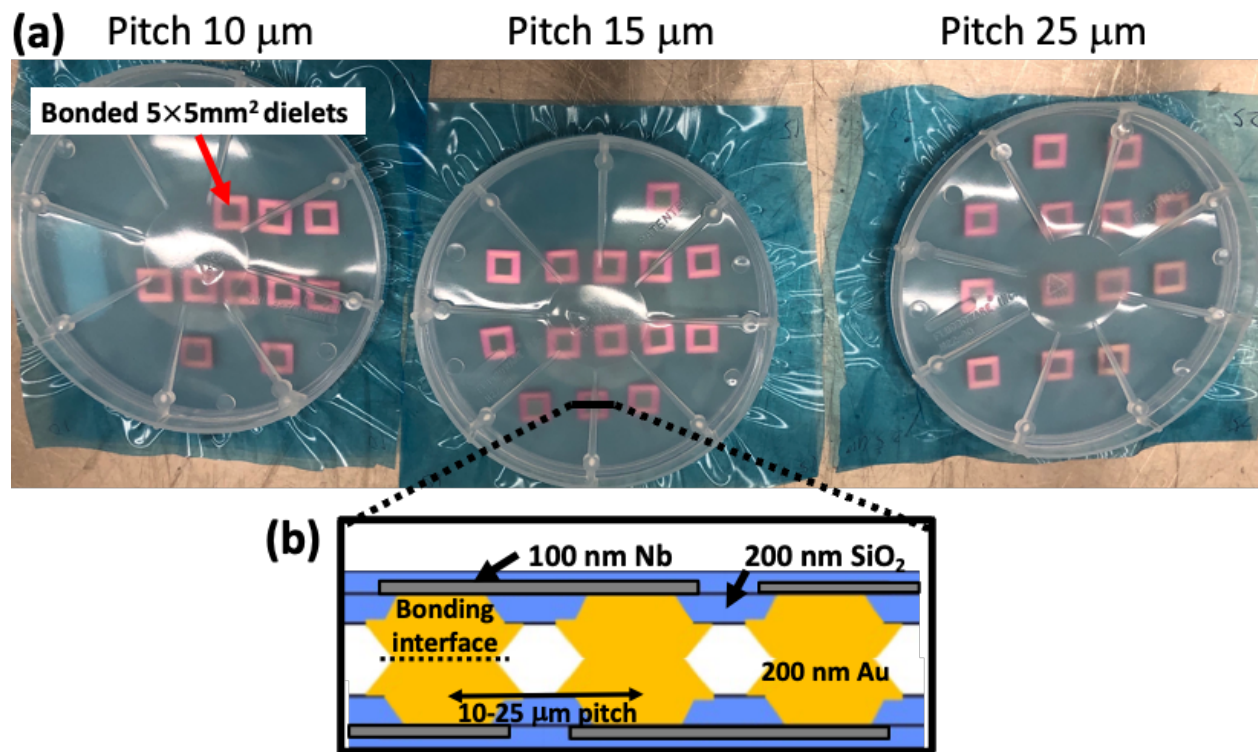


Figure 50. MIT Lincoln lab samples: (a) Top view of 5 × 5 mm^2 daisy-chain bonded dielets with pitches of 10/15/25 μm, (b) Cross-section scheme of pads on the die

The Au-Au low-temperature bonding process (bonding force: 300N, bonding time: 1.5s, and the interfacial bonding temperature 140°C) was applied to these samples and around 40 dielets were successfully bonded to corresponding substrates with an alignment overlay of $\pm$ 1 μm. Part of the bonded dies is presented in Figure 50 (a) and Figure 50(b) shows the cross-section scheme of the bonded structure. The bonded dies were first electrically tested at 300 K, as shown in Figure 51, and then shipped back to MIT Lincoln Lab for wire-bonding on designated PCB to conduct low-temperature measurements, as shown in Figure 52.

On the dies and substrates, there are two layers: the first layer is 100 nm Nb with a SiO_2 inter-layer dielectric; the second layer includes the relevant UBM layer (20 nm Ti/ 50 nm Pd), not shown in the image, and a 200 nm SiO_2 inter-layer dielectric with a 200 nm Au layer protruding above the SiO_2 layer. Within a flip-chip bonded assembly, there are 44 daisy chains and the length per daisy chain on 10-μm, 15-μm, and 25-μm pitches of dielets are around 24 mm, 16 mm, and 13 mm, respectively. For 10-μm, 15-μm, and 25-μm pitches of dielets, the corresponding calculated room-temperature resistance with Au-Au bonded joints and Nb wires are 8.75 kΩ, 4.3 kΩ, and 1.2 kΩ at 300 K. The measured results at 300 K are around 7 kΩ, 5 kΩ, and 1.25 kΩ, as shown in Figure 53(a), and all of them match the calculated values. Low-temperature measurements down to 2 K are presented in Figure 53(b) with a Tc at 9 K and a residual resistance of few hundreds of Ohms. Many assemblies were tested to be fully connected at low temperature and the data of one out of the fully-connected dies with the 10-μm pitch is presented in Figure 53(c). It shows superconducting behavior and keeps staying in the low-resistance state below Tc with smooth curves, meaning the large number of joints (up to a count of 100,000) in the bonded samples are connected. This again justifies that the developed Au-Au bonding is reliable mechanically with a 10-μm pitch.

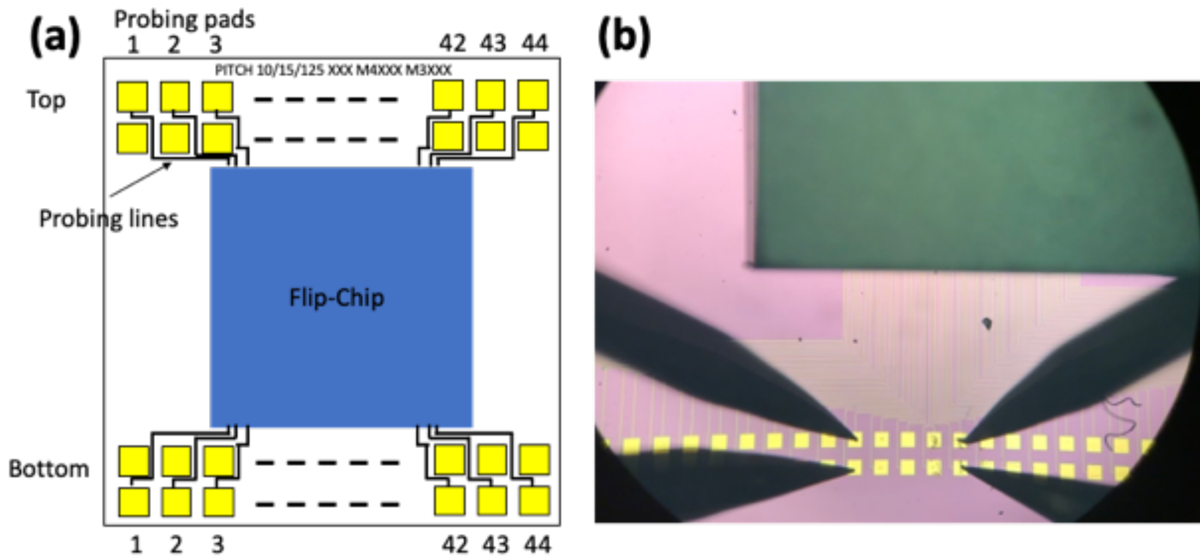


Figure 51. (a) Configuration of probing pads after a 5 X 5 mm^2 dielet flip-chip bonded Si substrate. (b) Image of four-probe measurements at 300 K under microscope.

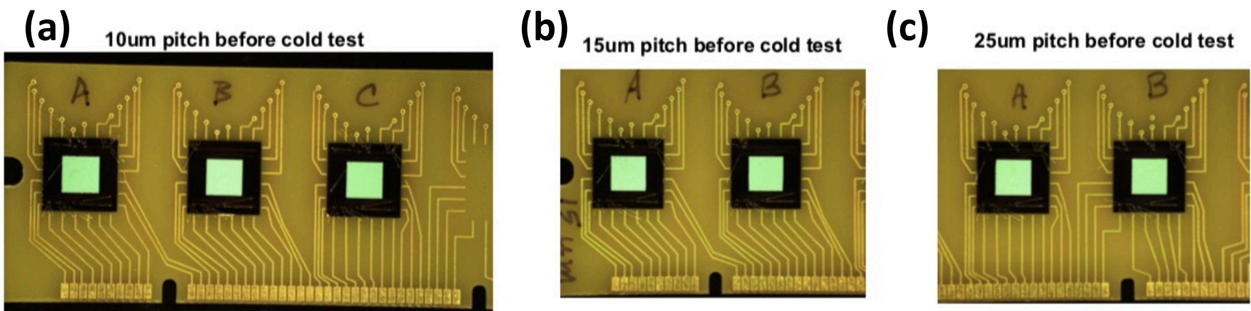
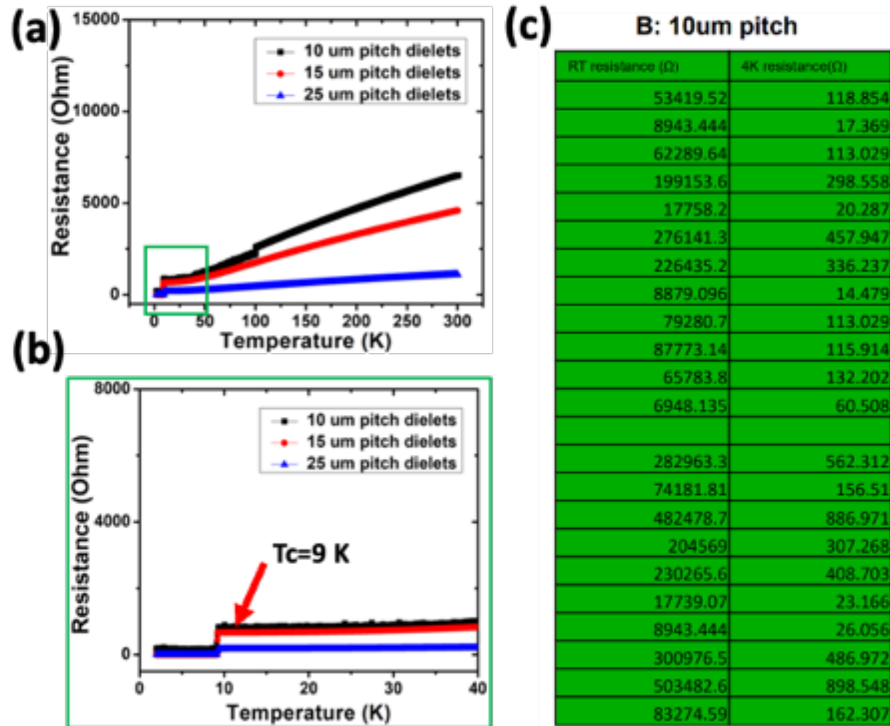


Figure 52. MIT Lincoln lab samples wire-bonded to designated PCB for low-temperature measurements: (a) 10-µm pitch of bonded dielets, (b) 15-µm pitch of bonded dielets, (c) 25-µm pitch of bonded dielets



RT resistance (Ω)	4K resistance(Ω)
53419.52	118.854
8943.444	17.369
62289.64	113.029
199153.6	298.558
17758.2	20.287
276141.3	457.947
226435.2	336.237
8879.096	14.479
79280.7	113.029
87773.14	115.914
65783.8	132.202
6948.135	60.508
282963.3	562.312
74181.81	156.51
482478.7	886.971
204569	307.268
230265.6	408.703
17739.07	23.166
8943.444	26.056
300976.5	486.972
503482.6	898.548
83274.59	162.307

Figure 53. (a) R-T measurements of the dies with 10-µm, 15-µm, and 25-µm pitches from 300K cooling down to 2K. (b) zoom-in of (a) from 40 K to 2 K. (c) One set of the measurement results probing all pads on one die with the 10-µm pitch: all of the results show superconducting behavior.

5.4 Wafer-level assembly on Superconducting-IF

As mentioned in Chapter 4, to achieve a wafer-level system integration, a wafer-level surrogating superconducting dielet demonstration is a must to show the feasibility for future integrated quantum computing. The fabrication process is as follows: (1) ideal 100 nm Nb/5 nm Ir sputter deposition without breaking the chamber on oxidized wafers, (2) patterning 100 nm Nb/5 nm Ir through a subtractive Cl_2/Ar dry etching, (3) Au interlayer deposition and patterning through the previously mentioned Au lift-off process, (4) singulation of 2 × 2 mm^2, 3 × 3 mm^2, 4 × 4 mm^2, 5 × 5 mm^2 dies, (5) fabrication of mating bonding area on wafer-level Superconducting IF, (5) ultrasonicating cleaning and Ar plasms surface activation both parts, and (6) flip-chip bonding ~ 100 surrogate superconducting dies on the Superconducting-IF. On this Superconducting-IF, heterogeneity and flexibility in the die sizes and the distance between each flip-bonded die are achieved, as shown in Figure 54. The die size of 2 × 2 mm^2, 3 × 3 mm^2, 4 × 4 mm^2, and 5 × 5 mm^2 were bonded on the same wafer with varying inter-dielet distance from 60 µm to 20 mm. The configuration can be fully populated [71], partially populated, and irregular spaced. This irregular spacing can provide a flexible design and allocation of Si area for qubit noise cancellation approaches mentioned in the section 4.2. Another method to reduce qubit noise is to apply die stitching technique and is demonstrated on large 8 × 8 mm^2 die stitched into 16 smaller 2 × 2 mm^2 die with an inter-dielet spacing of 60 µm, providing an alternative approach to prevent phonon-related qubit noise from spreading out. Superconducting-IF provides a large Si area to accommodate area-hungry qubits, interconnects, and on-wafer miniaturized micro-connectors.

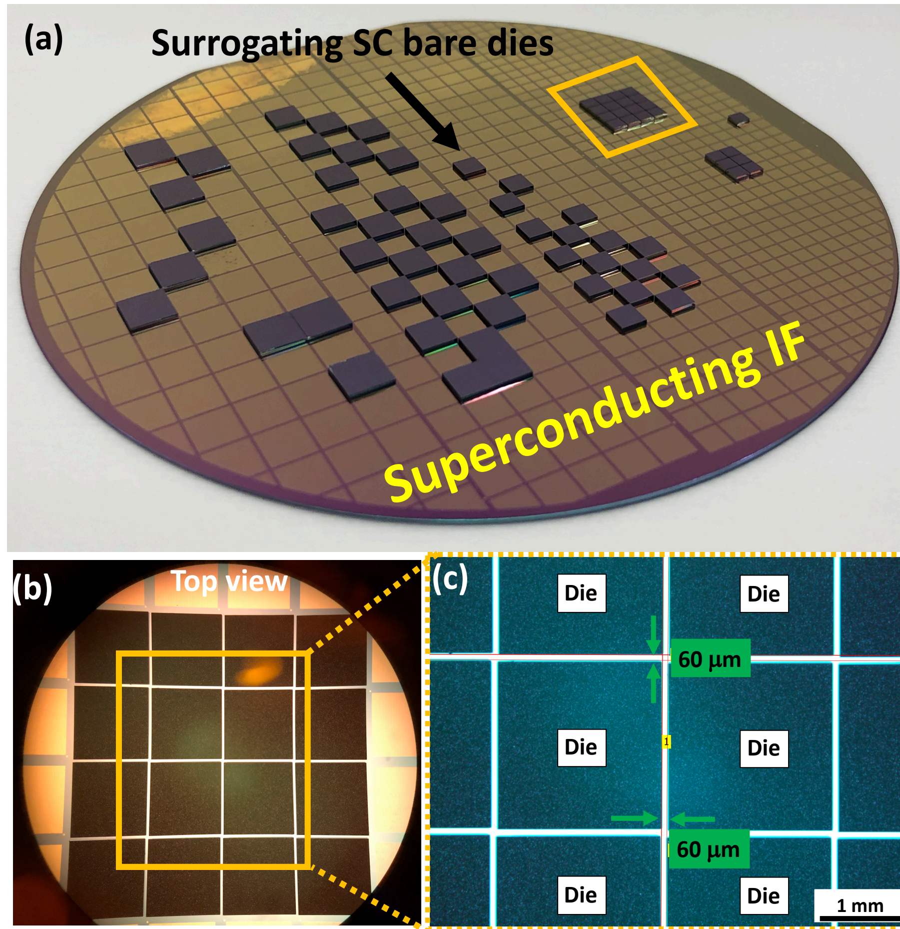


Figure 54. (a) SoW demonstration of integrated superconducting (SC) surrogating dielets with die sizes from 2 mm by 2 mm to 5 mm by 5mm. (b) Zoom-in image of highlight area in Fig. 4. The SC surrogating dies are assembled through the Au interlayer technology with an inter-dielet spacing of 60 μm. © 2021 IEEE

To assemble a huge number of dielets within a short period, an automated bonding process was developed and investigated, as shown in Figure 55(a). In the Au interlayer technology, one bonding cycle takes 8.5 seconds (die transfer of 2 seconds from a tray to the bonding head, die-to-wafer alignment of 2.5 seconds, temperature ramping up and down of 2.5 seconds, and bonding of 1.5 seconds), meaning the throughput of 425 dies per hour can be achieved. Within

the bonding period, a real thermal compression bonding only takes 1.5 second; the remaining duration of 2.5 seconds is for pre-bonding and post-bonding tooling preparation in the pick and place tool. During the 1.5 second main bonding duration, 80 MPa (blue line) and 140 °C (black line) is applied to the dies. The image of the pick& place tool used at UCLA shows in Figure 55 (b).

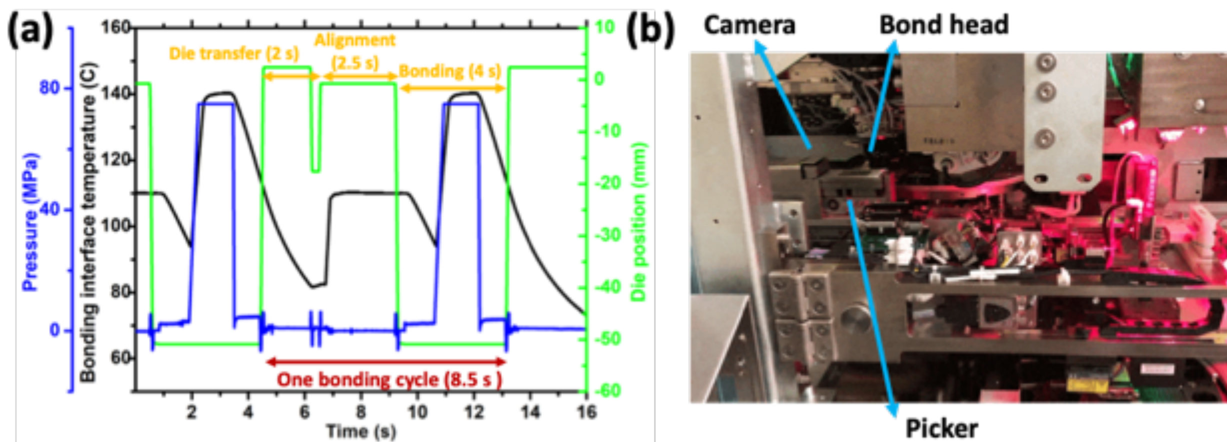


Figure 55. (a) Detail of the assembly process. In the Au interlayer technology, one bonding cycle is 8.5 second (die transfer of 2 seconds from a tray to the bonding head, die-to-wafer alignment of 2.5 seconds, temperature ramping up and down of 2.5 seconds, and bonding of 1.5 seconds). (b) Pick & place bonding tool. © 2021 IEEE

5.5 Summary

In this chapter, the deposition and fabrication of Nb interconnects is presented. The issue of Nb oxidation is detailed and the Ir is introduced on top of Nb as an effective passivation method through Tc measurements to reduce the Nb oxidation for processing. The logistics to select the Au interlayer as the flip-chip bonding medium is described for fine-pitch applications, instead of In/Sn eutectic bonding and In/Sn/Au solid-liquid interdiffusion bonding. The Au interlayer fine-pitch bonding technology was carefully examined with step-by-step pre-bonding check as follows: (1) Au diffusion into Nb during processing was verified to be insignificant and the superconductivity of Nb was unaltered after any processing, (2) the adhesion between Au and Nb/Ir was examined by the standard tape test to be excellent on a large area, and (3) the surface roughness was confirmed through AFM to have a Ra of around 3 - 4 nm, which was flat for thermal compression bonding. The Au-Au bonding DOE specific for low-thermal-budget applications was implemented to have finalized parameters: bonding interfacial temperature of 140 ℃, bonding pressure at least 35 MPa, and a main bonding duration of 1.5 seconds. As compared to existing solder integration approaches for quantum applications, the optimized Au-Au bonding can provide much higher bonding strength and much finer pitch.

Electrical characterization through the Kevin structure and the fine-pitch daisy chain structure shows that the Au interlayer technology is electrically reliable and robust with a contact resistance of 10^{-8} Ω-cm^2. As an important metric, critical current is characterized with a value of 30 mA. Although the Ic drops due to the non-superconducting Au interlayer, the most important part is that the Nb/Ir wire with the Au interlayer can still provide a critical current that is 7-8 X higher than the required current to drive the functional superconducting chip for both processing

and memory and can create a path to fulfill the dense I/O requirement for quantum computing. As compared to the Nb interconnects with In bumps, the demonstrated Au interlayer can induce less defect to the Nb superconducting interconnects and can provide a Jc closer to the theoretical value.

The samples form MIT Lincoln Lab again verified that the developed Au-Au bonding was mechanically reliable with a 10-μm I/O pitch on 5 X 5 mm^2 dies having very longer interconnects. A wafer-scale flip-chip assembly was demonstrated with the Au interlayer technology on Nb interconnects with various die sizes (2 × 2 mm^2, 3 × 3 mm^2, 4 × 4 mm^2, and 5 × 5 mm), emulating an integrated qubit control/readout system with fine-pitch interconnects and a proximal die spacing (≤ 100 μm). On this wafer-scale assembly, heterogeneity and flexibility were both demonstrated regarding the die sizes, the distance between each flip-bonded die, and configuration. A large Si area can be provided for qubit noise cancellation methods, area-hungry qubits, fan-out interconnects, and on-wafer miniaturized micro-connectors for communication to the outside world. To populate the wafer-scale assembly, an automated bonding procedure was introduced with an average bonding duration of 8.5 s for each dielet, having a high throughput of 425 dies per hour.

Chapter 6 RF characterization on Superconducting-IF

As mentioned in Chapter 4, in the proposed concept of wafer-level integration on Superconducting-IF for large-scale quantum computing, qubit array and periphery JJ-based superconducting control/readout dies are co-integrated on the system-on-wafer platform. Within the assembly, in addition to DC bias trees for the power of qubit array and periphery systems, RF electrical signals are also required to manipulate quantum states, to exchange quantum information, and to communicate between dielets for on-chip-like qubit control/readout. Those RF signals in the range of 5-10 GHz for qubit manipulation and of up to 20 GHz for pulse control [130], [131] are carried through superconducting transmission lines, mostly in the form of coplanar waveguides. For the inter-dielet communication, it includes SFQ signals departing from I/O pads on a previous superconducting dielet, transmitting through superconducting coplanar waveguides between dielets on the Superconducting-IF, and arriving at I/O pads on a next superconducting dielet. The overall signal transmission distance, or the link length, contains the previously mentioned inter-dielet spacing ($\leq$ 100 μm) adding I/O depth from dielet edges. Since the on-chip quantum information (a signal amplitude of few hundreds to few tens of μV) and the on-chip JJ-based controlling signals (a signal amplitude of on average 300 μV [132], [133]) are delicate and both of them should be well-preserved during inter-die communication, the passive RF behavior within superconducting coplanar waveguide links, including insertion loss and crosstalk, is important to characterize

In this thesis, the goal focuses on a broadband (from 50 MHz to 20 GHz) RF characterization, including insertion loss and crosstalk, of various link lengths and wire line & space (L/S) within superconducting coplanar waveguides. The links are designed with three various lengths: 125

μm (extremely short link length on Si-IF [91]), 500 μm (short link length on Si-IF [91]), and 1750 μm (long link length on Si-IF) and with two types of L/S: 2/2 μm and 5/5 μm. RF simulation was first conducted under previously mentioned dimensions and the results are presented in the section 6.1. Experiments of RF characterization were implemented and are detailed in the section 6.2 (1) the setup of measurement apparatus (subsection 6.2.1), (2) the logistics and the design of sample holders (subsection 6.2.2) and test vehicles, (3) sample fabrication (subsection 6.2.3), and (4) measurement results at 4 K (subsection 6.2.4).

Although the Nb superconducting interconnects with the Au interlayer integration is characterized at 4 K, a temperature higher than the operating temperature of superconducting qubits (sub 100 mK), the RF behaviors of the interconnects should be close to that at sub 100 mK. The RF behavior (characteristic impedance and insertion loss) of the superconducting transmission line is determined by the resistance (R), inductance (L), conductance (G), and capacitance (C). For R, the Nb film already transitions to a low-resistance superconducting state both at 4 K and at mK region and the experimental input shows in Figure 73 and Figure 74. For L, the permeability of nonmagnetic material (Si, SiO_2, air, and Nb used in this thesis) does not vary significantly within few Kelvin difference between 4 K and mK. One thing to note that the kinetic inductance should be considered and added on top of magnetic inductance in the superconducting region. In the subsection 6.1.1, it is explained that through proper dimension selection, the kinetic inductance can be neglected. For C, the permittivity varies much less than 1 % at few Kelvin in SiO_2[134], Si [135], [136] and air. For G, the dielectric conductance is reduced at cryogenic temperature and does not vary much within few Kelvin difference between 4 K and mK. At 4 K and mK, the overall value of the characteristic impedance is close to $\sqrt{\frac{L}{C}}$.

The insertion loss includes R, G and loss tangent loss and at 4 K and mK, the R, G, and the loss tangent loss [135] are close. The RF crosstalk behavior is mostly related to the dimension of transmission line, which is not a temperature-dependent term. Due to much faster turnaround time during measurement, RF characterization at 4K is selected. The broadband measurement frequency is chosen from 0.05 GHz up to 20 GHz to cover the frequency span for qubit excitation and JJ-based controlling. In the section 6.3, passive dies with resonators integrated on Superconducting-IF are presented but due to the tool issue, only simulation results are presented. The summary of this chapter is presented in section 6.4.

6.1 Superconducting coplanar waveguide simulation

In this section, superconducting coplanar waveguide simulation was conducted and is detailed in the following sequence: simulation assumption & approximation (subsection 6.1.1), model establishment (subsection 6.1.2), results & analysis (subsection 6.1.3).

6.1.1 Simulation assumption & approximation

As compared to the normal metal RF simulation using classical transmission line theory in ANSYS HFSS, superconducting RF simulation needs to consider two extra metrics: one is the penetration length compared to the superconducting film thickness; the other is the kinetic inductance. The first metric determines whether the traditional transmission line theory or a modified layer-by-layer integration [137] method should be applied for the calculation of the characteristic impedance and the scattering metrics. The second metric increases the value of overall inductance on top of magnetic inductance, affecting the characteristic impedance in the simulation model. In the next two paragraphs, with adequate selection of the film thickness and the dimension of the structure, it is stated that the two above-mentioned metrics (penetration

length compared to the superconducting film thickness and the kinetic inductance) can be ignored in the simulation model and the experimental realization. Following the above-mentioned approximation, the classical transmission line theory in ANSYS HFSS is still applicable to the superconducting RF simulation in this thesis without the need of dramatic change.

In the first metric, the penetration length is a depth scale describing how deep the external magnetic field can penetrate a superconducting film. Under an ideal case, based on the Meissner effect, external magnetic fields can be completely excluded from percolating into a superconducting material. In a real case, magnetic fields can penetrate superconducting film in a small scale, known as the penetration length. The penetration length is a thickness-dependent metric, pointed by Gubin *et. al* [138], and is inversely proportional to the thickness of a superconducting film (Nb in this paper). This indicates that with a thicker film, the penetration length gets thinner. Also indicated by Gubin *et. al* [138], when the Nb film is thicker than 200 nm, the penetration length is reduced to around 80 nm. Considering both the top and the bottom superconducting surfaces, there is an impenetrable superconducting film in the middle. This matches the definition of a bulk superconductor reported by Gao [137] and classical transmission line theory can be applied in the superconducting RF simulation with a film thickness over 200 nm. Based on this, Nb film with a thickness of 200 nm is selected in the simulation and the later testing vehicle.

The second metric, the kinetic inductance, should be considered when samples are in the superconducting region. Inductance includes both magnetic inductance and kinetic inductance. The former comes from magnetic energy in magnetic fields, while the latter originates from the kinetic energy of the surface superconducting current distributed within the penetration depth. In

a line width of micrometer scale [137] with a Nb thickness of 200 nm [139], the ratio between the kinetic inductance and the whole inductance is only few percent, which is a negligible number. This can justify the use of classical transmission line theory in the ANSYS HFSS solver to simulate the superconducting RF behavior in this thesis.

6.1.2 Simulation model establishment

As mentioned previously, superconducting coplanar waveguides are used to simulate the insertion loss and crosstalk with various link lengths and wire L/S. The link length is defined as the length scale between two edge I/Os on two nearby dielets and three various lengths (125 μm, 500 μm, and 1750 μm) are investigated. Both the link length of 125 μm and 500 μm are the focus in this dissertation since both of them can only be achieved through advanced packaging; the link length of 1750 μm is used as the reference for comparison. The wire L/S is defined as the interconnect width (L) and the gap between interconnect (S) in coplanar waveguides. In the following models for insertion loss and crosstalk, the characteristic impedance in HFSS is tweaked to match around 50 Ω.

To simulate RF behavior of insertion loss, a ground-signal-ground (GSG) configuration was selected, as shown in Figure 56 (a)-(c). The material stack at the coplanar waveguide part from the bottom to the top is Si substrate, 500 nm of SiO2, and 200 nm of Nb/Ir. At the two opposite ends of the superconducting waveguide, there are flip-chip bonding sites, at which the material stack is Si substrate, 500 nm of SiO2, and 200 nm of Nb/Ir, 50 nm of Ti, 1000 nm of Au, 50 nm of Ti, and 200 nm of Nb/Ir connecting the previous and next die. The coplanar wave excitation is input at the two ends. Since there is no superconducting related module in the ANSYS HFSS solver, the superconducting Nb interconnects are interpreted as interconnects with an ultra-high

electrical conductivity, which is input from an empirical value of 1x10^{11} Siemens/m. This value is three orders higher than Cu interconnects.

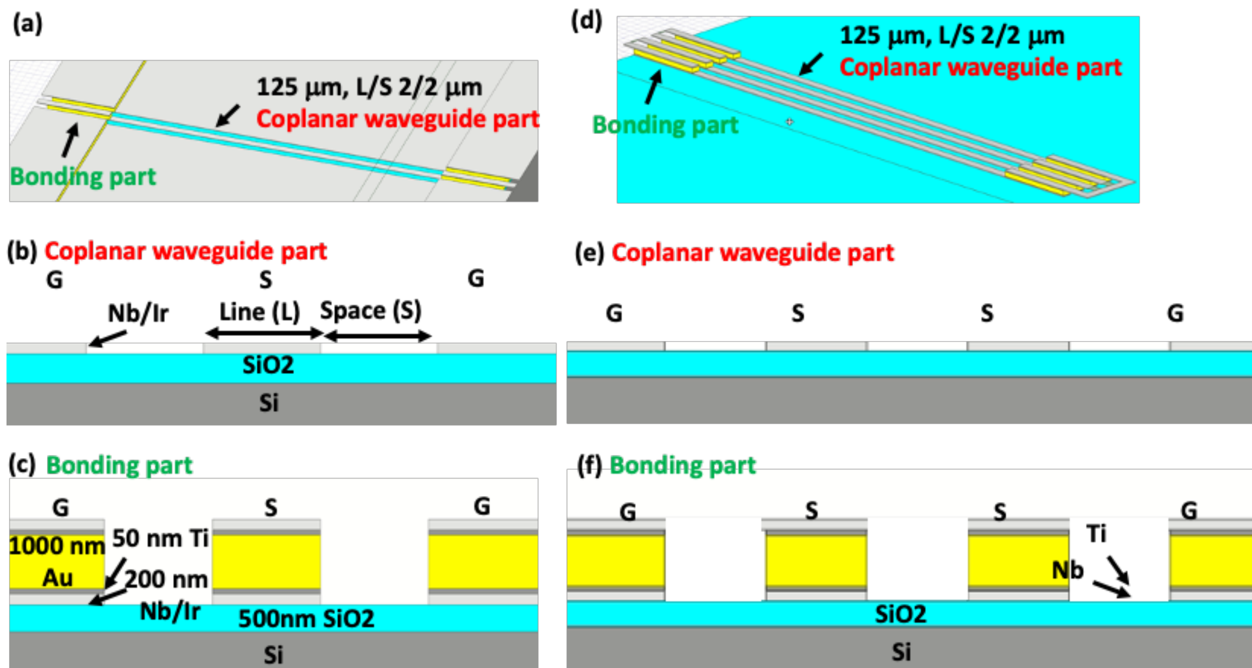


Figure 56. Simulation model: (a) Isomeric view of GSG 125 μm link length with a L/S 2/2 μm. (b) GSG cross section at the coplanar waveguide part with material stack from bottom to top: 500 μm Si (not to full scale) , 500 nm SiO_2, 200 nm Nb/Ir. (c) GSG cross section at the bonding part with a material stack from bottom to top: 500 μm Si (not to full scale), 500 nm SiO_2, 200 nm Nb/Ir, 50 nm Ti, 1000 nm Au, 50 nm Ti, 200 nm Nb/Ir. (d) Isomeric view of GSSG 125 μm link length with a L/S 2/2 μm. (e) GSSG cross section at the coplanar waveguide part with a material stack from bottom to top: 500 μm Si (not to full scale) , 500 nm SiO_2, 200 nm Nb/Ir. (f) GSSG cross section at the bonding part with material stack from bottom to top: 500 μm Si (not to full scale) , 500 nm SiO_2, 200 nm Nb/Ir, 50 nm Ti, 1000nm Au, 50 nm Ti, 200 nm Nb/Ir. © 2022 IEEE

To simulate RF behavior of crosstalk, a ground-signal-signal-ground (GSSG) configuration was selected, as shown in Figure 56 (d)-(f). Two signal paths are brought close to each other and the material stack and parameters are the same as that in the insertion loss: Si substrate, 500 nm of SiO2, and 200 nm of Nb/Ir (from the bottom to the top) at the coplanar waveguide part; at the flip-chip bonding sites, the material stack is Si substrate, 500 nm of SiO2, and 200 nm of Nb/Ir, 50 nm of Ti, 1000 nm of Au, 50 nm of Ti, and 200 nm of Nb/Ir connecting the previous and next die.

6.1.3 Simulation results & analysis

The simulated insertion loss of three various lengths (125 µm, 500 µm, and 1750 µm) with two L/S (2/2 and 5/5 µm) are presented in Figure 57.

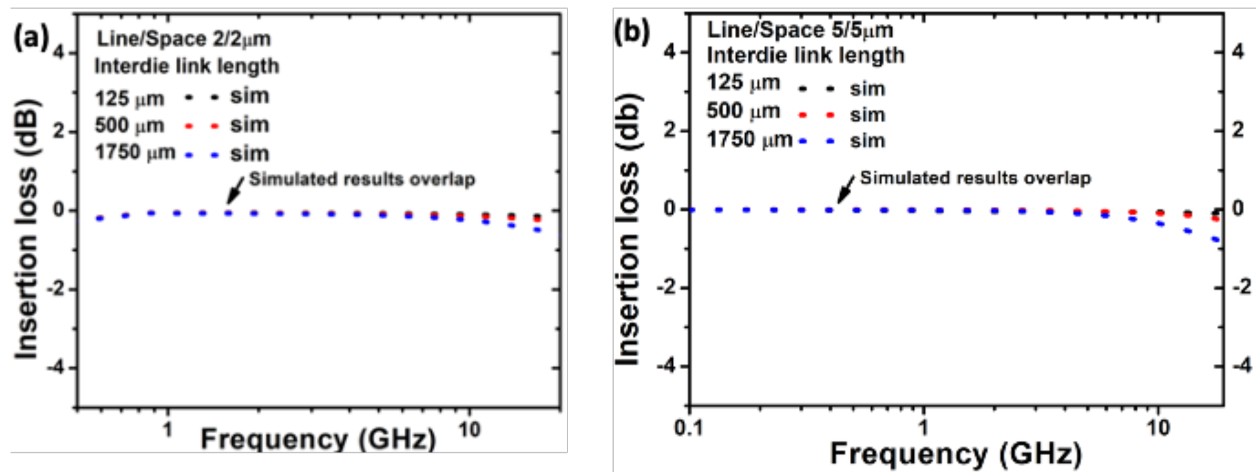


Figure 57. Simulated insertion loss results (a) L/S 2/2 µm with link lengths of 125 µm, 500 µm, and 1750 µm. (b) L/S 5/5 µm with link lengths of 125 µm, 500 µm, and 1750 µm. © 2022 IEEE

Insertion loss is a loss term comprised of the metal loss (finite conductivity within frequency-dependent skin depth) and the dielectric loss (dielectric conductance and loss tangent loss).

Dielectric loss is the dominant part when Nb is in the high-conductivity superconducting region as metal loss is inverse square root proportional to the conductivity. For a longer link with a higher frequency, the dielectric loss increases linearly, leading to an overall higher insertion loss with frequency. As an example in Figure 57(a), for a 125 μm link length at 10 and 20 GHz with a L/S of 2/2 μm, the value of insertion loss are 0.02 dB and 0.06 dB. For a 1750 μm link length at 10 and 20 GHz with a L/S of 2/2 μm, the value of insertion loss are 0.16 dB and 0.5 dB. As for a larger L/S, the insertion loss increases since more electromagnetic field distributes in lossy SiO_2 and Si dielectrics. As an example in Figure 57(a) & (b), for a L/S of 2/2 and 5/5 μm with 125 μm link length at 20 GHz, the value of insertion loss of are 0.06 dB and 0.1 dB. For a L/S of 2/2 and 5/5 μm with 1750 μm link length at 20 GHz, the value of insertion loss are 0.5 dB and 0.9 dB. The simulated value of insertion loss with a link length of 125/500/1750 μm and a L/S of 2/2 μm and 5/5 μm is very low. The comparison between the simulated and the measured loss values is presented in the subsection 6.2.4.

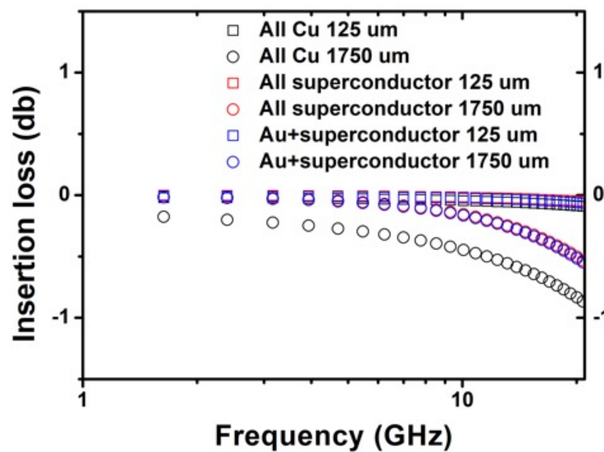


Figure 58. Results of simulated insertion loss with 125 and 1750 μm link lengths and different materials (all Cu, all superconductor, and Au interlayer with superconducting wires in this thesis). © 2022 IEEE

Jangam *et al.* carries out similar analysis but with the normal metal Cu [91], instead of superconducting material. The comparison of simulated insertion loss between the results in [91] and this thesis is presented in Figure 58. In the simulated results, there are three conditions regarding the material used in both the waveguide part and the bonding part: (1) all Cu, (2) superconductor in the waveguide part with the Au interlayer in the bonding part, and (3) all superconductors. Based on the simulated results, the signal loss using Cu interconnects is similar to that using the Au interlayer with Nb superconducting wires and using all superconductor. This comes from the advantage of ultra-short link length (125 μm), meaning signals are exposed to much less metal loss and dielectric loss during the transceiving process. Insertion loss, however, is not the only factor that determines what materials are suitable for quantum computing. Thermalization and thermal phonon noise are the two key metrics as emphasized in the Chapter 1.

Table 10 A case study for Nb-based and Cu based interconnect

Per link (L X W X T)	(L) 125 um (W) 2 um (T) 2 um		(L) 10.5 mm (W) 600 um (T) 70 um	
Technology	**Nb-based Si-IF**	**Cu-based Si-IF**	**PCB**	**limit**
DC resistance per link (mΩ)	0.1	125	1	≤ 1
I^2 X R Joule heating of 300 μA (nW)	0.01	11.3	0.09	≤ 0.1
2000 edge links per 5 X 5 mm² die with 10 um I/O pitch (nW)	20.0	22600 (> cooling power)	180	≤ 200

As a case study in Table 10, the resistance limit per interconnect should at most be limited to 1 mΩ (estimated from [82] with a corresponding operating current of 300 μA [140]). In this way, the I^2 X R joule heating per link and 2000 edge links per 5 X 5 mm^2 die with 10 μm I/O pitch can be restricted within 0.1 nW and 200 nW. When using Cu interconnects on PCB, to satisfy the

resistance threshold and long wiring length, wires with width of 600 μm and thick Cu film of 70 μm are required. This wiring is acceptable on large-pitch PCB, but to scale down interconnect pitches, the Cu resistance increases and unavoidably exceeds the resistance limitation. As an example, when the Si-IF technology is adopted, the dimension of interconnects can be pushed to extreme to achieve width of 2 μm, inter-die link length of 125 μm, and thickness of 2 μm. For Cu-based Si-IF, the resistance per interconnect increases to 125 mΩ; the corresponding 2000 edge links adds up to 22,600 nW, which excels the cooling power of 20 μW [13] at sub 100 mK mixing chamber. In addition to the system thermalization problem, the huge thermal noise below 1 K within the Cu-based Si-IF interconnect is around 1.18 μV, which is considered noisy and is very high compared to a noise level in the scale of nV [140]. With the Nb-based Si-IF, the resistance per interconnect can be reduced to 0.1 mΩ and 20 nW for 2000 edge links, which is significantly lower than the power limitation, and the thermal noise is around 0.03 nV. Lower resistance of Nb interconnects is favorable for future large-scale quantum applications. Therefore, it can be considered that interconnect scaling for quantum computing is enabled by changing material from Cu to Nb.

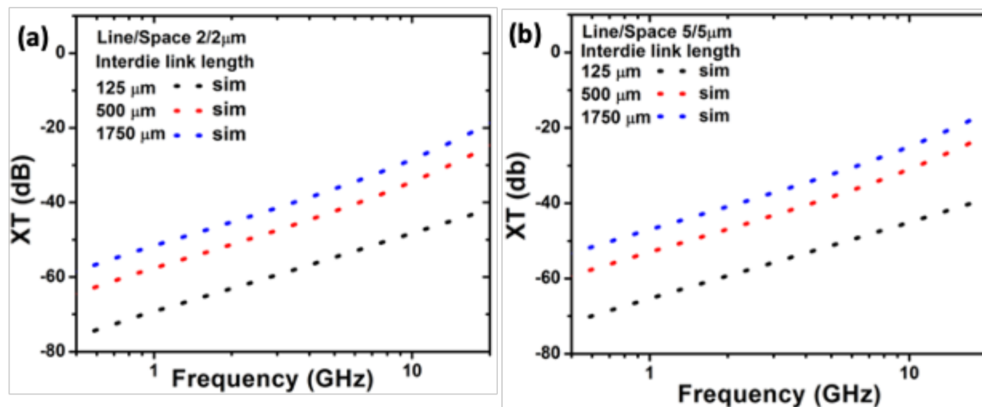


Figure 59. Simulated crosstalk results (a) L/S 2/2 μm with link lengths of 125 μm, 500 μm, and 1750 μm. (b) L/S 5/5 μm with link lengths of 125 μm, 500 μm, and 1750 μm. © 2022 IEEE

Crosstalk is an electromagnetic energy sharing behavior between neighboring signal paths. In general, the value of crosstalk increase as the gap between two signal paths reduces and the length of signal paths increases. As expected, the simulated crosstalk adds up with longer link length and higher frequency. The simulated crosstalk results of three various lengths (125 µm, 500 µm, and 1750 µm) with two L/S (2/2 and 5/5 µm) are presented in Figure 59. In Figure 59(a), for a 125 mm link length at 10 and 20 GHz with L/S of 2/2 µm, the value of crosstalk are -50 dB and -42 dB. For a 1750 µm link length at 10 and 20 GHz with L/S of 2/2 µm, the value of crosstalk are -29 dB and -19 dB. The simulated crosstalk has comparable value in L/S of 2/2 µm and 5/5 µm. In Figure 59(a) & (b), for a 125 µm link length at 20 GHz with L/S of 2/2 and 5/5 µm, the value of crosstalk are -42 dB and -38 dB. For a 1750 mm link length at 20 GHz with L/S of 2/2 and 5/5 µm, the value of crosstalk are -19 dB and -16 dB. The simulated value of crosstalk with a link length of 125/500/1750 µm and a L/S of 2/2 µm and 5/5 µm is very low. The comparison between the simulated and measured loss values is presented in subsection 6.2.4.

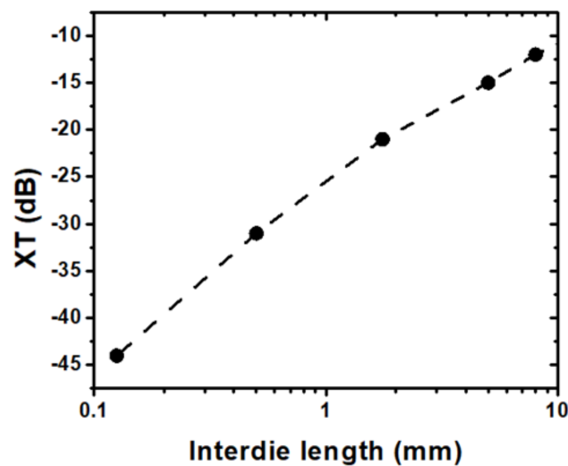


Figure 60. Crosstalk versus the length of inter-die coplanar waveguide at qubit operating frequency of 5 GHz

Since one of the future applications of the crosstalk effect in coupled transmission line can be used in superconducting qubit readout circuits, the correlation between the crosstalk and the length of inter-die coplanar waveguide is summarized in Figure 60. The value of crosstalk increases with a longer length and reaches around -15 dB in a waveguide length of 5 mm. This length-dependent increase mainly comes from capacitive coupling and is briefed in the next paragraph.

Based on the coupled transmission line theory with lossless lines and non-magnetic material [141]–[144], the crosstalk originates from the capacitive and inductive coupling of the two lines. As the length of inter-die coplanar waveguide increases, not only the capacitance and the self inductance in a single line add up, but also the mutual capacitance and the mutual inductance between two coupled transmission lines can rise. The combination of capacitance and inductance and the change of characteristic impedance cause the capacitive coupling to increase much faster than the inductive coupling since inductance is only a weak function of line geometry and is primarily controlled by the location of current return paths.

6.2 RF characterization of the superconducting interconnects on the Superconducting-IF

In this section, to characterize the RF behavior of the superconducting interconnects on the Superconducting-IF at 4K, dilution fridge setup for measurements (subsection 6.2.1), PCB sample holder design (subsection 6.2.2), passive superconducting die and Superconducting-IF design (subsection 0), fabrication process and tweaking (subsection 6.2.3), measurement results and discussion (subsection 6.2.4) are presented.

6.2.1 Dilution fridge setup

As mentioned previously, due to much faster turnaround time during measurements, RF characterization at 4 K was chosen and a dilution fridge was selected due to its capability for four two-port RF measurements per cooling down cycle, maximizing the efficiency of liquid Helium (LHe) usage. The Oxford Instrument wet system KelvinoxMX is the main dilution fridge used in this thesis. The main chamber is installed underground and the cylinder, on which the sample is installed, is controlled by a mechanical arm, through which the cylinder can be loaded into and unloaded from the chamber within three hours (including cooling down) in each measurement. To cool down the system, liquid nitrogen is first used to cool down the chamber to 77K, followed by LHe cooling down to 4 K. Each time the dilution fridge consumes around 100 liter LHe to replace a sample since the chamber needs to be brought back to room temperature and re-cooling. LHe can leak during storage in the fridge and during transferring between the LHe dewar and inside the fridge. To do measurement through Vector Network Analyzer (VNA 8720es), a series of cabling and connectors are required: (1) VNA terminated with two-port

SMA female, (2) two coaxial cables with 3.5 mm male and 2.4 mm female, (3) two adaptors with 2.4 mm male and 2.4 mm male, (4) eight fridge interface with 2.4 mm female and 2.4 mm male, (5) eight three-inch coaxial cables with 2.4 mm female and 3.5 mm male, and (6) eight SMA female on PCB, as shown in Figure 61. The PCB sample holder is supported and is heat-conducted by hand-formable coaxial cables (UT-085-FORM with a center metal of Silver-Plated Copper Weld, SPCW), having a thermal conductivity of 300-400 W/m X K at 4K [13]). When cooling down, the Oxford Instrument wet system KelvinoxMX is protected by a vacuum can and by a sliding can. The vacuum can is installed on the cylinder and is sealed by an Indium wire with vacuum glue at the interface, through which a vacuum level of 10^{-6} torr is achieved to reduce water vapor condensation into ice at low temperature around the sample. The sliding can is made with the G-10 glass that can withstand huge thermal gradient during system loading and unloading and provides mechanical support, as shown in Figure 62 (a). The loading and unloading process is presented in Figure 62 (b)& (c).

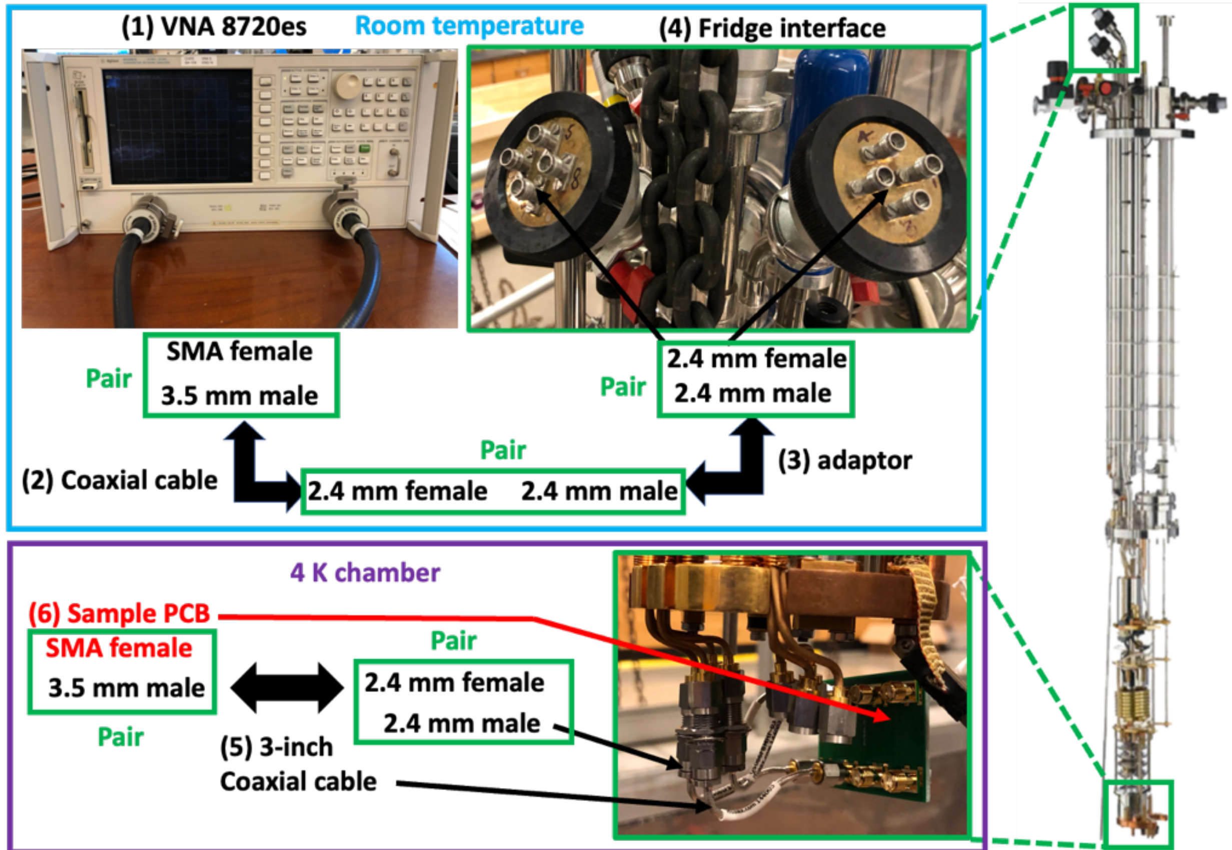
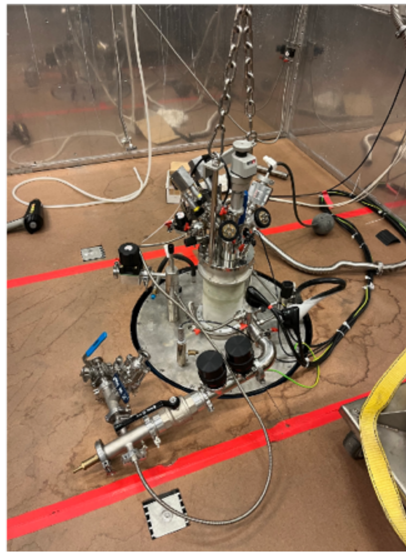


Figure 61. A series of cabling and connectors to connect from VNA 8720es at room temperature to PCB at 4 K: (1) VNA terminated with SMA female, (2) coaxial cables with 3.5 mm male and 2.4 mm female, (3) adaptors with 2.4 mm male and 2.4 mm male, (4) fridge interface [145] with 2.4 mm female at room temperature and 2.4 mm male at 4 K, (5) three-inch coaxial cables with 2.4 mm female and 3.5 mm male, and (6) SMA female on PCB.



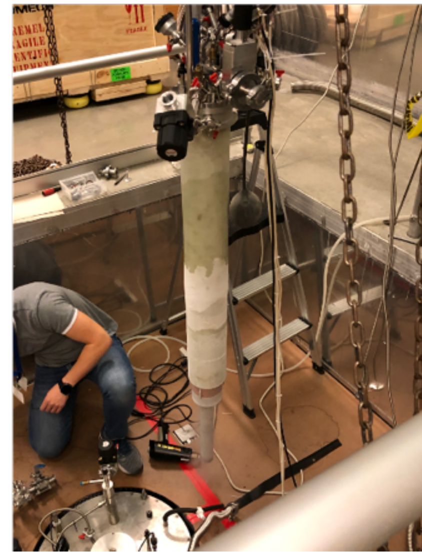


Figure 62. Main dilution refrigerator used in this thesis: (a) vacuum preparation of the cylinder before loading the sample to fridge chamber. (b) Cylinder inside the fridge chamber. (c) Cylinder unloading from the chamber.

As a backup dilution fridge, the BLUEFORS LD dilution fridge is a dry system providing two two-port RF channels, as shown in Figure 63. Instead of an underground system, the BLUEFORS LD dilution fridge is supported by four over-ground standing arms and is shielded by layers of stainless steel. For each time loading and unloading, it requires three human powers and one to two days of cooling and warming. The sample holder is designed, as shown in Figure 64, and is made with Oxygen-Free High thermal Conductivity (OFHC) Cu that can provide high thermal conductivity down to millikelvin. The test run is shown in Figure 65 with temperature sensors in each cryostat.

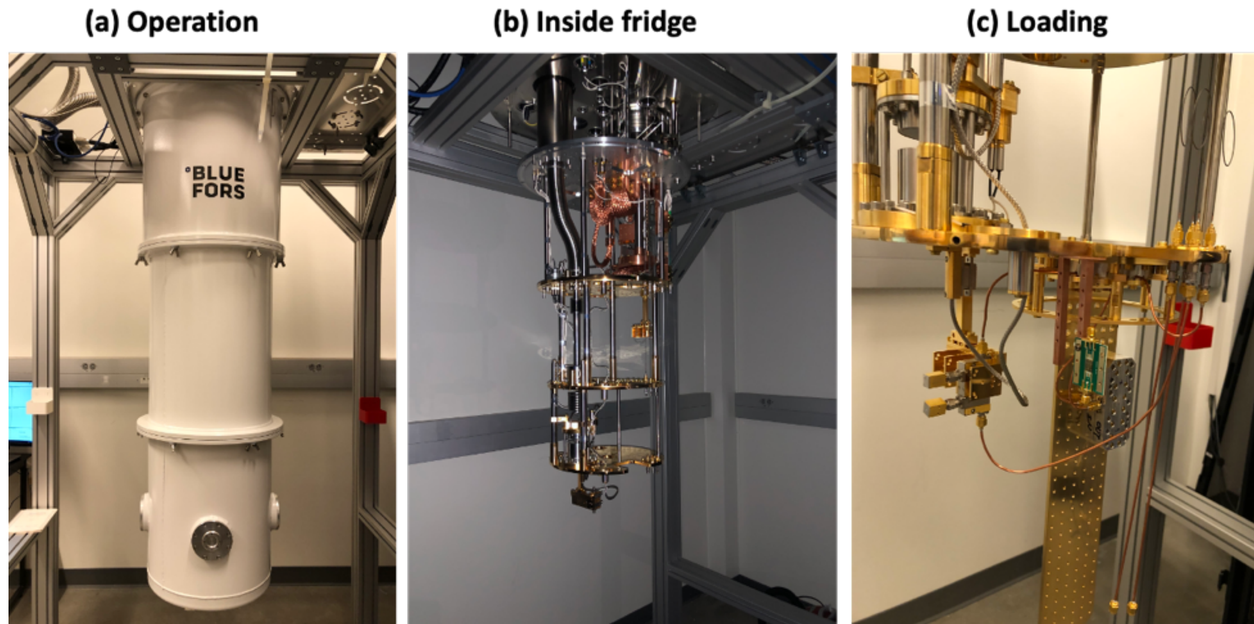


Figure 63. BLUEFORS LD dilution refrigerator: (a) during operation, (b) chamber inside, and (c) sample loading

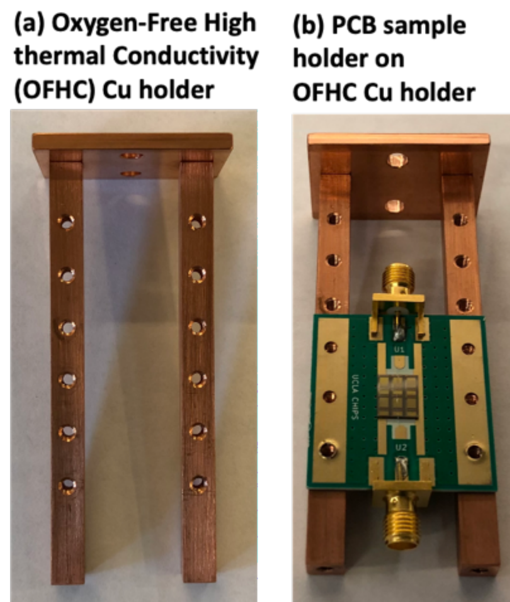


Figure 64. (a) Oxygen-Free High thermal Conductivity (OFHC) Cu holder with screw hole to fix PCB holder. (b) PCB sample holder on Cu holder.

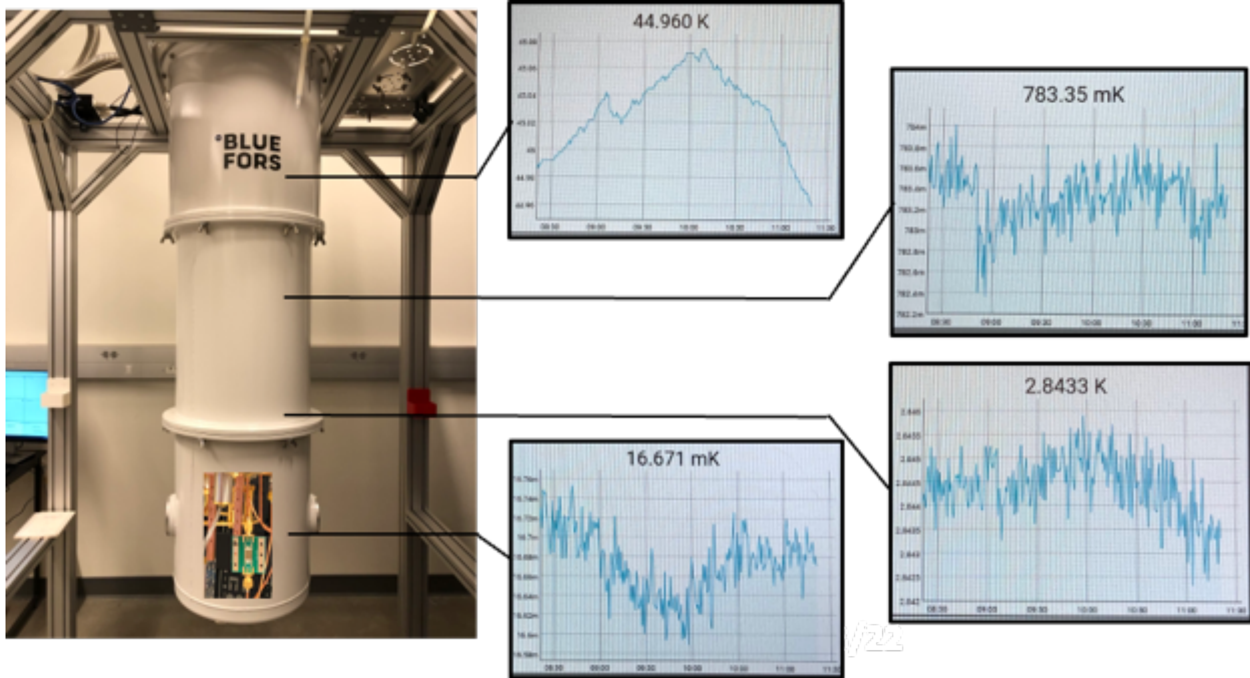


Figure 65. BLUEFORS LD dilution fridge in a test run with progressively reduced temperature in each cryostat.

6.2.2 PCB sample holder design

Figure 66. The design specs of the PCB carrier with 50 Ω coplanar waveguides with ground plane in the KiCAD PCB EDA.

Due to the limit of Oxford Instrument dilution fridge, instead of cryogenic probes that contact samples directly, four two-port RF 2.4 mm I/O within the fridge are used. To connect to the fridge 2.4 mm male I/O at 4 K chamber, coaxial cables with 2.4 mm female bulkhead on one side and 3.5 mm male on the other side are utilized to connect SMA connectors a PCB sample mount.

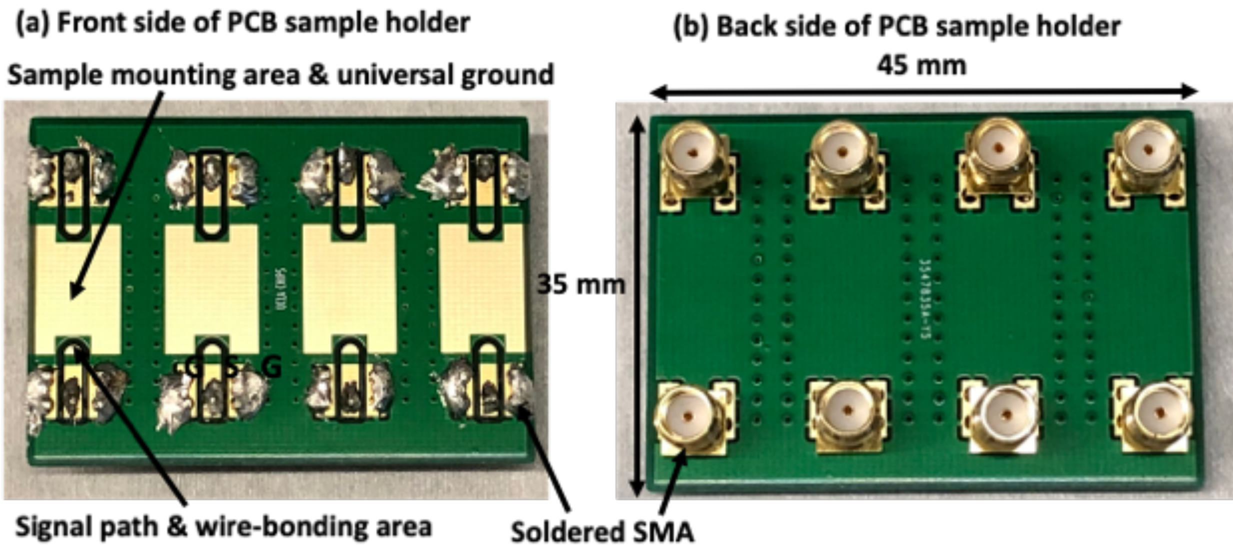


Figure 67. Two-layer FR-4 PCB carrier with a sample mounting area, signal path and wire-bonding 50 Ω traces, out-of-plane SMA soldering area, and an universal ground. Test die and IF design.

To design a simple two-layer FR-4 PCB (front side layer to house Si samples; part of front side and whole back side layer to function as a universal ground), the software called KiCAD PCB EDA is used. The design specs of the PCB carrier with 50 Ω coplanar waveguides with ground plane are as follows: (1) relative dielectric constant: 4.5, (2) loss tangent: 0.02, (3) Cu resistivity: 1.72 X 10^{-8}, (4) substrate thickness: 1.6 mm, (5) Cu film thickness: 0.035 mm, (6) signal width: 2.3 mm, and (7) gap between signal path and ground: 0.85 mm, as shown in Figure

66. The PCB carrier is designed with a sample mounting area, signal path and wire-bonding 50 Ω traces, SMA soldering area, and an universal ground. To fit into the cylinder shape of fridge chamber (diameter and height: ~60 mm), the SMA connectors (Amphenol RF, model: 132322) are designed to solder onto the PCB carrier (35 mm X 45 mm) in a out-of-plane direction, as shown in Figure 67.

6.2.3 Passive superconducting dielets and -IF deisgn and fabrication

To measure insertion loss and crosstalk, the test vehicles of GSG and GSSG on dielets and Superconducting-IF with the Au interlayer were designed. Three lengths of links (125 μm, 500 μm, and 1750 μm) with two L/S (2/2 and 5/5 μm) were designed to be cascaded on 2 X 2 mm^2 dielets and were tweaked to have a characteristic impedance of 50 Ω. For the crosstalk (Far-end crosstalk, FEXT, and Near-end crosstalk, NEXT) characterization, it requires two signal paths to be brought close. Due to the fridge I/O limit, the on-chip transmission line coupler was selected as the structure for characterization and was designed to have a 50 Ω impedence matching resistor (Ti/Pd, 3/75 nm) at 4K on one end of each signal path. Due to the fridge space limit, it cannot house a cryogenic RF calibration kit with switch [146] that can calibrate out background noise inside the low-temperature chamber. An on-chip de-embeded through stucture was also designed to filter out part of the background noises.

The fabrication process of the test vehicle is as follows: (1) 200 nm Nb/5 nm Ir consecutive sputter deposition on 500 nm thermal oxidized wafers, (2) patterning through dry etching, (3) Ti/Au interlayer lift-off process on designated bonding area, (4) 2×2 mm^2 dies and 10×10 mm^2 substrates after dicing, (5) surface ultrasonicating wet cleaning and surface Ar plasma activation, and (6) using flip-chip fine-pitch thermal compression bonding assembly, the cascaded GSG and

GSSG daisy chain structure are completed. Figure 68, Figure 69, and Figure 70 present the test vehicle for insertion loss and crosstalk characterization before flip-chip bonding.

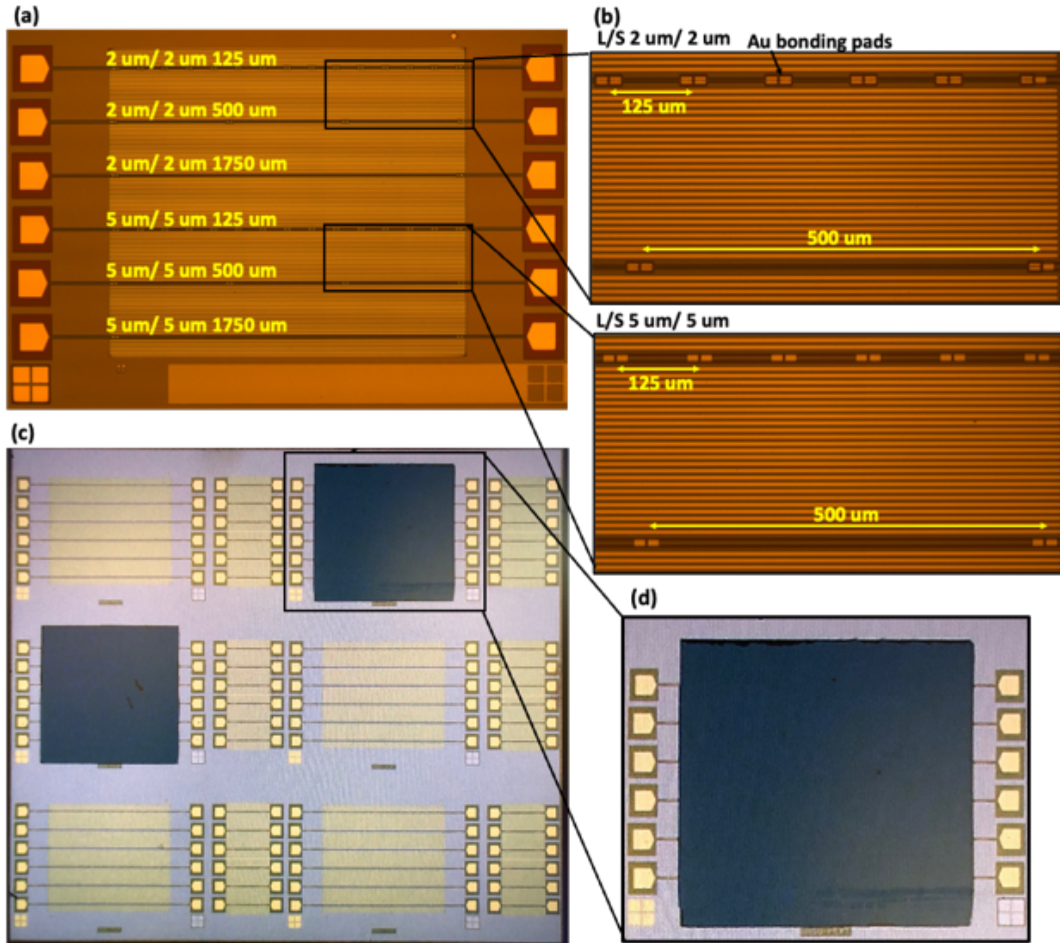


Figure 68. Insertion loss characterization: (a) three lengths of links (125 μm, 500 μm, and 1750 μm) with two L/S (2/2 and 5/5 μm) are cascaded on 2 X 2 mm^2 dielets, (b) zoom-in of (a), (c) flip-chip bonded multiple 2 X 2 mm^2 dielets, (d) zoom-in of (c).

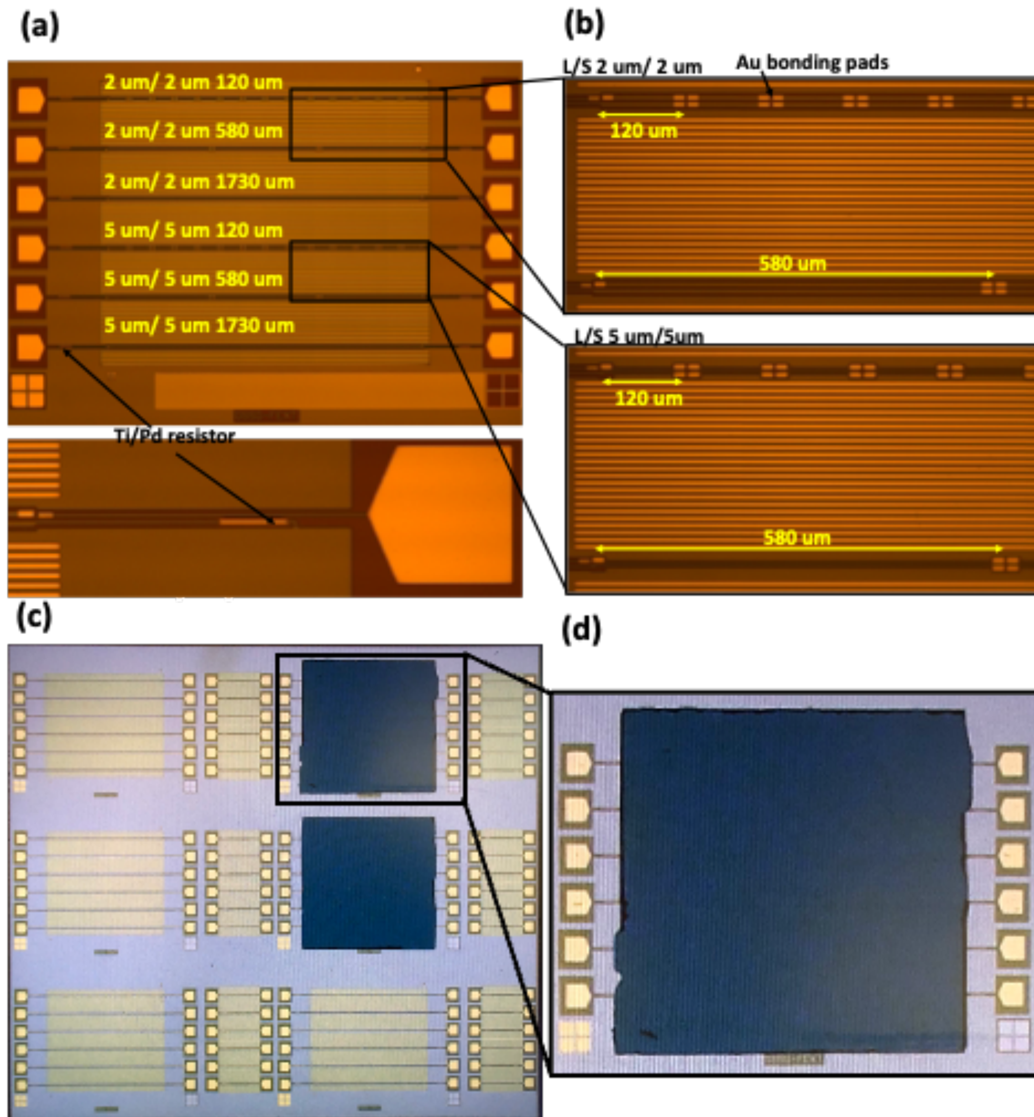


Figure 69. FEXT characterization: (a) three lengths of links (125 μm, 500 μm, and 1750 μm) with two L/S (2/2 and 5/5 μm) are cascaded on 2 X 2 mm^2 dielets, (b) zoom-in of (a), (c) flip-chip bonded multiple 2 X 2 mm^2 dielets, (d) zoom-in of (c).

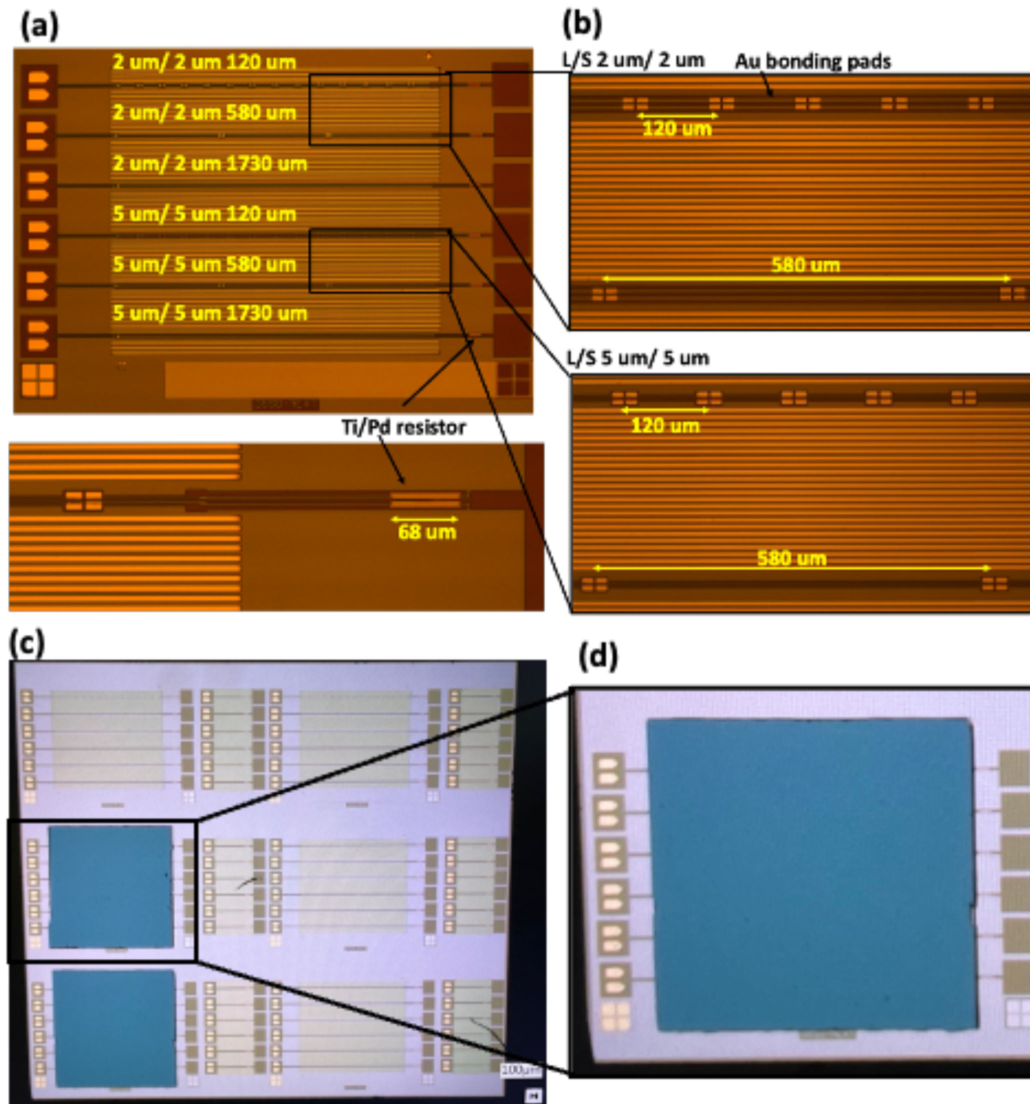


Figure 70. NEXT characterization: (a) three lengths of links (125 μm, 500 μm, and 1750 μm) with two L/S (2/2 and 5/5 μm) are cascaded on 2 X 2 mm^2 dielets, (b) zoom-in of (a), (c) flip-chip bonded multiple 2 X 2 mm^2 dielets, (d) zoom-in of (c).

As mentioned in the previous paragraph, the cryogenic impedence matching resistor (Ti/Pd, 3/75 nm) was also fabricated and tweaked during the design phase to match 50 Ω. Literature reports [37], [39], Ti/Pd is used for the qubit application with a thickness of 3/25 nm. Initially, this thickness was tested on a balnket wafer and the sheet resistance was measured, having a

sheet resistance distribution across the wafer at room temperature, as shown in Figure 71 (a). The wafer was then cleaved in to 10 X 10 mm^2 and was cooled down to 2 K. The resistance saturates below 25 K and has a value of 2.15 at 2 K as compared to 300 K, as shown in Figure 71 (b), using which the resistor is designed with a proper dimension. The ratio of resistance drop between 300K and 2 K is close to the value reported by the literature [147]. After the whole fabrication process was done, a testing de-embedded sample was measured at 2 K before flip-chip bonding but the pattern resistance was way higher than its original design. To remedy this issue, a thicker Pd film was used and the thickness of Pd increases from 25 nm to 50/75/100 nm. The bulk sheet resistance and the pattern ressitance are measured and the final thickness of Ti/Pd resistor is determined to be 3/75 nm with a pattern resistance of 50 Ω, as shown in Figure 72.

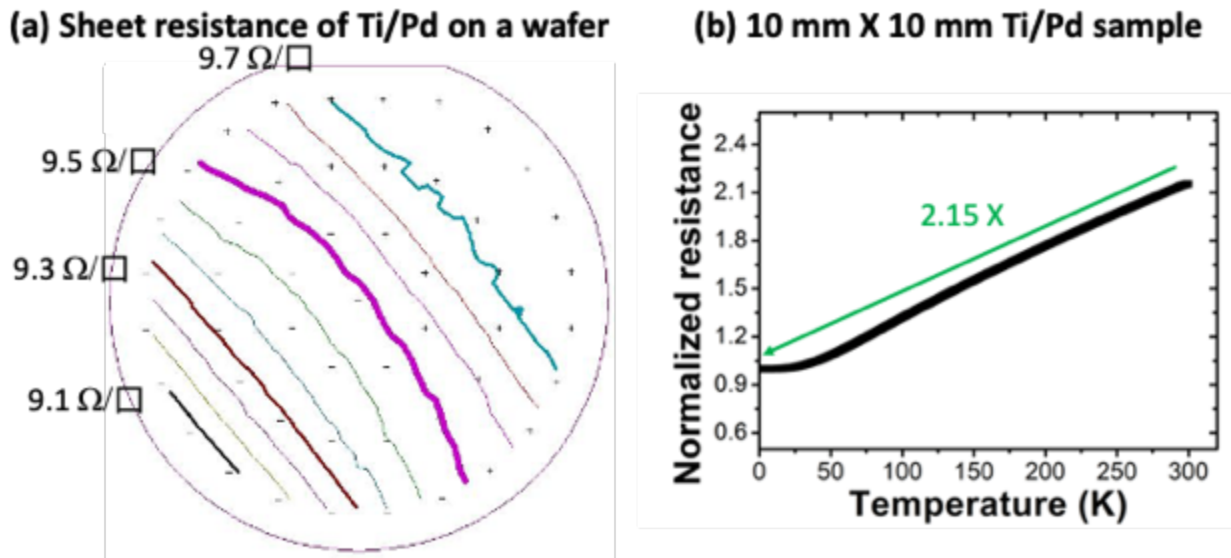


Figure 71. (a) the sheet resistance measurement on a blanket wafer with Ti/Pd 3/25 nm. (b) resistance drop as the temperature cools down to 2 K.

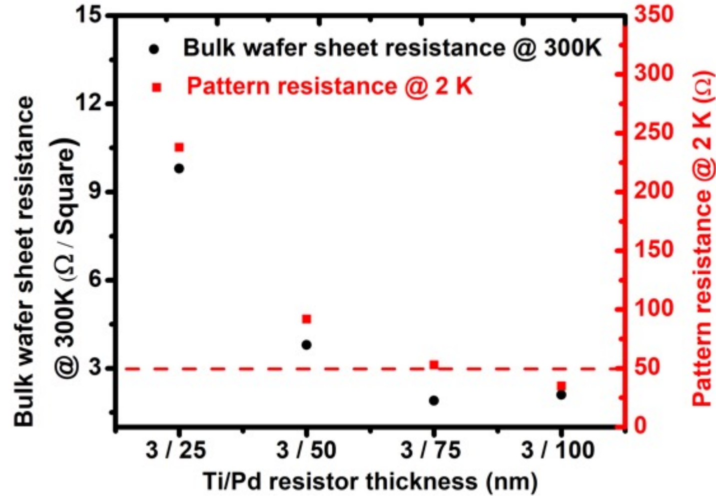


Figure 72. Tweaking of the pattern resistance of Ti/Pd with respect to different Pd thickness of 25/50/75/100 nm.

6.2.4 Pre-check and measurement results and discussion

In the preliminary DC electrical tests, results of insertion loss samples with L/S 5/5 μm, as shown in Figure 73(a), and of crosstalk samples with L/S 5/5 μm, as shown in Figure 73(b), show an expected curve both at 300K and at 4K. The samples with L/S 5/5 μm for insertion loss was also measured down to millikelvin region, as shown in Figure 74. The samples with L/S 2/2 μm showed a good result at 300K but was electrically open at 4K in all the link lengths. With a FIB-SEM investigation, as shown in Figure 75, a delamination issue is observed between the Ti and the Nb/Ir layer on the IF. This cryogenic adhesion problem was only observed in the links with L/S 2/2 μm, but not the links with L/S 5/5 μm, indicating a special method is required to increase the adhesion between Ti and Nb/Ir on small pads. In this section, only the measurement results of L/S 5/5 μm at 4 K is presented.

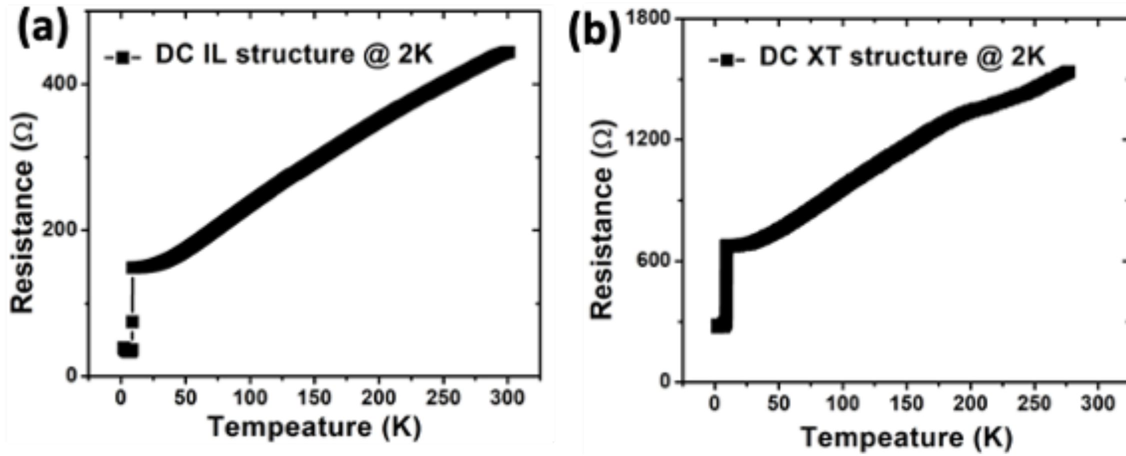


Figure 73. Preliminary DC electrical test result down to 2 K: (a) an insertion loss sample with Tc at 9 K and clear resistance drop below Tc, (b) a XT sample with Tc at 9K and clear resistance drop below Tc.

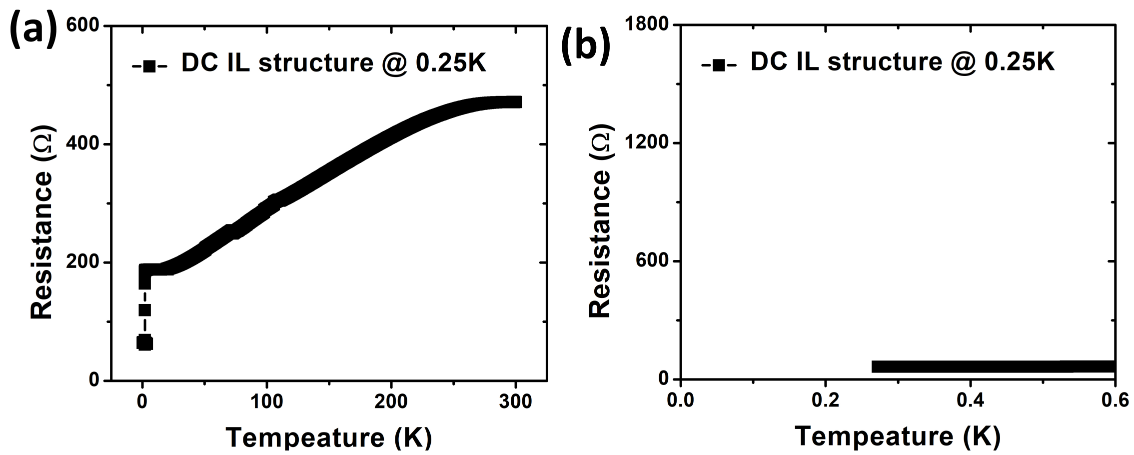


Figure 74. Preliminary DC electrical test result down to 0.25 K: (a) an insertion loss sample with Tc at 9K and clear resistance drop below Tc, (b) zoom-in of (a) between 0.25 K and 0.6 K.

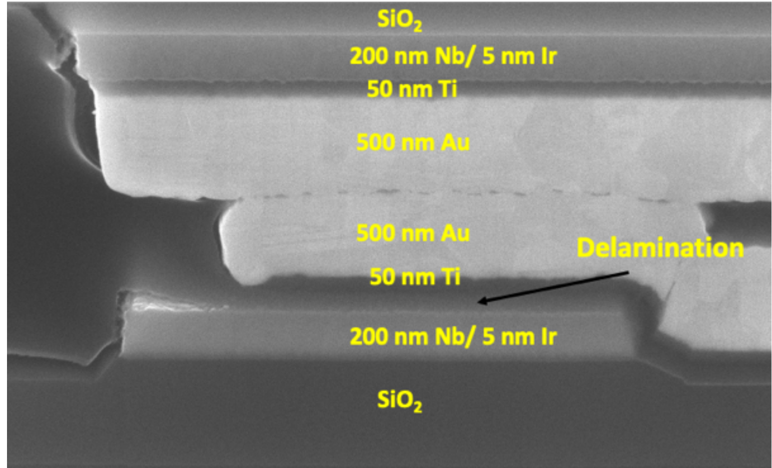


Figure 75. Image of the XT samples. A cryogenic adhesion problem is observed between the Ti film and the Nb/Ir film in the links with L/S 2/2 μm. © 2022 IEEE

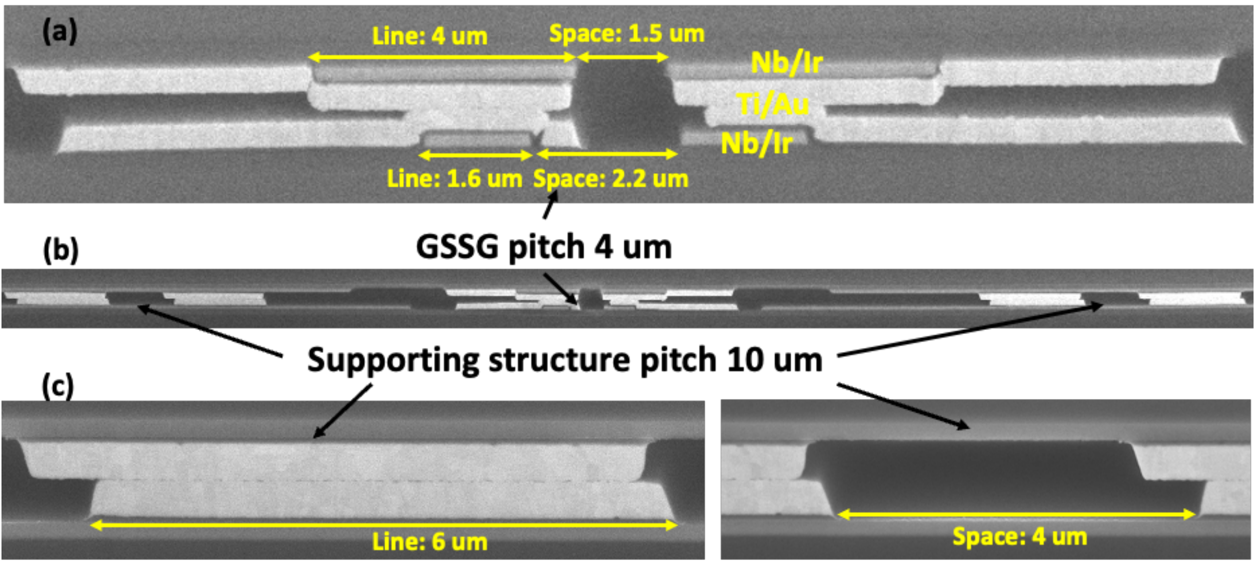


Figure 76. Image of the XT samples: (a) at the GSSG bonding site, (b) GSSG bonding site with the surrounding supporting structure, (c) zoom-in of the supporting bonding structure.

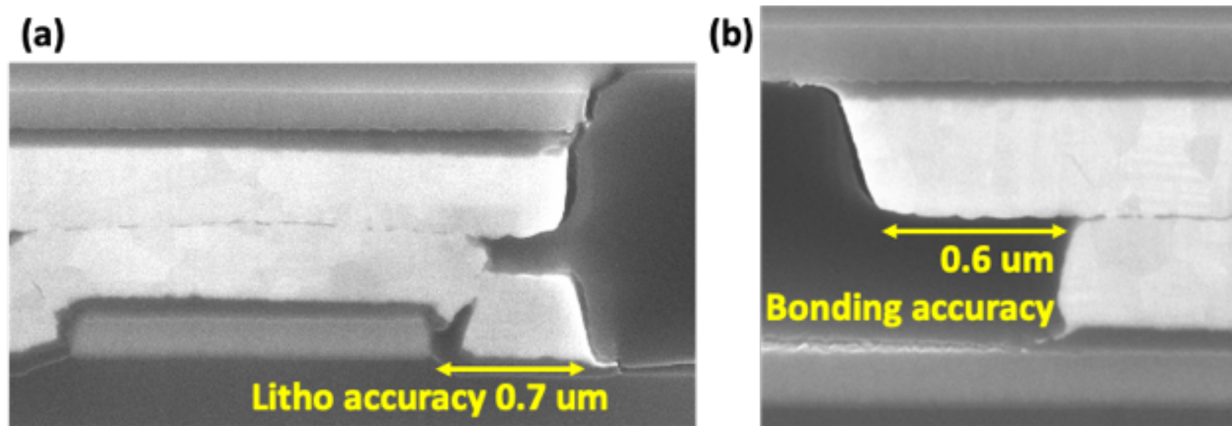


Figure 77. Image of the XT samples: (a) layer-to-layer lithography alignment overlay <1 μm, (b) assembly mechanical bonding overlay <1 μm.

In the following results, there exists a discrepancy between the measured and simulated results. This difference comes from two facts: (1) impedance mismatch, and (2) lack of cryogenic calibration before entering the Device-Under-Test (DUT) samples due to the fridge limit mentioned in the subsection 6.2.2. Although the test structure was designed to match 50 Ω, the fabrication and assembly process produced ±20% impedance, leading to mismatch of impedance and extra reflection loss. For the cryogenic calibration, due to the fridge space limit (a cylinder volume of 60-mm diameter with a height of 60 mm) at the mixing chamber, a cryogenic calibration kit [146] cannot fit into the chamber for cryogenic RF calibration. Instead, the RF calibration was conducted at the VNA side at room temperature with the tool kit Agilent 85052D using TRL calibration (Thru, Reflect, and line) before the fridge, not on the dielet at 4 K. The on-chip Thru structure was used as the de-embedded structure, but the loss from all adaptors, connections, and the coaxial cables was still partially included.

After the samples passed DC low-temperature test, they were adhered on the PCB carrier with CMRdirect GE Varnish for better cryogenic heatsinking, as shown in Figure 78, and was

installed on the fridge cylinder, as shown in Figure 79. After obtaining the raw data, de-embedding techniques was implemented through the MATLAB codes. The code logistics are as follows: (1) using the de-embedding code package in the RF toolbox, (2) importing the measured S parameter touchstone file in the de-embedded structure into the the de-embeding code, extracting the background noise in each SMA part assuming both sides are symmetric, (3) importing measured S parameter of cascaded DUTs with the background noise in each SMA part into MATLAB, (4) applying the the de-embedding code, extracting the S parameters of cascaded DUTs, (5) converting the S parameter of cascaded DUTs into the S parameter of a single DUT through ABCD metrics multiplication. Figure 80 shows the 4 K measurement of insertion loss with samples of L/S 5/5 μm and link lengthes of 125/500/1750 μm. The dash line is the simulation result, while the solid line is the measured results. For quantum applications, a low insertion loss value is defined as less than 1 dB loss [148]. All the simulated and measured results are within 1 dB loss in three link lengths, meaning the ultra-short, short and long links in a compact assembly on Superconducting-IF are suitable to carry quantum information and SFQ signals with low loss.

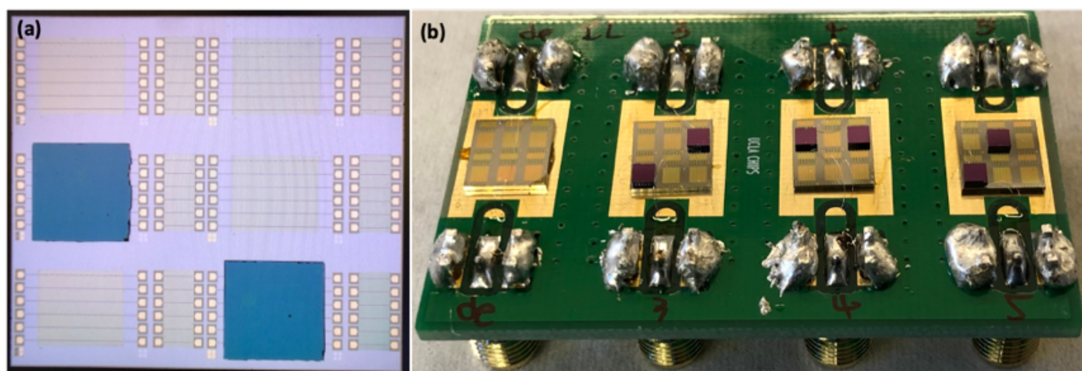


Figure 78. (a) image of flip-chip bonded samples, (b) image of flip-chip bonded samples adhered on PCB and with wire-bonding. © 2022 IEEE

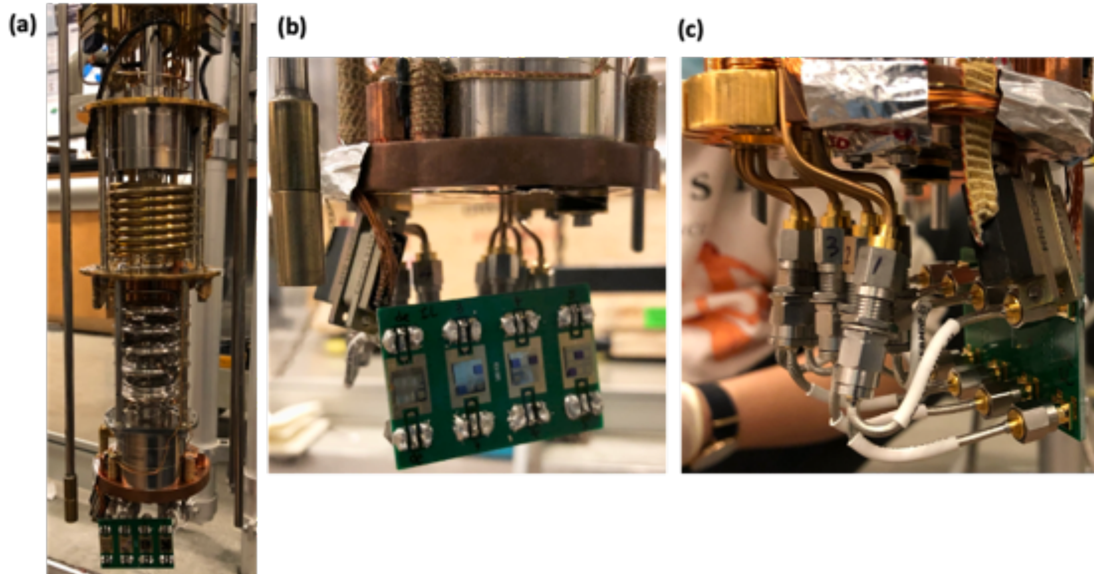


Figure 79. (a) PCB carrier installed on the fridge cylinder, (b) front side image. (c) back side image. © 2022 IEEE

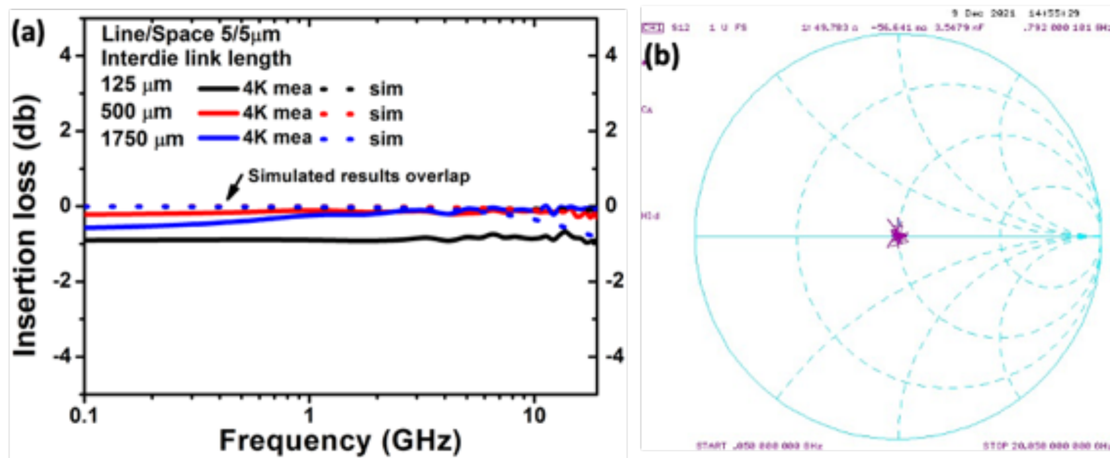


Figure 80. (a) 4 K measurement of insertion loss with samples of L/S 5/5 μm and link lengthes of 125/500/1750 μm. © 2022 IEEE (b) impedence mismatching in the samples due to fabrication process.

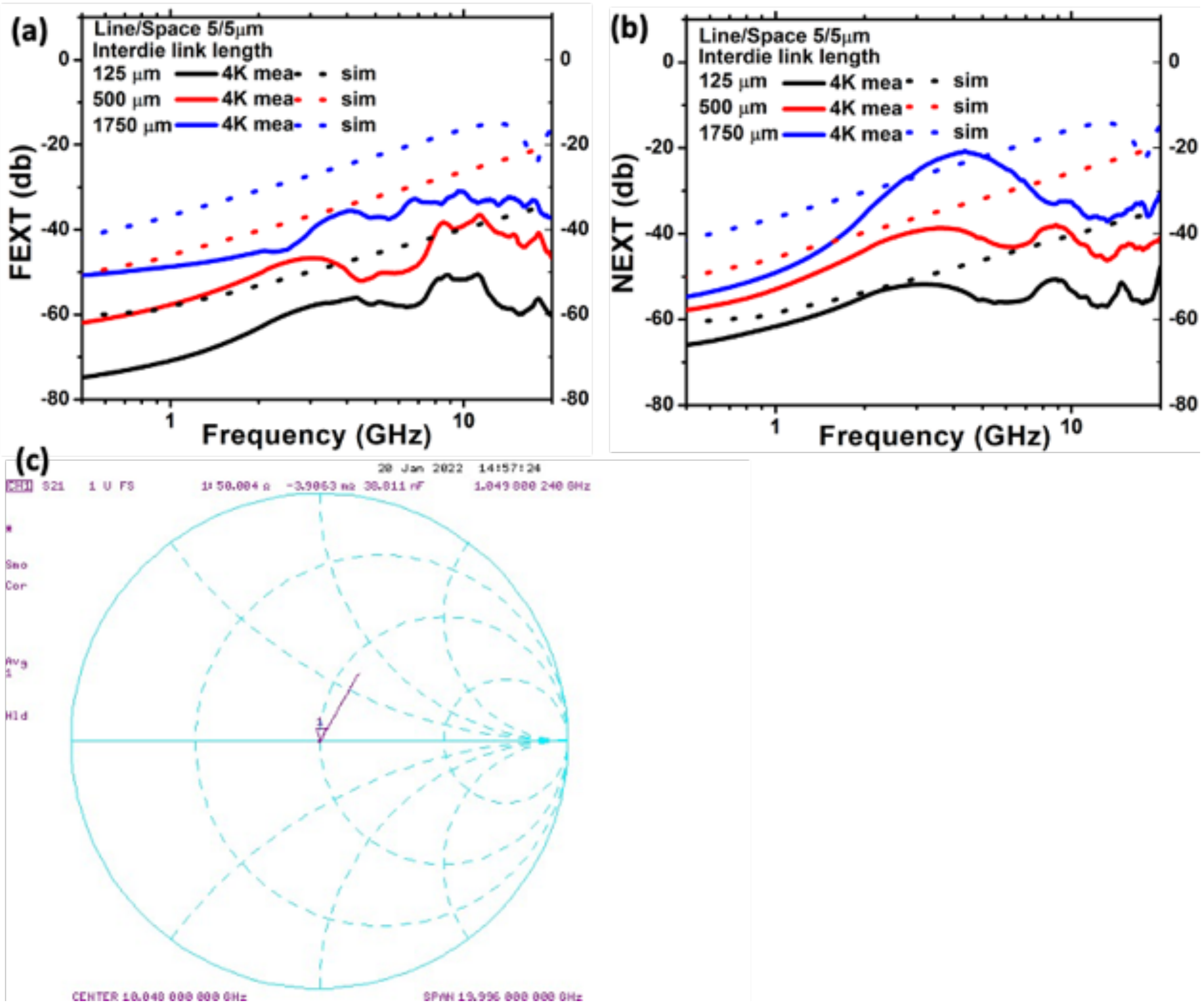


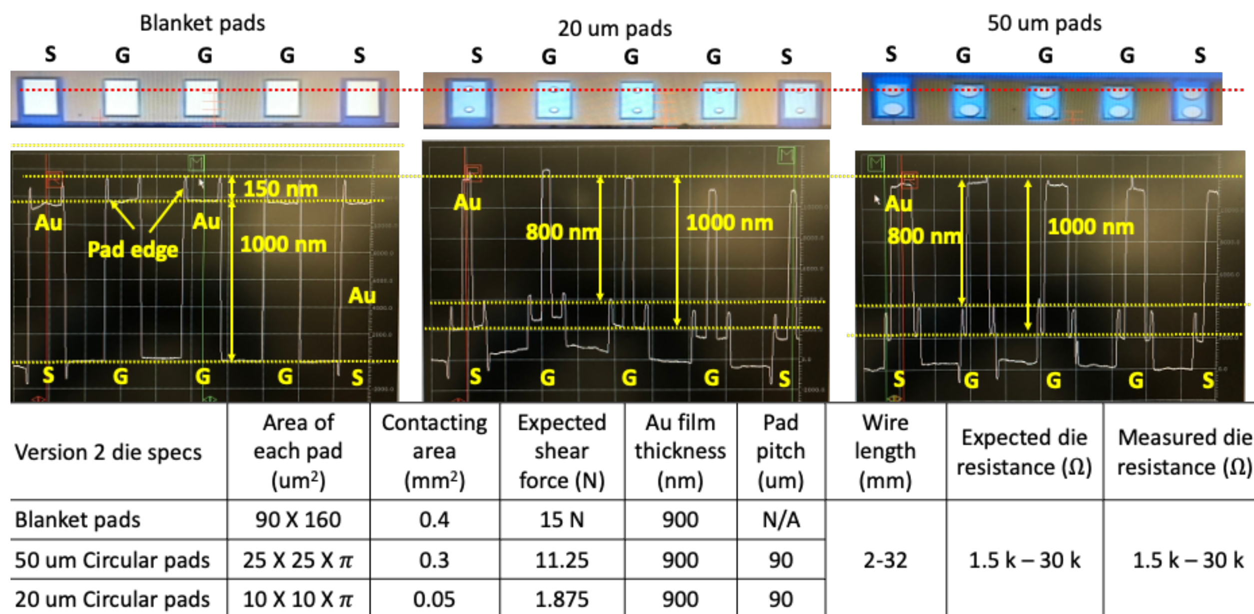
Figure 81. 4 K measurement of (a) FEXT (b) NEXT with samples of L/S 5/5 μm and link lengthes of 125/500/1750 μm. (c) Impedence mismatching due to fabrication process. © 2022 IEEE

Figure 81 shows the 4K measurement of crosstalk with samples of L/S 5/5 μm and link lengthes of 125/500/1750 μm. The dash line is the simulation result, while the solid line is the measured results. The crosstalk values for the link length of 125, 500, and 1750 μm from 0.5 GHz to 20 GHz are -75 to -35 dB, -62 to -38 dB, and -50 to -30 dB, respectively. The value for acceptable crosstalk is that as compared to the thermal Johnson noise from resistors, the value of

crosstalk should not be larger than the thermal noise [149] and should not become the major source of noises. Resistors with a large resistance value in superconducting electronics are mainly used for current biasing (43 Ω) and power on superconducting circuits [16]. As one example in [16], the thermal noise from a 43 Ω resistor at 4 K with a bandwidth of 200 GHz causes -23 dB under the condition a single on-chip SFQ pulse is 300 μV. As shown in Figure 81, in the short links (≤500 μm) and almost the long link at low frequency, the simulated and measured results regarding the FEXT and NEXT are below -23 dB. This concludes that the short and long links links from proximal integration of dielets on Superconducting-IF are suitable for inter-dielet communication.

6.3 Passive die with resonator integrated on Superconducting-IF

Part of this project collaborates with SeeQC, through which three types of passive dies were shipped to UCLA. These three types of passive dies was designed with 20 Ω impedance and included passive transmission line resonators between 10-20 GHz with blanket Au bonding pads, 20 μm cylinder Au bonding pads, and 50 μm circular Au bonding pads. The detailed die information shows in Figure 82 with a Au film of 1000 nm on the bonding pads and a root-mean-square surface roughness (Rq) of 6.69 nm.



Version 2 die specs	Area of each pad (um^2)	Contacting area (mm^2)	Expected shear force (N)	Au film thickness (nm)	Pad pitch (um)	Wire length (mm)	Expected die resistance (Ω)	Measured die resistance (Ω)
Blanket pads	90 X 160	0.4	15 N	900	N/A	2-32	1.5 k – 30 k	1.5 k – 30 k
50 um Circular pads	25 X 25 X π	0.3	11.25	900	90			
20 um Circular pads	10 X 10 X π	0.05	1.875	900	90			

Figure 82. Detailed die information with blanket Au bonding pads, 20 μm cylinder Au bonding pads, and 50 μm circular Au bonding pads.

Since at the edge of blanket Au pad there exists Au protruding of 150 nm, to compensate this surface height difference, a thick Au film is used to provide around 15% deformation of the overall film thickness [150], [151], allowing the whole blanket Au pad to contact its mating pad. To match the dies with 20Ω impedance to the measurement tool with 50Ω impedance through the

Superconducting-IF, a quarter-wavelength-transmission-line impedance transformer was designed and fabricated on the Superconducting-IF. Among all the types of the transmission line, ungrounded coplanar waveguide was selected for processing with less challenges.

The quarter-wavelength-coplanar-waveguide impedance transformer has a characteristic impedance value equal to square root of the impedance of dies multiplying the impedance of the measurement tool. Through calculation, the impedance transformer should have a characteristic impedance of 31.6 Ω. The ungrounded coplanar waveguide sits on a substrate having 500 nm SiO_2 and 500 μm Si, leading to an effective dielectric constant of 9.36 through calculation [152]. Using the ANSYS HFSS solver, with a 31.6 Ω ungrounded coplanar waveguide and, the width of the center signal path and the gap between the signal path and the ground plane are simulated to be 28 μm and 2 μm respectively, as shown in Figure 83.

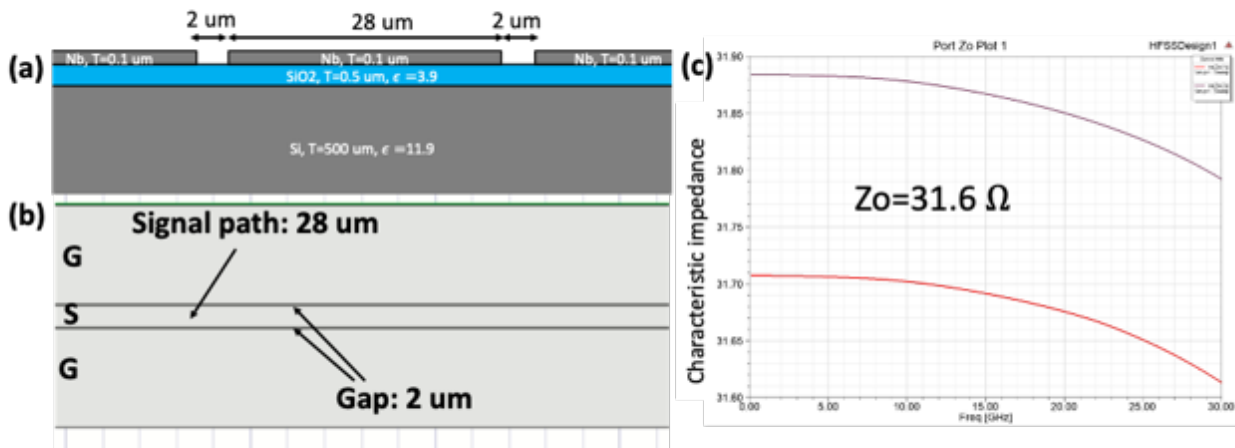


Figure 83. (a) Cross section view for a 31.6 Ω ungrounded coplanar waveguide. The width of the center signal path and the gap between the signal path and the ground plane are 28 μm and 2 μm. (b) Top view of the ungrounded coplanar waveguide in a GSG configuration. (c) Simulation result shows a characteristic impedance of around 31.6 Ω. Each curve indicates the impedance in each port.

To have the on-chip transmission line resonator resonating between 10-20 GHz, the quarter-wavelength-coplanar-waveguide impedance transformer should also resonate at the same frequency range. The phase speed equals to the light speed divided by square root of half of the effective dielectric constant adding one ($\upsilon = \frac{c}{\sqrt{\frac{(9.36+1)}{2}}}$). This corresponding to a quarter wavelength of 2533 μm. The design is presented in Figure 84 with 20 Ω termination in one side and 50 Ω termination in the other side. Since the designed dimension of the coplanar waveguide can vary during the fabrication process, the dimension of the center signal path and the gap between the signal path and the ground plane are changed advertently in the model to investigate a proper range of dimensions with the desired resonance. With the dimension of center signal path ranging in 24-30 μm and the gap ranging in 1.8-2.6 μm, the resonance is preserved and the simulation results of the S parameter are presented in Figure 85 and Figure 86.

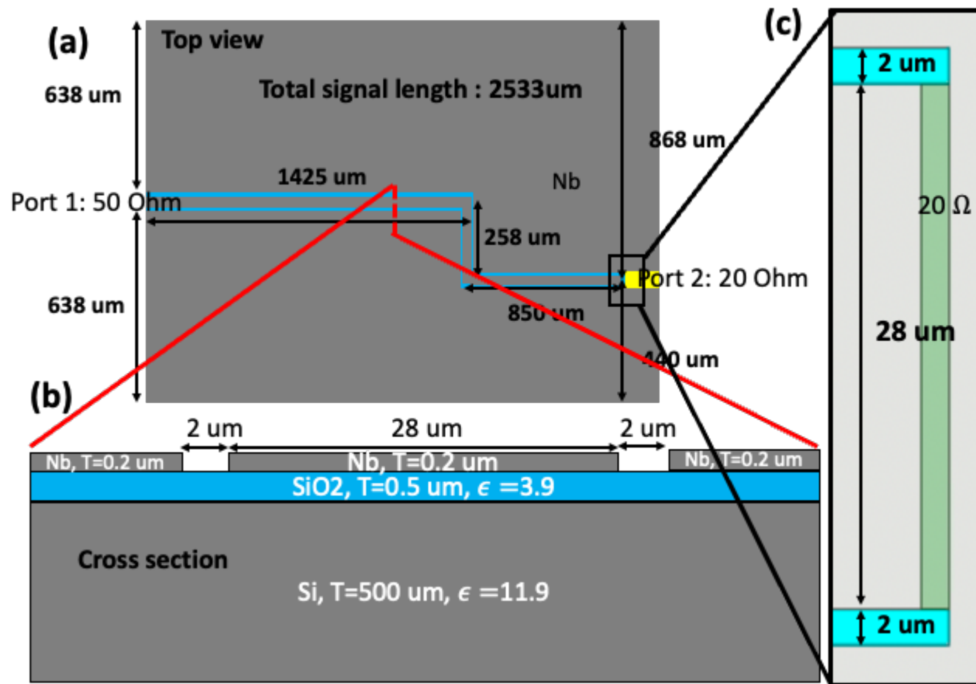


Figure 84. (a) Top view of the design quarter-wavelength-coplanar-waveguide impedance transformer on the Superconducting-IF. (b) Cross section view of the coplanar-waveguide with a 28-μm center signal path and a 2-μm gap between the signal path and the ground plane.

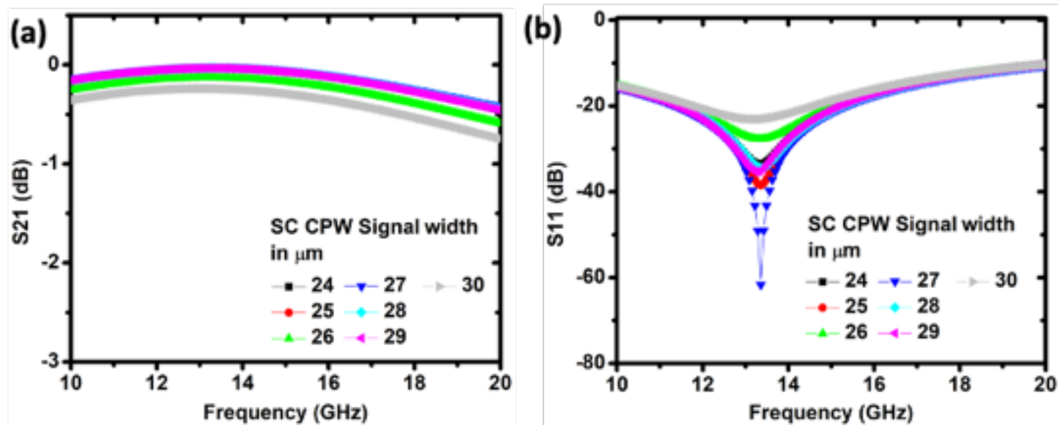


Figure 85. Simulation results of S parameter with the varying signal width from 24 to 30 μm. The resonance is at 13 GHz.

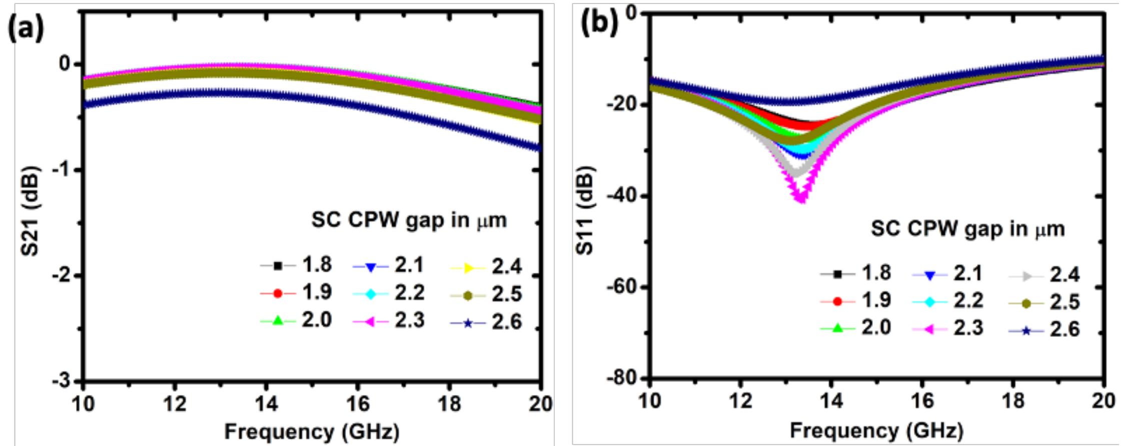


Figure 86. Simulation results of S parameter with varying gap width from 1.8 to 2.6 μm. The resonance is at 13 GHz.

Following the simulation results and the well-established fabrication process, the fabricated Superconducting-IF, 5 X 5 mm^2 die, and flip-chip bonded Superconducting-IF are presented in Figure 87. The dimension of the quarter-wavelength-coplanar-waveguide impedance transformers on Superconducting-IF is double-checked by SEM image, as shown in Figure 88. The image shows the fabricated dimension is within the simulated range.

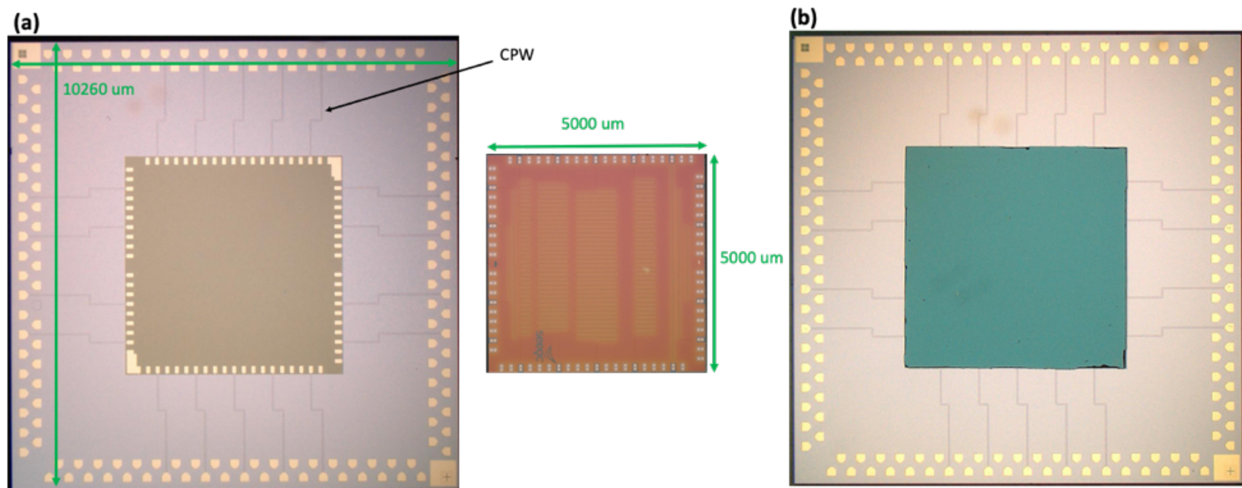


Figure 87. (a) Fabricated Superconducting-IF with quarter-wavelength-coplanar-waveguide impedance transformers. (b) SeeQC dies. (c) flip-chip bonded Superconducting-IF.

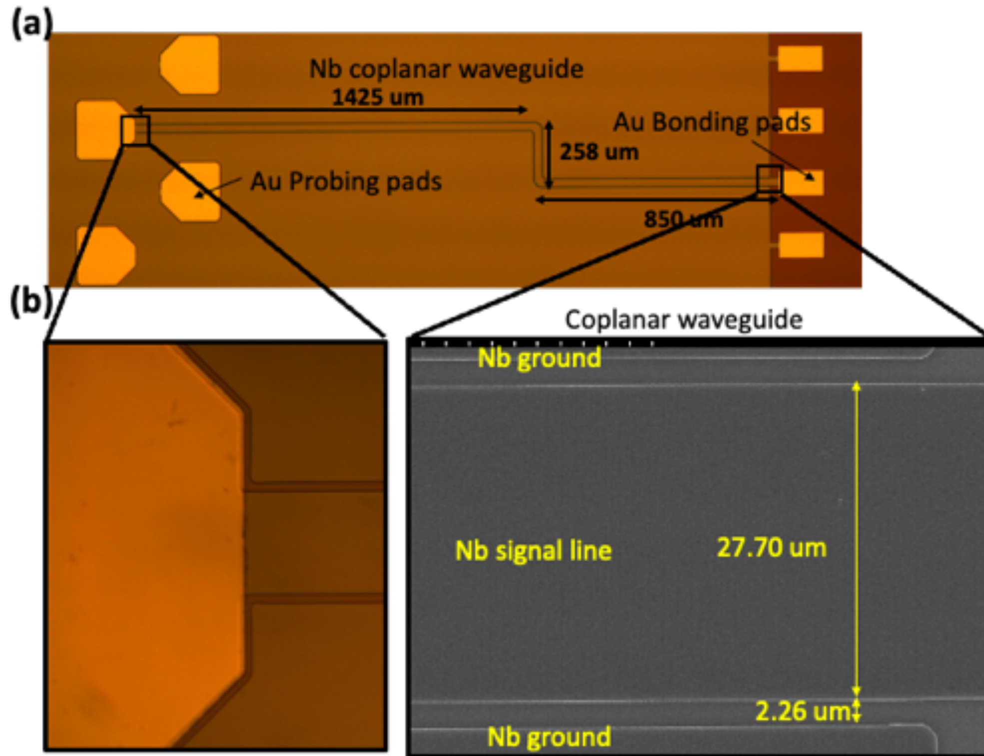
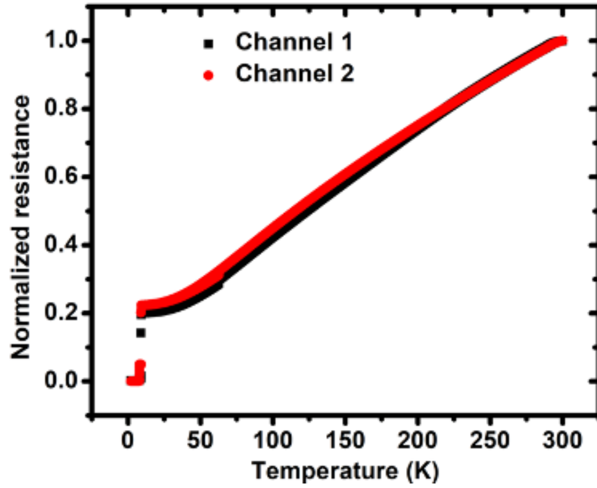


Figure 88. (a) Optical and (b) SEM images of the quarter-wavelength-coplanar-waveguide impedance transformers on Superconducting-IF.

The DC measurements of flip-chip-bonded Superconducting-IF with blanket Au pads and 20-μm Au pads down to 2 K are implemented as a pre-RF-measurement check and the results are presented in Figure 89 and Figure 90. In both cases, the superconductivity exists in all the samples at 9K with smooth curves. Unfortunately, when the bonded samples are shipped back to SeeQC, although the dimension of the Superconducting-IF has been double-checked ahead with the company in the design phase, the size of the sample cannot fit into the holder for measurements. Therefore, there is only simulation results without measurement results in this subsection.



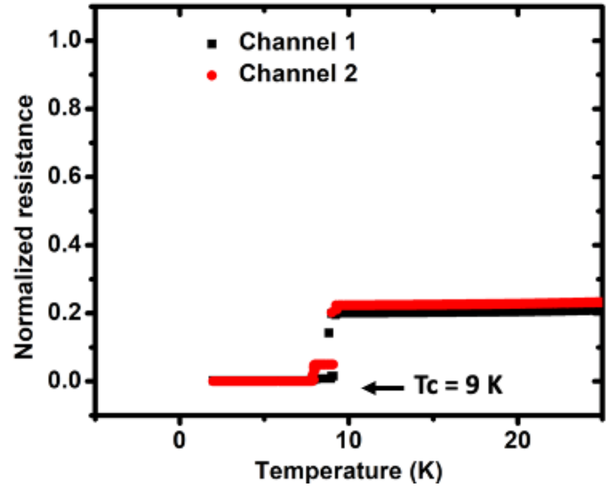
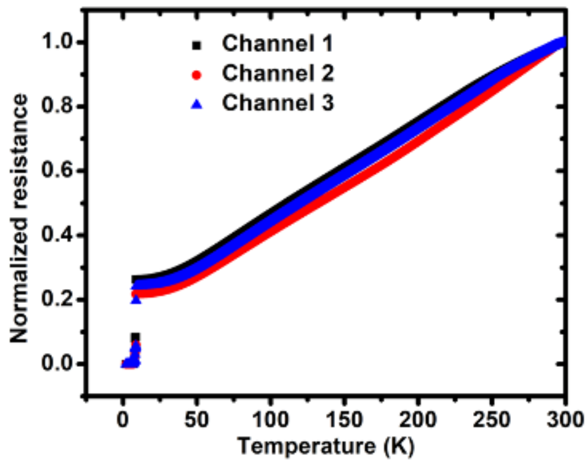


Figure 89. DC measurements of flip-chip-bonded Superconducting-IF with blanket Au pads. Tc exists in all the samples at 9K with smooth curves



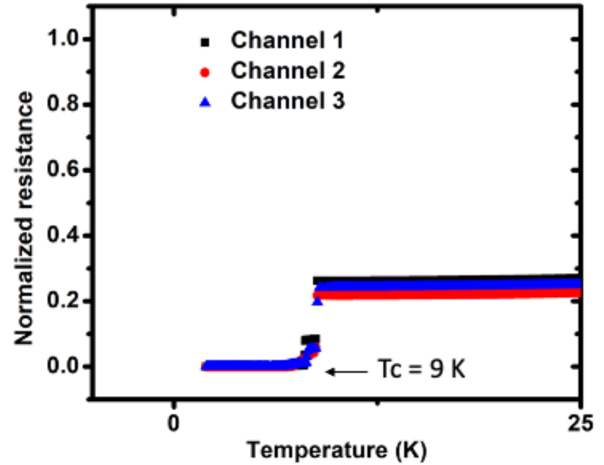


Figure 90. DC measurements of flip-chip-bonded Superconducting-IF with 20-um Au pads. Tc exists in all the samples at 9K with smooth curves

6.4 Summary

In this chapter, to realize the proposed concept of wafer-level integration on Superconducting-IF for large-scale quantum computing with qubit array and periphery JJ-based superconducting control/readout dies, superconducting RF characterization for inter-dielet communication is implement. Before any real design and fabrication, RF simulation with different link lengths (short links, ≤ 500 µm, and long links, 1750 µm) and L/S (2/2 µm, and 5/5 µm) was first carried out up to 20 GHz. Simulation assumption, approximation, model establishment, and analysis are presented. Ahead of characterization of RF behavior on testing samples, dilution fridge apparatus and PCB sample holders were set up. Next, test vehicles were designed and fabricated to characterize insertion loss and crosstalk on the Superconducting-IF. During DC testing, a cryogenic adhesion problem is observed and is measured electrically open on dies with a L/S of 2/2 µm. Only dies with a L/S of 5/5 µm were measured at 4K. All the simulated and measured results regarding insertion loss are within 1 dB loss in three link lengths up to 20 GHz, meaning the ultra-short, short and long links in a compact assembly on Superconducting-IF are suitable to carry quantum information and SFQ signals. All the simulated and measured results regarding crosstalk are lower than on-chip thermal noise and the values are almost below -23 dB up to 20 GHz, indicating that the short and long links links from proximal integration of dielets on Superconducting-IF are suitable for inter-dielet communication. At the end of the chapter, to have SeeQC dies with 20 Ω impedance match to measurement tools with 50 Ω impedance, simulations were implemented on quarter-wavelength-coplanar-waveguide impedance transformers resonating at 13 GHz. Since the designed dimension of the coplanar waveguide can vary during the fabrication process, the dimensions of the coplanar waveguide are changed

intentionally in the model. With the dimension of the center signal path ranging in 24-30 μm and the gap ranging in 1.8-2.6 μm, the resonance is preserved. The SeeQC dies are integrated on Superconducting-IF and are pre-tested successfully but due to the company tool issue, only simulation results are presented.

Chapter 7

Conclusions

7.1 Conclusions

This dissertation introduces the mechanism of quantum computing and various physical vehicles to potentially realize a large-scale quantum computing. This thesis points out the principal difference between a transistor computing and a qubit computing and organizes the hardware challenges, especially the I/O, latency and thermalization, to build a large-scale quantum computer from the viewpoint of electronic packaging. The objective of this dissertation is to minimize the challenges through cryogenic heterogeneous integration and advanced packaging approaches.

In Chapter 2, four types of architectures for quantum computing are presented: the room-temperature-CMOS architecture, the CryoCMOS architectures of Google, Microsoft, Intel, the 3D-stacking architecture through superconducting TSV from MIT, and the JJ-based architecture featuring ultra-low-power-dissipation that can integrate with qubits. To further investigate the JJ-based architecture, various types of JJ-based circuits and their power dissipation are presented.

In Chapter 3, the packaging technologies for room-temperature applications, including traditional PCB assembly, advanced assembly through flip-chip heterogeneous integration and Silicon Interconnect Fabric, are overviewed. The packaging technologies for cryogenic applications, including wire-bonding and flip-chip, and their concerns for the large-scale quantum computing are introduced.

In Chapter 4, the cryogenic version of the Silicon Interconnect Fabric- Superconducting-IF is introduced and it will inherit all the features of Si-IF. The detailed concept of quantum system

integration on the Superconducting-IF from the aspects of thermal limits of qubit chips (< 500 nW) and the deep cryogenic chamber (< 20 μW) is described. The comparison between the proposed concept in this thesis and the existing schemes is presented. Through the development of flexible cables and the superconducting multiplexing, controlling over one million qubit arrays is possible using the quantum system integration on the Superconducting-IF.

In Chapter 5, the logistics and the experiments to use the Au interlayer for quantum applications are presented. The Au interlayer bonding technology is demonstrated to be fine-pitch, mechanically robust, and electrically reliable, and has a quantum compatible process with a critical current value 7-8 X higher than the required current to drive functional superconducting chips for both processing and memory. On the Superconducting-IF platform, heterogeneity and flexibility are both demonstrated regarding the die sizes, the distance between each flip-bonded die, and configuration.

In Chapter 6, to realize the proposed concept of wafer-level integration on Superconducting-IF for large-scale quantum computing with qubit array and periphery JJ-based superconducting control/readout dies, superconducting RF characterization, including insertion loss and crosstalk for inter-dielet communication is implemented. All the simulated and measured results regarding insertion loss and crosstalk are below the threshold, meaning the superconducting interconnects are suitable to carry future delicate quantum information.

In conclusion, to scale up the qubit count to one million and a large-scale integration, the proposed integrated architecture on the Superconducting Silicon Interconnect Fabric is one alternative approach besides the CryoCMOS methods and our proposed approach can potentially unsheathe the full strength of quantum computing.

This dissertation achieves the following intellectual contributions: (1) among the first to

combine heterogeneous integration and advanced packaging with quantum applications, (2) the first work to propose the architecture of integrated system-on-wafer quantum hardware, (3) demonstration of the Au interlayer bonding to be fine-I/O pitch, mechanically robust, electrically reliable, and quantum compatible, and (4) the first work to achieve low-loss, low-crosstalk, and 20GHz-broadband RF capability on superconducting short links through advanced packaging.

7.2 Outlook

The superconducting version of Si-IF technology is a superior alternative to realize a large-scale heterogeneous quantum computing as compared to PCBs. Although the potential advantages of the proposed scheme that uses fine-pitch integration on the superconducting Silicon Interconnect Fabric are significant, several challenges remain and four major directions for the future are suggested below:

1. Bonding material that can achieve fine-pitch applications, can be superconducting and can have quantum-compatible fabrication process (<150 ℃):

   For quanutm computing, a lower resistance is always a better way. This is the reason indium bumps are adopted for its superconductivity at millikevin. Indium bumps, however, cannot achieve fine-pitch applications for a future compact quantum system due to solder extrution during flip-chip bonding process. Therefore, the low-temperature Au interlayer bonding technology is developed to achieve 10-µm I/O pitch and its resistance has been optimized to reach an excellent specific contact resistance in the order of $10^{-8}\ \Omega - \text{cm}^2$, which is the limitation to this material because Au is not a superconductor and the resistance of Au does not drop abruptly at certain temperature. The Au interlayer bonding technology demonstrated in this thesis is the first phase of proof of concept. The second phase of proof of concept can explore different material, such as Pb and Al, to achieve superconducting fine-pitch applications and with quantum-compatible fabrication process (<150 ℃)

2. Qubits and superconducting control/readout circuits co-integration on superconducting-IF:

   Designing one qubit and relevant superconducting control/readout circuits and integrating them on a substrate has been demonstrated by University of Wisconsin-Madison/Syracuse

University in 2018/2019. The qubit fidelity can achieve at least 95%, which is a good fidelity. Through lab collaboration, after securing the qubit die and relevant superconducting control/readout die, integrating all of them using the Au interlayer bonding on the Superconducting-IF with a fine I/O pitch can be a good starter to demonstrate the proposed architecture in Chapter 4.

3. Superconducting integrated flexible cables that can replace stand-alone coaxial cables and corresponding cryogenic miniaturized connectors:

   At Auburn university, it has been demonstrated that multiple superconducting wires are fabricated and integrated compactly on flexible polyimide substrates[153], which can potentially replace multiple coaxial cables connecting from room-temperature electronics to the deep cryogenic chamber. The superconducting flexible cables are anchored on Molybdenum through their cryogenic connectors[154]. In this way, each wire on average dissipates less heat, meaning more wires can be installed inside the fridge chamber and allowing more qubits to be integrated. These technologies can be applied on the Superconducting-IF for further increase of qubits and decrease the chamber heat load.

4. Development of superconducting multiplexing:

   Superconducting multiplexing technique enables multiple qubits control with one wire. State-of-the-art superconducting multiplexing can achieve 1-to-8 multiplexing or 8-bit multiplexer [100]. The bottleneck is the phase matching and the power on each wire after multiplexed. The phase matching issue can possibly be resolved through a careful design, while the power can be amplified through Josephson parametric amplifier or on-chip integrated circulators. Through these methods, 16 bits of superconducting multiplexer should

be possible to attain.

5. Development of high Tc superconductor:

   As compared to the low-temperature superconductor, Nb, used in this thesis, high Tc superconductor, such as Yttrium barium copper oxide (YBCO), can possibly come into play on the platform of the superconducting silicon interconnect fabric for high conductivity and large-scale quantum computing applications. Although thermal noise and other noise will increase due to a higher chamber temperature at 77 K after adopting high Tc superconductor as superconducting interconnects, the available cooling power of the fridge, however, can increase to a couple of watts, instead of a few microwatts. This significant increase in the cooling power allows housing more coaxial cables and relevant components inside the fridge for independent qubit controls/readout and is an alternative approach to scale up the qubit count.

References

[1] R. N. Das *et al.*, "Large Scale Cryogenic Integration Approach for Superconducting High-Performance Computing," in *2017 IEEE 67th Electronic Components and Technology Conference (ECTC)*, 2017, pp. 675–683. doi: 10.1109/ECTC.2017.54.

[2] S. K. Tolpygo, "Superconductor digital electronics: Scalability and energy efficiency issues (Review Article)," *Low Temperature Physics*, vol. 42, no. 5, pp. 361–379, May 2016, doi: 10.1063/1.4948618.

[3] T. D. Ladd, F. Jelezko, R. Laflamme, Y. Nakamura, C. Monroe, and J. L. O'Brien, "Quantum computers," *Nature*, vol. 464, no. 7285, pp. 45–53, 2010, doi: 10.1038/nature08812.

[4] E. Knill, R. Laflamme, and G. J. Milburn, "A scheme for efficient quantum computation with linear optics," *Nature*, vol. 409, no. 6816, pp. 46–52, 2001, doi: 10.1038/35051009.

[5] D. J. Wineland, C. Monroe, W. M. Itano, D. Leibfried, B. E. King, and D. M. Meekhof, "Experimental Issues in Coherent Quantum-State Manipulation of Trapped Atomic Ions," *J Res Natl Inst Stand Technol*, vol. 103, no. 3, pp. 259–328, 1998, doi: 10.6028/jres.103.019.

[6] S. Olmschenk, D. N. Matsukevich, P. Maunz, D. Hayes, D. L.-M., and C. Monroe, "Quantum Teleportation Between Distant Matter Qubits," *Science (1979)*, vol. 323, no. 5913, pp. 486–489, Jan. 2009, doi: 10.1126/science.1167209.

[7] O. Morsch and M. Oberthaler, "Dynamics of Bose-Einstein condensates in optical lattices," *Rev Mod Phys*, vol. 78, no. 1, pp. 179–215, Feb. 2006, doi: 10.1103/RevModPhys.78.179.

[8] D. J. Reilly, "Challenges in Scaling-up the Control Interface of a Quantum Computer," in

2019 IEEE International Electron Devices Meeting (IEDM), 2019, pp. 31.7.1-31.7.6. doi: 10.1109/IEDM19573.2019.8993497.

[9] B. S. Landman and R. L. Russo, "On a Pin Versus Block Relationship For Partitions of Logic Graphs," *IEEE Transactions on Computers*, vol. C–20, no. 12, pp. 1469–1479, 1971, doi: 10.1109/T-C.1971.223159.

[10] M. A. Nielsen and I. L. Chuang, "Quantum computation and quantum information," *Phys. Today*, vol. 54, no. 2, p. 60, 2001.

[11] A. J. Salim, "SFQ-enabled Scalable Digital Quantum Computing," *Electronic Components and Technology Conference*, Jun. 2020.

[12] F. Arute *et al.*, "Quantum supremacy using a programmable superconducting processor," *Nature*, vol. 574, no. 7779, pp. 505–510, 2019, doi: 10.1038/s41586-019-1666-5.

[13] S. Krinner *et al.*, "Engineering cryogenic setups for 100-qubit scale superconducting circuit systems," *EPJ Quantum Technol*, vol. 6, no. 1, Dec. 2019, doi: 10.1140/epjqt/s40507-019-0072-0.

[14] B. Abdo *et al.*, "Active protection of a superconducting qubit with an interferometric Josephson isolator," *Nat Commun*, vol. 10, no. 1, Dec. 2019, doi: 10.1038/s41467-019-11101-3.

[15] axon cable & interconnect, "Coaxial cables." https://www.axon-cable.com/publications/Coaxial-cables.pdf (accessed Mar. 02, 2022).

[16] A. M. Savin, J. P. Pekola, D. v Averin, and V. K. Semenov, "Thermal budget of superconducting digital circuits at subkelvin temperatures," *J Appl Phys*, vol. 99, no. 8, p. 084501, Apr. 2006, doi: 10.1063/1.2187276.

[17] Y. Salathé *et al.*, "Low-Latency Digital Signal Processing for Feedback and Feedforward

in Quantum Computing and Communication," *Phys Rev Appl*, vol. 9, no. 3, p. 34011, Mar. 2018, doi: 10.1103/PhysRevApplied.9.034011.

[18] D. T. McClure, H. Paik, L. S. Bishop, M. Steffen, J. M. Chow, and J. M. Gambetta, "Rapid Driven Reset of a Qubit Readout Resonator," Mar. 2015, doi: 10.1103/physrevapplied.5.011001.

[19] C. C. Bultink *et al.*, "Active resonator reset in the nonlinear dispersive regime of circuit QED," Apr. 2016, doi: 10.1103/physrevapplied.6.034008.

[20] J. C. Bardin, "A Low-Power CMOS Quantum Controller for Transmon Qubits," in *2020 IEEE International Electron Devices Meeting (IEDM)*, 2020, pp. 25.5.1-25.5.4. doi: 10.1109/IEDM13553.2020.9372108.

[21] J. C. Bardin *et al.*, "29.1 A 28nm Bulk-CMOS 4-to-8GHz ¡2mW Cryogenic Pulse Modulator for Scalable Quantum Computing," in *2019 IEEE International Solid- State Circuits Conference - (ISSCC)*, 2019, pp. 456–458. doi: 10.1109/ISSCC.2019.8662480.

[22] J. C. Bardin *et al.*, "Design and Characterization of a 28-nm Bulk-CMOS Cryogenic Quantum Controller Dissipating Less Than 2 mW at 3 K," *IEEE J Solid-State Circuits*, vol. 54, no. 11, pp. 3043–3060, 2019, doi: 10.1109/JSSC.2019.2937234.

[23] S. J. Pauka *et al.*, "A cryogenic CMOS chip for generating control signals for multiple qubits," *Nat Electron*, vol. 4, no. 1, pp. 64–70, Jan. 2021, doi: 10.1038/s41928-020-00528-y.

[24] J.-S. Park *et al.*, "13.1 A Fully Integrated Cryo-CMOS SoC for Qubit Control in Quantum Computers Capable of State Manipulation, Readout and High-Speed Gate Pulsing of Spin Qubits in Intel 22nm FFL FinFET Technology," in *2021 IEEE International Solid- State Circuits Conference (ISSCC)*, 2021, vol. 64, pp. 208–210. doi:

10.1109/ISSCC42613.2021.9365762.

[25] S. K. Tolpygo, V. Bolkhovsky, T. J. Weir, L. M. Johnson, M. A. Gouker, and W. D. Oliver, “Fabrication Process and Properties of Fully-Planarized Deep-Submicron Nb/Al–AlOx/Nb Josephson Junctions for VLSI Circuits,” *IEEE Transactions on Applied Superconductivity*, vol. 25, no. 3, pp. 1–12, 2015, doi: 10.1109/TASC.2014.2374836.

[26] C. Macklin *et al.*, “A near–quantum-limited Josephson traveling-wave parametric amplifier,” *Science (1979)*, vol. 350, no. 6258, pp. 307–310, Oct. 2015, doi: 10.1126/science.aaa8525.

[27] D. R. W. Yost *et al.*, “Solid-state qubits integrated with superconducting through-silicon vias,” *npj Quantum Inf*, vol. 6, no. 1, p. 59, 2020, doi: 10.1038/s41534-020-00289-8.

[28] Y. Okuma, N. Takeuchi, Y. Yamanashi, and N. Yoshikawa, “Design and Demonstration of an Adiabatic-Quantum-Flux-Parametron Field-Programmable Gate Array Using Josephson-CMOS Hybrid Memories,” *IEEE Transactions on Applied Superconductivity*, vol. 29, no. 8, pp. 1–6, 2019, doi: 10.1109/TASC.2019.2938577.

[29] O. A. Mukhanov, “Energy-Efficient Single Flux Quantum Technology,” *IEEE Transactions on Applied Superconductivity*, vol. 21, no. 3, pp. 760–769, 2011, doi: 10.1109/TASC.2010.2096792.

[30] Y. Yamanashi, T. Nishigai, and N. Yoshikawa, “Study of LR-Loading Technique for Low-Power Single Flux Quantum Circuits,” *IEEE Transactions on Applied Superconductivity*, vol. 17, no. 2, pp. 150–153, 2007, doi: 10.1109/TASC.2007.898608.

[31] Q. P. Herr, A. Y. Herr, O. T. Oberg, and A. G. Ioannidis, “Ultra-low-power superconductor logic,” *J Appl Phys*, vol. 109, no. 10, p. 103903, May 2011, doi: 10.1063/1.3585849.

[32] M. E. Çelik *et al.*, "25-GHz Operation of ERSFQ Time-to-Digital Converter," *IEEE Transactions on Applied Superconductivity*, vol. 31, no. 5, pp. 1–5, 2021, doi: 10.1109/TASC.2021.3065282.

[33] T. Ortlepp, O. Wetzstein, S. Engert, J. Kunert, and H. Toepfer, "Reduced Power Consumption in Superconducting Electronics," *IEEE Transactions on Applied Superconductivity*, vol. 21, no. 3, pp. 770–775, 2011, doi: 10.1109/TASC.2011.2117410.

[34] C. L. Ayala, T. Tanaka, R. Saito, M. Nozoe, N. Takeuchi, and N. Yoshikawa, "MANA: A Monolithic Adiabatic iNtegration Architecture Microprocessor Using 1.4-zJ/op Unshunted Superconductor Josephson Junction Devices," *IEEE J Solid-State Circuits*, vol. 56, no. 4, pp. 1152–1165, 2021, doi: 10.1109/JSSC.2020.3041338.

[35] S. Hasuo, *Josephson Microprocessors*. 1993. Accessed: Mar. 04, 2022. [Online]. Available: http://www.physics.umd.edu/courses/Phys798S/AnlageSpring16/Josephson%20Microprocessors.pdf

[36] S. Intiso, J. Pekola, A. Savin, Y. Devyatov, and A. Kidiyarova-Shevchenko, "Rapid single-flux-quantum circuits for low noise mK operation," *Supercond Sci Technol*, vol. 19, no. 5, pp. S335–S339, May 2006, doi: 10.1088/0953-2048/19/5/S36.

[37] E. Leonard *et al.*, "Digital Coherent Control of a Superconducting Qubit," *Phys Rev Appl*, vol. 11, no. 1, p. 14009, Jan. 2019, doi: 10.1103/PhysRevApplied.11.014009.

[38] C. J. Howington, "Digital Readout and Control of a Superconducting Qubit," 2019. Accessed: Mar. 06, 2022. [Online]. Available: https://surface.syr.edu/cgi/viewcontent.cgi?article=2122&context=etd

[39] E. Leonard Jr. *et al.*, "Digital coherent control of a superconducting qubit," Jun. 2018, doi:

10.1103/physrevapplied.11.014009.

[40] C. Howington, A. Opremcak, R. McDermott, A. Kirichenko, O. A. Mukhanov, and B. L. T. Plourde, "Interfacing Superconducting Qubits With Cryogenic Logic: Readout," *IEEE Transactions on Applied Superconductivity*, vol. 29, no. 5, pp. 1–5, 2019, doi: 10.1109/TASC.2019.2908884.

[41] K. Li, R. McDermott, and M. G. Vavilov, "Hardware-Efficient Qubit Control with Single-Flux-Quantum Pulse Sequences," *Phys Rev Appl*, vol. 12, no. 1, p. 14044, Jul. 2019, doi: 10.1103/PhysRevApplied.12.014044.

[42] D. v Averin, K. Rabenstein, and V. K. Semenov, "Rapid ballistic readout for flux qubits," *Phys Rev B*, vol. 73, no. 9, p. 94504, Mar. 2006, doi: 10.1103/PhysRevB.73.094504.

[43] N. v Klenov, A. v Kuznetsov, I. I. Soloviev, S. v Bakurskiy, M. v Denisenko, and A. M. Satanin, "Flux qubit interaction with rapid single-flux quantum logic circuits: Control and readout," *Low Temperature Physics*, vol. 43, no. 7, pp. 789–798, Jul. 2017, doi: 10.1063/1.4995627.

[44] K. G. Fedorov, A. v Shcherbakova, M. J. Wolf, D. Beckmann, and A. v Ustinov, "Fluxon Readout of a Superconducting Qubit," *Phys Rev Lett*, vol. 112, no. 16, p. 160502, Apr. 2014, doi: 10.1103/PhysRevLett.112.160502.

[45] C. J. Howington, "Digital Readout and Control of a Superconducting Qubit," 2019. Accessed: Mar. 06, 2022. [Online]. Available: https://surface.syr.edu/cgi/viewcontent.cgi?article=2122&context=etd

[46] S. S. Iyer, "Heterogeneous Integration for Performance and Scaling," *IEEE Trans Compon Packaging Manuf Technol*, vol. 6, no. 7, pp. 973–982, 2016, doi: 10.1109/TCPMT.2015.2511626.

[47] E. Perfecto and K. Srivastava, “Technology Trends: Past, Present, and Future,” in *Springer US*, 2013, pp. 23–52.

[48] P. Cheng, W. C. Wu, W. J. Wang, and T. M. Pai, “Challenge and process optimization of Thermal Compression bonding with Non Conductive Paste,” in *2015 IEEE 65th Electronic Components and Technology Conference (ECTC)*, 2015, pp. 484–489. doi: 10.1109/ECTC.2015.7159635.

[49] RUSHPCB, “CAUSES OF WARPAGE IN PCBS.” https://rushpcb.com/causes-of-warpage-in-pcbs/ (accessed Mar. 09, 2022).

[50] A. Manasson *et al.*, “Indium bump deposition for flip-chip micro-array image sensing and display applications,” in *Proc.SPIE*, May 2018, vol. 10639. doi: 10.1117/12.2303735.

[51] MacRumors, “New cores (A15),” 2021. https://forums.macrumors.com/threads/new-cores-a15.2310871/page-15 (accessed Apr. 23, 2022).

[52] R. H. Dennard, F. H. Gaensslen, H.-N. Yu, V. L. Rideout, E. Bassous, and A. R. LeBlanc, “Design of ion-implanted MOSFET’s with very small physical dimensions,” *IEEE J Solid-State Circuits*, vol. 9, no. 5, pp. 256–268, 1974, doi: 10.1109/JSSC.1974.1050511.

[53] J. F. McDonald, E. H. Rogers, K. Rose, and A. J. Steckl, “The trials of wafer-scale integration: Although major technical problems have been overcome since WSI was first tried in the 1960s, commercial companies can’t yet make it fly,” *IEEE Spectr*, vol. 21, no. 10, pp. 32–39, 1984, doi: 10.1109/MSPEC.1984.6370295.

[54] Cerebras, “Wafer Scale Engine: Why We Need Big Chips for Deep Learning.” https://cerebras.net/blog/cerebras-wafer-scale-engine-why-we-need-big-chips-for-deep-learning/ (accessed Mar. 10, 2022).

[55] EE Times China, “Cerebras全球最大芯片WSE升级二代：参数翻番，功耗不变.”

https://www.eet-china.com/news/202104211708.html (accessed Mar. 13, 2022).

[56] B. Hedrick *et al.*, “End-to-End Integration of a Multi-die Glass Interposer for System Scaling Applications,” in *2016 IEEE 66th Electronic Components and Technology Conference (ECTC)*, 2016, pp. 283–288. doi: 10.1109/ECTC.2016.261.

[57] S. Y. Hou *et al.*, “Wafer-Level Integration of an Advanced Logic-Memory System Through the Second-Generation CoWoS Technology,” *IEEE Trans Electron Devices*, vol. 64, no. 10, pp. 4071–4077, 2017, doi: 10.1109/TED.2017.2737644.

[58] H. Pu, H. J. Kuo, C. S. Liu, and D. C. H. Yu, “A Novel Submicron Polymer Re-Distribution Layer Technology for Advanced InFO Packaging,” in *2018 IEEE 68th Electronic Components and Technology Conference (ECTC)*, 2018, pp. 45–51. doi: 10.1109/ECTC.2018.00015.

[59] IEEE Electronics Packaging Society, “Heterogeneous Integration Roadmap 2020 Edition.” https://eps.ieee.org/technology/heterogeneous-integration-roadmap/2020-edition.html (accessed Mar. 13, 2022).

[60] S. Pal, D. Petrisko, R. Kumar, and P. Gupta, “Design Space Exploration for Chiplet-Assembly-Based Processors,” *IEEE Trans Very Large Scale Integr VLSI Syst*, vol. 28, no. 4, pp. 1062–1073, 2020, doi: 10.1109/TVLSI.2020.2968904.

[61] S. Pal, D. Petrisko, M. Tomei, P. Gupta, S. S. Iyer, and R. Kumar, “Architecting Waferscale Processors - A GPU Case Study,” in *2019 IEEE International Symposium on High Performance Computer Architecture (HPCA)*, 2019, pp. 250–263. doi: 10.1109/HPCA.2019.00042.

[62] Lisa Su, “Delivering the future of high-performance computing.” https://eri-summit.darpa.mil/docs/Su_Lisa_AMD_Final.pdf (accessed Mar. 13, 2022).

[63] S. Pal, D. Petrisko, A. A. Bajwa, P. Gupta, S. S. Iyer, and R. Kumar, "A Case for Packageless Processors," in *2018 IEEE International Symposium on High Performance Computer Architecture (HPCA)*, 2018, pp. 466–479. doi: 10.1109/HPCA.2018.00047.

[64] John Lau, "MCM, SiP, SoC, and Heterogeneous Integration Defined and Explained." https://www.3dincites.com/2017/08/mcm-sip-soc-and-heterogeneous-integration-defined-and-explained/ (accessed Mar. 13, 2022).

[65] V. Boyadzhyan and J. Choma, *Multichip Module Technologies*. 2000. Accessed: Mar. 13, 2022. [Online]. Available: https://picture.iczhiku.com/resource/eetop/shkHqyAdKLIURbMb.pdf

[66] J. Lee, C. Y. Lee, C. Kim, and S. Kalchuri, "Micro Bump System for 2nd Generation Silicon Interposer with GPU and High Bandwidth Memory (HBM) Concurrent Integration," in *2018 IEEE 68th Electronic Components and Technology Conference (ECTC)*, 2018, pp. 607–612. doi: 10.1109/ECTC.2018.00096.

[67] J. Kim, "Active Si interposer for 3D IC integrations," in *2015 International 3D Systems Integration Conference (3DIC)*, 2015, p. TS11.1.1-TS11.1.3. doi: 10.1109/3DIC.2015.7334619.

[68] K. Oi *et al.*, "Development of new 2.5D package with novel integrated organic interposer substrate with ultra-fine wiring and high density bumps," in *2014 IEEE 64th Electronic Components and Technology Conference (ECTC)*, 2014, pp. 348–353. doi: 10.1109/ECTC.2014.6897310.

[69] Newsroom, "Apple unveils M1 Ultra, the world's most powerful chip for a personal computer." https://www.apple.com/newsroom/2022/03/apple-unveils-m1-ultra-the-worlds-most-powerful-chip-for-a-personal-computer/ (accessed Mar. 10, 2022).

[70] NEWS DIRECTORY, “Apple’s M1 Ultra chip ‘assembles’ the secret recipe found in the patent – Apple.” https://www.newsdirectory3.com/apples-m1-ultra-chip-assembles-the-secret-recipe-found-in-the-patent-apple/ (accessed Mar. 15, 2022).

[71] S. Jangam, “Heterogeneous Integration on Silicon-Interconnect Fabric using fine-pitch interconnects (≤10 μm),” 2020.

[72] R. Leadbetter, “Inside the next Xbox: Project Scorpio tech revealed”, Accessed: Mar. 13, 2022. [Online]. Available: https://www.eurogamer.net/articles/digitalfoundry-2017-project-scorpio-tech-revealed

[73] NVIDIA, “Nvidia. Hot chips 30: Nvidia xavier soc.” https://old.hotchips.org/hc30/1conf/1.12_Nvidia_XavierHotchips2018Final_814.pdf (accessed Mar. 13, 2022).

[74] D. Yang and S. Wegner, “Apple iPhone Xs Max Teardown.” https://www.techinsights.com/blog/apple-iphone-xs-max-teardown (accessed Mar. 13, 2022).

[75] Qualcomm, “Qualcomm Datacenter Technologies Announces Commercial Shipment of Qualcomm Centriq 2400 – The World’s First 10nm Server Processor and Highest Performance Arm-based Server Processor Family Ever Designed.” https://www.qualcomm.com/news/releases/2017/11/08/qualcomm-datacenter-technologies-announces-commercial-shipment-qualcomm (accessed Mar. 13, 2022).

[76] I. Cutress, “3DFabric: The Home for TSMC’s 2.5D and 3D Stacking Roadmap.” https://www.anandtech.com/show/16051/3dfabric-the-home-for-tsmc-2-5d-and-3d-stacking-roadmap (accessed Mar. 13, 2022).

[77] J. I. Colless and D. J. Reilly, “Cryogenic high-frequency readout and control platform for

spin qubits," *Review of Scientific Instruments*, vol. 83, no. 2, p. 023902, Feb. 2012, doi: 10.1063/1.3681195.

[78] J. I. Colless and D. J. Reilly, "Modular Cryogenic Interconnects for Multi-Qubit Devices," May 2014, doi: 10.48550/arxiv.1405.3371.

[79] J. v Migacz and M. E. Huber, "Thermal annealing of Nb/Al-AlO/sub x//Nb Josephson junctions," *IEEE Transactions on Applied Superconductivity*, vol. 13, no. 2, pp. 123–126, 2003, doi: 10.1109/TASC.2003.813661.

[80] A. B. Pavolotsky, D. Dochev, and V. Belitsky, "Aging- and annealing-induced variations in Nb/Al–AlOx/Nb tunnel junction properties," *J Appl Phys*, vol. 109, no. 2, p. 024502, Jan. 2011, doi: 10.1063/1.3532040.

[81] J. v Gates, M. A. Washington, and M. Gurvitch, "Critical current uniformity and stability of Nb/Al-oxide-Nb Josephson junctions," *J Appl Phys*, vol. 55, no. 5, pp. 1419–1421, Mar. 1984, doi: 10.1063/1.333235.

[82] B. Foxen *et al.*, "Qubit compatible superconducting interconnects," *Quantum Sci Technol*, vol. 3, no. 1, p. 014005, Jan. 2018, doi: 10.1088/2058-9565/aa94fc.

[83] X. Li *et al.*, "Vacuum-gap transmon qubits realized using flip-chip technology," Jun. 2021, doi: 10.1063/5.0068255.

[84] K. Fujiwara and M. Asahi, "Characterization of Intermetallic Compound Formation on In/Bi/Sn Solder Bumps Used in Pb-Alloy Josephson Chip Packaging," *IEEE Transactions on Components, Hybrids, and Manufacturing Technology*, vol. 10, no. 2, pp. 263–266, 1987, doi: 10.1109/TCHMT.1987.1134729.

[85] K. E. Yokoyama, G. Akerling, A. D. Smith, and M. Wire, "Robust superconducting die attach process," *IEEE Transactions on Applied Superconductivity*, vol. 7, no. 2, pp. 2631–

2634, 1997, doi: 10.1109/77.621779.

[86] S. B. Kaplan, V. Dotsenko, and D. Tolpygo, "High-Speed Experimental Results for an Adhesive-Bonded Superconducting Multi-Chip Module," *IEEE Transactions on Applied Superconductivity*, vol. 17, no. 2, pp. 971–974, 2007, doi: 10.1109/TASC.2007.897376.

[87] C. R. Conner *et al.*, "Superconducting qubits in a flip-chip architecture," *Appl Phys Lett*, vol. 118, no. 23, p. 232602, Jun. 2021, doi: 10.1063/5.0050173.

[88] Silicon Valley Microelectronics (SVM), "300mm Silicon Wafer specification." https://svmi.com/wp-content/uploads/2020/09/SV025.pdf (accessed Mar. 09, 2022).

[89] S. Jangam and S. S. Iyer, "Silicon-Interconnect Fabric for Fine-Pitch (≤10 µm) Heterogeneous Integration," *IEEE Trans Compon Packaging Manuf Technol*, vol. 11, no. 5, pp. 727–738, 2021, doi: 10.1109/TCPMT.2021.3075219.

[90] A. A. Bajwa *et al.*, "Demonstration of a Heterogeneously Integrated System-on-Wafer (SoW) Assembly," in *2018 IEEE 68th Electronic Components and Technology Conference (ECTC)*, 2018, pp. 1926–1930. doi: 10.1109/ECTC.2018.00288.

[91] S. Jangam, A. A. Bajwa, K. K. Thankkappan, P. Kittur, and S. S. Iyer, "Electrical Characterization of High Performance Fine Pitch Interconnects in Silicon-Interconnect Fabric," in *2018 IEEE 68th Electronic Components and Technology Conference (ECTC)*, 2018, pp. 1283–1288. doi: 10.1109/ECTC.2018.00197.

[92] S. Jangam, S. Pal, A. Bajwa, S. Pamarti, P. Gupta, and S. S. Iyer, "Latency, Bandwidth and Power Benefits of the SuperCHIPS Integration Scheme," in *2017 IEEE 67th Electronic Components and Technology Conference (ECTC)*, 2017, pp. 86–94. doi: 10.1109/ECTC.2017.246.

[93] Computer History Museum, "Birth of the Computer."

https://www.computerhistory.org/revolution/birth-of-the-computer/4/78 (accessed Mar. 16, 2022).

[94] A. M. Savin, J. P. Pekola, D. v Averin, and V. K. Semenov, “Thermal budget of superconducting digital circuits at subkelvin temperatures,” *J Appl Phys*, vol. 99, no. 8, p. 084501, Apr. 2006, doi: 10.1063/1.2187276.

[95] S. J. Pauka *et al.*, “A cryogenic interface for controlling many qubits,” *ArXiv*, pp. 525–528, 2019.

[96] FUJITSU, “FUJITSU Processor A64FX.” https://www.fujitsu.com/downloads/SUPER/a64fx/a64fx_datasheet.pdf (accessed Mar. 16, 2022).

[97] RIKEN Center for Computational Science, “About Fugaku.” https://www.r-ccs.riken.jp/en/fugaku/about/ (accessed Mar. 16, 2022).

[98] M. Ito *et al.*, “Design and implementation of SFQ programmable clock generators,” *Physica C Supercond*, vol. 412–414, pp. 1550–1554, 2004, doi: https://doi.org/10.1016/j.physc.2004.02.220.

[99] Y. He *et al.*, “Compact RSFQ microwave pulse generator based on an integrated RF module for controlling superconducting qubits,” *Appl Phys Lett*, vol. 120, no. 6, p. 062601, Feb. 2022, doi: 10.1063/5.0083972.

[100] Lizhen Zheng, “High-Speed Rapid-Single-Flux-Quantum Multiplexer and Demultiplexer Design and Testing,” 2007. Accessed: Mar. 16, 2022. [Online]. Available: https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.417.3480&rep=rep1&type=pdf

[101] M. A. Castellanos-Beltran *et al.*, “Single-Flux-Quantum Multiplier Circuits for Synthesizing Gigahertz Waveforms With Quantum-Based Accuracy,” *IEEE Transactions*

on Applied Superconductivity, vol. 31, no. 3, pp. 1–9, 2021, doi: 10.1109/TASC.2021.3057013.

[102] J. Aumentado, “Superconducting Parametric Amplifiers: The State of the Art in Josephson Parametric Amplifiers,” *IEEE Microw Mag*, vol. 21, no. 8, pp. 45–59, 2020, doi: 10.1109/MMM.2020.2993476.

[103] Y. J. Rosen *et al.*, “Protecting superconducting qubits from phonon mediated decay,” *Appl Phys Lett*, vol. 114, no. 20, p. 202601, May 2019, doi: 10.1063/1.5096182.

[104] R.-P. Riwar *et al.*, “Normal-metal quasiparticle traps for superconducting qubits,” *Phys Rev B*, vol. 94, no. 10, p. 104516, Sep. 2016, doi: 10.1103/PhysRevB.94.104516.

[105] B. Abdo *et al.*, “Active protection of a superconducting qubit with an interferometric Josephson isolator,” *Nat Commun*, vol. 10, no. 1, p. 3154, 2019, doi: 10.1038/s41467-019-11101-3.

[106] M. Rymarz, S. Bosco, A. Ciani, and D. P. DiVincenzo, “Hardware-Encoding Grid States in a Non-Reciprocal Superconducting Circuit,” Feb. 2020, doi: 10.1103/physrevx.11.011032.

[107] A. Dasgupta, A. Alam, G. Ouyang, S. Jangam, and S. S. Iyer, “Antenna on Silicon Interconnect Fabric,” in *2020 IEEE 70th Electronic Components and Technology Conference (ECTC)*, 2020, pp. 1788–1794. doi: 10.1109/ECTC32862.2020.00279.

[108] S. Lahiri *et al.*, “Packaging Technology for Josephson Integrated Circuits,” *IEEE Transactions on Components, Hybrids, and Manufacturing Technology*, vol. 5, no. 2, pp. 271–280, 1982, doi: 10.1109/TCHMT.1982.1135958.

[109] ARDENT CONCEPTS, “ SP Production Proven Compression Interconnect.” https://www.ardentconcepts.com/spring-probe/#.YjTZ2BDML0o (accessed Mar. 17,

2022).

[110] M. Hamilton, “Packaging and Interconnect Technologies for Cryogenic & Quantum Systems.” https://r6.ieee.org/scv-eps/wp-content/uploads/sites/58/2021/02/eps2104b.pdf (accessed Mar. 15, 2022).

[111] O. Mukhanov, “Digital Quantum Computing based on Superconducting Electronics,” in *NY CREATES Seminar*, May 2020. Accessed: Mar. 17, 2022. [Online]. Available: https://ny-creates.org/wp-content/uploads/NY-CREATES-Seminar_Mukhanov_20May2020_distr.pdf

[112] K. J. S. Sokhey, S. K. Rai, and G. S. Lodha, “Oxidation studies of niobium thin films at room temperature by X-ray reflectivity,” *Appl Surf Sci*, vol. 257, no. 1, pp. 222–226, 2010, doi: https://doi.org/10.1016/j.apsusc.2010.06.069.

[113] J. Halbritter, “On the oxidation and on the superconductivity of niobium,” *Applied Physics A*, vol. 43, no. 1, pp. 1–28, 1987, doi: 10.1007/BF00615201.

[114] R. W. Powers and M. v Doyle, “Diffusion of Interstitial Solutes in the Group V Transition Metals,” *J Appl Phys*, vol. 30, no. 4, pp. 514–524, Apr. 1959, doi: 10.1063/1.1702398.

[115] J. Völkl, G. Schaumann, and G. Alefeld, “Diffusional relaxation of H and D in V and Nb,” *Journal of Physics and Chemistry of Solids*, vol. 31, no. 8, pp. 1805–1809, 1970, doi: https://doi.org/10.1016/0022-3697(70)90171-X.

[116] Okamoto H., *Binary Alloy Phase Diagrams*, 2nd ed. ASM International, 1990.

[117] A. Gold *et al.*, “Entanglement Across Separate Silicon Dies in a Modular Superconducting Qubit Device,” Feb. 2021, doi: 10.48550/arxiv.2102.13293.

[118] V. Gupta *et al.*, “Atomic Layer Deposited Materials as Barrier Layers for Preservation of Nb Superconductivity in Multilayered Thin-Film Structures,” *IEEE Transactions on*

Applied Superconductivity, vol. 31, no. 5, pp. 1–4, 2021, doi: 10.1109/TASC.2021.3055752.

[119] D. Ablitzee, “Diffusion of niobium, iron, cobalt, nickel and copper in niobium,” *The Philosophical Magazine: A Journal of Theoretical Experimental and Applied Physics*, vol. 35, no. 5, pp. 1239–1256, May 1977, doi: 10.1080/14786437708232950.

[120] J. H. Junkins, R. L. Farrar, E. J. Barber, and H. A. Bernhardt, “Preparation and Physical Properties of Niobium Pentafluoride1,” *J Am Chem Soc*, vol. 74, no. 14, pp. 3464–3466, Jul. 1952, doi: 10.1021/ja01134a002.

[121] R. Esser and A. Christou, “Gold/Niobium Thin Film Metallizations for GaAs Devices and Circuits,” *MRS Proceedings*, vol. 695, p. L6.23.1, 2001, doi: DOI: 10.1557/PROC-695-L6.23.1.

[122] M. Liehr *et al.*, “Comparative study of Nb and TiW barrier layers between Au and a-SiO2,” *Journal of Vacuum Science & Technology A*, vol. 2, no. 2, pp. 288–291, Apr. 1984, doi: 10.1116/1.572583.

[123] Quantum Design NORTH AMERICA, “ Physical Property Measurement System.” https://www.qdusa.com/products/dynacool.html (accessed Mar. 22, 2022).

[124] Y. Asada, “Superconductivity of Niobium Films,” *J Physical Soc Japan*, vol. 26, pp. 347–354, 1969.

[125] K. Il’in *et al.*, “Influence of thickness, width and temperature on critical current density of Nb thin film structures,” *Physica C Supercond*, vol. 470, no. 19, pp. 953–956, 2010, doi: https://doi.org/10.1016/j.physc.2010.02.042.

[126] K. Ilin, D. Henrich, Y. Luck, Y. Liang, M. Siegel, and D. Yu. Vodolazov, “Critical current of Nb, NbN, and TaN thin-film bridges with and without geometrical nonuniformities in a

magnetic field," *Phys Rev B*, vol. 89, no. 18, p. 184511, May 2014, doi: 10.1103/PhysRevB.89.184511.

[127] M. Y. Kupryanov and V. F. Lukichev, "Temperature dependence of pair-breaking current in superconductors," vol. 6:4, 1980, [Online]. Available: https://www.osti.gov/biblio/6826630

[128] R. N. Das *et al.*, "Large Scale Cryogenic Integration Approach for Superconducting High-Performance Computing," in *2017 IEEE 67th Electronic Components and Technology Conference (ECTC)*, 2017, pp. 675–683. doi: 10.1109/ECTC.2017.54.

[129] W. T. Chen and C. W. Nelson, "Thermal Stress in Bonded Joints," *IBM J Res Dev*, vol. 23, no. 2, pp. 179–188, 1979, doi: 10.1147/rd.232.0179.

[130] Y. Nobumori *et al.*, "Design and Implementation of a Fully Asynchronous SFQ Microprocessor: SCRAM2," *IEEE Transactions on Applied Superconductivity*, vol. 17, no. 2, pp. 478–481, 2007, doi: 10.1109/TASC.2007.898658.

[131] T. v Filippov *et al.*, "20GHz Operation of an Asynchronous Wave-Pipelined RSFQ Arithmetic-Logic Unit," *Phys Procedia*, vol. 36, pp. 59–65, 2012, doi: https://doi.org/10.1016/j.phpro.2012.06.130.

[132] G. Matsuda, Y. Yamanashi, and N. Yoshikawa, "Design of an SFQ Microwave Chopper for Controlling Quantum Bits," *IEEE Transactions on Applied Superconductivity*, vol. 17, no. 2, pp. 146–149, 2007, doi: 10.1109/TASC.2007.898657.

[133] S. Razmkhah, A. Bozbey, and P. Febvre, "A Compact High Frequency Voltage Amplifier for Superconductor-Semiconductor Logic Interface," Nov. 2020, doi: 10.1088/1361-6668/abdedb.

[134] J. Krupka, K. Derzakowski, M. Tobar, J. Hartnett, and R. G. Geyer, "Complex

permittivity of some ultralow loss dielectric crystals at cryogenic temperatures," *Meas Sci Technol*, vol. 10, no. 5, pp. 387–392, May 1999, doi: 10.1088/0957-0233/10/5/308.

[135] J. Krupka, J. Breeze, A. Centeno, N. Alford, T. Claussen, and L. Jensen, "Measurements of Permittivity, Dielectric Loss Tangent, and Resistivity of Float-Zone Silicon at Microwave Frequencies," *IEEE Trans Microw Theory Tech*, vol. 54, no. 11, pp. 3995–4001, 2006, doi: 10.1109/TMTT.2006.883655.

[136] K. Zhou, S. Caroopen, Y. Delorme, M. Batrung, M. Gheudin, and S. Shi, "Dielectric Constant and Loss Tangent of Silicon at 700–900 GHz at Cryogenic Temperatures," *IEEE Microwave and Wireless Components Letters*, vol. 29, no. 7, pp. 501–503, 2019, doi: 10.1109/LMWC.2019.2920532.

[137] J. Gao and J. Zmuidzinas, "The Physics of Superconducting Microwave Resonators," California Institute of Technology, Ann Arbor, 2008. [Online]. Available: https://www.proquest.com/dissertations-theses/physics-superconducting-microwave-resonators/docview/1080814168/se-2?accountid=14512

[138] A. I. Gubin, K. S. Il'in, S. A. Vitusevich, M. Siegel, and N. Klein, "Dependence of magnetic penetration depth on the thickness of superconducting Nb thin films," *Phys Rev B*, vol. 72, no. 6, p. 64503, Aug. 2005, doi: 10.1103/PhysRevB.72.064503.

[139] B. Cooper, "SIMULATION TECHNIQUES FOR SUPERCONDUCTING MICROWAVE DEVICES," 2020. Accessed: Mar. 23, 2022. [Online]. Available: https://www.research.manchester.ac.uk/portal/files/160457551/FULL_TEXT.PDF

[140] J. Bardin, "Beyond-Classical Computing Using Superconducting Quantum Processors," in *2022 IEEE International Solid- State Circuits Conference (ISSCC)*, 2022, vol. 65, pp. 422–424. doi: 10.1109/ISSCC42614.2022.9731635.

[141] H. You and M. Soma, “Crosstalk Analysis of Interconnection Lines and Packages in High-Speed Integrated Circuits,” *IEEE Trans Circuits Syst*, vol. 37, no. 8, pp. 1019–1026, 1990, doi: 10.1109/31.56075.

[142] J. Kim and J. F. Mcdonald, “Transient and Crosstalk Analysis of Slightly Lossy Interconnection Lines for Wafer Scale Integration and Wafer Scale Hybrid Packaging—Weak Coupling Case,” *IEEE Trans Circuits Syst*, vol. 35, no. 11, pp. 1369–1382, 1988, doi: 10.1109/31.14461.

[143] T. Mido and K. Asada, “Crosstalk Noise in High Density and High Speed Interconnections due to Inductive Coupling,” 1997.

[144] K. Agarwal, D. Sylvester, and D. Blaauw, “Modeling and analysis of crosstalk noise in coupled RLC interconnects,” in *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, May 2006, vol. 25, no. 5, pp. 892–901. doi: 10.1109/TCAD.2005.855961.

[145] Oxford Instrument, “Dilution refrigerator KelvinoxMX.” https://nanoscience.oxinst.com/products/kelvinoxmx (accessed Mar. 23, 2022).

[146] Mouser Electronics, “Coaxial Switches Cryogenic SP6T Ramses SMA.” https://www.mouser.com/ProductDetail/Radiall/R583423141?qs=%2Fha2pyFadugGk96rmd8OgyVSctFw7IHbTBAoePIDklHwqoCoghbeTA%3D%3D (accessed Mar. 24, 2022).

[147] X. Zhang *et al.*, “Fabrication and measurement of Nb-based SQUID magnetometer,” *Physica C: Superconductivity and its Applications*, vol. 548, pp. 1–4, 2018, doi: https://doi.org/10.1016/j.physc.2018.01.016.

[148] B. Abdo *et al.*, “Active protection of a superconducting qubit with an interferometric Josephson isolator,” *Nat Commun*, vol. 10, no. 1, p. 3154, 2019, doi: 10.1038/s41467-019-

11101-3.

[149] M. Currie, R. Sobolewski, and T. Y. Hsiang, “High-frequency crosstalk in superconducting microstrip waveguide interconnects,” *IEEE Transactions on Applied Superconductivity*, vol. 9, no. 2, pp. 3602–3605, 1999, doi: 10.1109/77.783809.

[150] L. Condra, J. Svitak, and A. Pense, “The High Temperature Deformation Properties of Gold and Thermocompression Bonding,” *IEEE Transactions on Parts, Hybrids, and Packaging*, vol. 11, no. 4, pp. 290–296, 1975, doi: 10.1109/TPHP.1975.1135079.

[151] M. Yamamoto *et al.*, “Effect of Au Film Thickness and Surface Roughness on Room-Temperature Wafer Bonding and Wafer-Scale Vacuum Sealing by Au-Au Surface Activated Bonding,” *Micromachines (Basel)*, vol. 11, no. 5, p. 454, Apr. 2020, doi: 10.3390/mi11050454.

[152] R. SIMONS, *Coplanar Waveguide Circuits, Components, and Systems*. 2001.

[153] A. Shah *et al.*, “Towards Cable-to-Cable Connectors for Flexible Thin-Film Superconducting Transmission Lines,” *IEEE Transactions on Applied Superconductivity*, vol. 31, no. 5, pp. 1–6, 2021, doi: 10.1109/TASC.2021.3062566.

[154] A. Shah *et al.*, “Superconducting Molybdenum Multi-Chip Module Approach for Cryogenic and Quantum Applications”, doi: 10.1109/ECTC51906.2022.00052.