Ultra-Low-Power Superconductor Logic

Quentin P. Herr\textsuperscript{a}, Anna Y. Herr, Oliver T. Oberg, and Alexander G. Ioannidis
Northrop Grumman Systems Corp., Baltimore, MD 21240
(Dated: January 15, 2013)

We have developed a new superconducting digital technology, Reciprocal Quantum Logic, that uses AC power carried on a transmission line, which also serves as a clock. Using simple experiments we have demonstrated zero static power dissipation, thermally limited dynamic power dissipation, high clock stability, high operating margins and low BER. These features indicate that the technology is scalable to far more complex circuits at a significant level of integration. On the system level, Reciprocal Quantum Logic combines the high speed and low-power signal levels of Single-Flux-Quantum signals with the design methodology of CMOS, including low static power dissipation, low latency combinational logic, and efficient device count.

Power consumption has increasingly become a limiting factor in high performance digital circuits and systems. According to a U.S. Environmental Protection Agency study\textsuperscript{[1]}, the demand of servers and data centers in the U.S. is approaching 12 GW, equivalent to the output of 25 power plants. Here we show a new logic family, Reciprocal Quantum Logic, that combines the low energy and high clock rates of superconductor devices with the essential qualities of CMOS, including low static power dissipation, low latency combinational logic, and efficient device count. On a system level, this yields a factor of 300 reduction in power compared to projected nano-scale CMOS, while taking in account the overhead of cryogenic operating temperature.

Superconducting digital electronics has long been considered the ultimate low energy alternative to CMOS\textsuperscript{[2]} based on the fundamental advantages of lossless interconnect and fast, low energy signal levels. Passive superconducting interconnect allows data transmission without signal amplification. Superconducting interconnects have a typical bandwidth of 700 GHz, which has allowed serial data rates up to 60 Gbs\textsuperscript{−1} for chip-to-chip communication\textsuperscript{[3]}. Lossless interconnects would enable large systems with high computational density, as compared to conventional systems where interconnect dominates the total power budget.

Unlike transistor circuits, where dissipated power is set by device size and materials, superconductor circuits are in the regime where device size and power dissipation is set by the thermal noise limit. The active device, the Josephson junction, generates quantum accurate digital information in the form of Single Flux Quanta (SFQ) of magnetic field \(\Phi_0 = h/2e = 2.06 \times 10^{-15}\) Wb. Using equivalent units \(\Phi_0 \approx 2\) mV ps = 2 mApH illustrates that the SFQ can exist as a transient voltage pulse across the Josephson junction \(V = \int_0^\infty Vdt \equiv \Phi_0\) or as a persistent current in a superconducting inductive loop. For a typical minimum critical current of 0.1 mA at liquid helium temperature, the SFQ pulse energy \(E_{SFQ} = \int_0^{\infty} IVdt \approx I_c\Phi_0\) is only \(1 \times 10^{-19}\) J. This is only about three orders of magnitude above the fundamental thermal Boltzmann limit, \(k_B T\), and is the practical limit for classical digital circuits operating with low bit error rate. Beyond this limit there are only reversible computing and quantum computing.

Numerous circuits with record-breaking clock rates have been demonstrated in the Rapid-Single-Flux-Quantum (RSFQ)\textsuperscript{[4]} logic family, including a static digital divider operating up to 770 GHz\textsuperscript{[5]}, digital signal processors clocked at 20-40 GHz\textsuperscript{[6], [7]}, and a serial microprocessor at 20 GHz\textsuperscript{[8]}. Cryocooler-mounted prototypes have included a digital receiver used for satellite communications\textsuperscript{[9]} and high-end router components operating at 47 Gbs\textsuperscript{−1} port speed\textsuperscript{[10]}. However, the technology suffers from high overhead in static power dissipation and device count, which offset the energy advantages of SFQ data encoding. RSFQ circuits use DC power delivered on a common voltage rail via bias resistors, which is analogous to TTL logic and inferior to CMOS due to static power dissipation. Ten times more power is dissipated in the bias resistors than in the active devices even in a fully active RSFQ circuit. While the power rail voltage is only about 1 mV, it draws significant current, reaching 1 A for a circuit with 10,000 Josephson junctions. This results in high parasitic heat load in the cryopackage. Additional overhead is incurred in the timing design; RSFQ uses an active clock distribution network, which leads to significant accumulated jitter and timing variations based on device parameters and data pattern statistics. Timing design of high speed circuits results in a total device count that is dominated by the clock distribution network\textsuperscript{[11]}. RSFQ is pipelined on the gate level, which enables high clock rates but also incurs high latency. The new logic family described here circumvents each of these limitations while preserving the fundamental property of SFQ data encoding.

We report a new superconducting logic family, Reciprocal Quantum Logic (RQL), that eliminates static power by replacing bias resistors with inductive coupling to an AC transmission line that effectively powers the devices...
in series and eliminates large ground return current. The AC power also serves as a stable clock reference signal, preventing accumulated clock jitter. The novel power supply is paired with a novel data encoding. A logical "one" is encoded as a reciprocal pair of SFQ pulses of opposite polarity. During the positive half cycle, the logic operation involves storage and routing of SFQ data pulses. While the gates have internal state with respect to the positive pulse, the trailing negative-polarity SFQ pulse serves as a reset. This greatly simplifies gate design and produces combinational logic behavior. Similar to CMOS, these combinational gates allow multiple levels of logic per stage for low latency. Overall, RQL combines the high speed and low-power signal levels of SFQ signals with the design methodology of CMOS.

Using simple experiments involving logic gates and a 1600-device shift register and logic gates, we have demonstrated that RQL is at once high speed and low energy with a low bit-error rate. We measure energy dissipation to be within a factor of 1000 of the thermal limit at clock rates in the range 2-10 GHz for the shift register, and negligible BER of less than $10^{-40}$ for the logic gates while maintaining operating margins of $\pm 30\%$. AC power is supplied to the circuit on superconductor microstrip transmission line, which also serves as a passive clock distribution network. We show high stability of the clock at frequencies up to 12 GHz. The technology scales to the one-million device level clocked at 6 GHz with only a 6 mW power supply, amounting to only 15 mA on a 50 $\Omega$ line with dynamic timing variation of only $\pm 1\%$ of the clock period. This indicates that the technology is scalable to complex circuits at a significant level of integration. Computational efficiency of the circuits is nearly three orders of magnitude higher in terms of operations per Joule compared to high performance CMOS. Taking into account that superconductor circuits require a cryocooler, with efficiency of 1,000 W/W achievable at 4.2 Kelvin, RQL circuits offer a system-level factor of 300 less wall-plug power dissipation. This makes RQL technology attractive for many applications, including high end computing.

**Results**

**Power dissipation.** RQL circuits have zero static power dissipation, so for the first time dynamic power dissipation in a superconducting SFQ circuit could be measured directly. The clock power is carried on 50 $\Omega$ lines that return to room temperature without termination on chip, allowing direct measurement of the relative amplitude of the output waveforms for an inactive and fully active circuit. Because dynamic power dissipation is so small, the experiment requires a circuit with a large number of Josephson junctions and a low AC power amplitude with relatively high coupling to the clock line.

The shift register was chosen as a convenient test vehicle.

**Fig. 1** shows the schematic and physical layout of one bit of the RQL shift register circuit. The four-phase clock is a fundamental feature that provides directionality. Without this, the positive pulse that moves forward during the positive half clock cycle would travel backward during the negative half of the cycle, annihilating the negative pulse. Instead, the positive pulse rides the leading edge of the clock from one phase to the next and arrives at the output after one cycle of delay, and the
of Josephson junctions and clock frequency. Measured power dissipation in the circuit is three times smaller than this estimate. Additional data points corresponding to a 6 Gbs⁻¹ pseudo-random input pattern show 0.6 µW total power dissipation in the 800 Josephson junctions on each clock line. This is half the power dissipation of the all “ones,” as expected, and is only three orders of magnitude above the von Neumann-Landauer thermal limit \[\mu B k T \ln 2\] per bit.

A model for SFQ dissipation based on the energy potential indicates that the work done on a switching junction is a function of bias current rather than critical current \[J_0\]. Physical-level simulation \[13\] of the circuit shows that in the range of interest, where clock period is much longer than the switching time of the Josephson junctions, data pulses pass through each stage early in the clock cycle under low bias conditions. This results in a low energy SFQ pulse with simulated dissipation of about 0.25 \(L \Phi_0\) at 6 Gbs⁻¹. Switching of the Josephson junctions is shifted to higher bias at higher frequency, producing a slightly non-linear frequency dependence. Experimental data agrees with this simulation result fit with a single prefactor of 1.35.

**Clock phase stability.** Switching of Josephson junctions not only attenuates the AC clock but also adds accumulative delay. The magnitude of this effect can be estimated using a simple linear model where the Josephson junction acts as an inductor if superconducting, and as a resistor if switching. The clock propagation time in the case of all digital “ones” is the same as for an isolated line \(\tau = \sqrt{L_c C_c} = 7.6 \text{ fs/µm}\), where \(L_c = 0.3 \text{ pH/µm}\) and \(C_c = 0.29 \text{ fF/µm}\) are the clock line inductance and capacitance in the circuit. In the case of all digital “zeros,” propagation time is \(\tau' = \sqrt{L_c' C_c'}\) with inductance \(L_c'\) given by the impedance transformation for inductive coupling

\[
\frac{L_c'}{L_c} = (1 - k^2) + 2 \frac{k^2}{1 + (L_g/L_b)},
\]

where \(L_b\) is the bias inductor and \(k = L_m/\sqrt{L_c L_b}\) is the magnetic coupling constant as shown in Fig. 10. \(L_g\) is the inductance of the RQL gate connected to the bias inductor. In the shift register, \(L_g = (L_{j1} + L_1)(L_{j2} + L_2)\) is the series and parallel combination of the interconnect and the Josephson inductances \(L_{j} = \Phi_0/2\pi I_c\). In general terms, data-dependent phase delay of the clock scales as \(k^2\) and can be minimized by reducing coupling to the clock line and increasing AC clock power.

Accumulated variable clock delay for the entire 200-bit shift register is 1.4±0.2 ps and is independent of frequency. Variable delay was directly observed on the clock return from the chip at 2-12 GHz on a sampling oscilloscope to compare the data patterns of all “ones” and all “zeros”. Accuracy was limited by drift between the two phase-locked synthesizers that clocked the chip and triggered the oscilloscope. This result is in agreement with the analytical estimate of Eq. 1 which gives 1.5 ps vari-

---

**FIG. 2: Power dissipation.** Power ratio of the clock output for the two data patterns, corresponding to all “ones” and all “zeros,” was measured for frequencies of 2-12 GHz. Power dissipation is derived from the directly-observed power ratio and the on-chip clock power of 12.5 µW, calculated as the geometric mean of applied and returned power. At 6 GHz and below, measurement error is within the size of the data points. At 8 GHz and above, the primary source of error is variation in clock attenuation on the different lines in the American Cryoprobe BCP-2 chip holder, producing visible spread between data points for clock 1 and clock 2. Additional data points correspond to a 6 Gbs⁻¹ pseudo-random input pattern that shows half the power dissipation compared to all “ones”, as expected. Measured power dissipation agrees well with the result from circuit simulation, \(P_{num}\), with a single multiplicative fitting parameter. However, the power is three times smaller than the analytical estimate that scales with circuit size and frequency as \(2nIL_0f\) and that would apply to dc-powered SFQ devices.
able clock delay for the complete 200-bit shift register circuit.

**Logic gates.** Routing and processing of pulse-based signals is distinct from transistor-based voltage-state logic, as shown in Fig. 3. Logical A-and-not-B (AnotB) means that an input pulse A will propagate to output Q unless a pulse on input B comes first. Logical And & Or (AndOr) means that the first input pulse, if any, goes to Q1 (logical OR), and the second input pulse goes to Q2 (logical AND). Inputs to the AnotB gate must satisfy the timing requirement that B arrive before A generates an output. There is no similar timing requirement for the AndOr gate.

The logical behavior of the gates is based on the reciprocal data encoding. Considering only the positive pulses, the gates are similar to the state machines of RSFQ logic, as input changes the internal flux state of the inductive loops. However, the trailing negative pulse erases the internal state every clock cycle and produces combinational logic behavior. The reset operation afforded by the negative pulse greatly simplifies the logic design, so that each gate consists of only two active devices with inductive interconnect. The gates have large parametric operating margins, including simulated tolerances on junction critical currents of at least ±50%.

On the physical level, both RQL gates have a bistable internal flux state corresponding to ±Φ0/2. The AnotB
The BER of the AnotB gate at 6 GHz is shown as function of its flux bias $I_{flux}$. A 32-bit input pattern generated with an Anritsu MP1763C was split and applied to the inputs with a 15 bit relative shift, and the XOR output was compared to the correct pattern with an Anritsu MP1764C error detector. Error bars on the lowest points correspond to counting statistics of 4 errors (left) and 5 errors (right). Near the center, no errors detected for a period of 30 hours gives an error floor below $10^{-15}$ for the entire circuit. The data fit to the error function extrapolate to a minimum BER of $10^{-480}$ at the optimal bias of 1.82 mA. Additional curves correspond to the BER scaled for reduced device size and power.

![Bit Error Rate (BER)](image.png)

FIG. 5: **Bit-error rate (BER).** The BER of the AnotB gate at 6 GHz is shown as function of its flux bias $I_{flux}$. A 32-bit input pattern generated with an Anritsu MP1763C was split and applied to the inputs with a 15 bit relative shift, and the XOR output was compared to the correct pattern with an Anritsu MP1764C error detector. Error bars on the lowest points correspond to counting statistics of 4 errors (left) and 5 errors (right). Near the center, no errors detected for a period of 30 hours gives an error floor below $10^{-15}$ for the entire circuit. The data fit to the error function extrapolate to a minimum BER of $10^{-480}$ at the optimal bias of 1.82 mA. Additional curves correspond to the BER scaled for reduced device size and power.

The BER of the AnotB gate at 6 GHz is shown as function of its flux bias $I_{flux}$. A 32-bit input pattern generated with an Anritsu MP1763C was split and applied to the inputs with a 15 bit relative shift, and the XOR output was compared to the correct pattern with an Anritsu MP1764C error detector. Error bars on the lowest points correspond to counting statistics of 4 errors (left) and 5 errors (right). Near the center, no errors detected for a period of 30 hours gives an error floor below $10^{-15}$ for the entire circuit. The data fit to the error function extrapolate to a minimum BER of $10^{-480}$ at the optimal bias of 1.82 mA. Additional curves correspond to the BER scaled for reduced device size and power.

The BER of the AnotB gate at 6 GHz is shown as function of its flux bias $I_{flux}$. A 32-bit input pattern generated with an Anritsu MP1763C was split and applied to the inputs with a 15 bit relative shift, and the XOR output was compared to the correct pattern with an Anritsu MP1764C error detector. Error bars on the lowest points correspond to counting statistics of 4 errors (left) and 5 errors (right). Near the center, no errors detected for a period of 30 hours gives an error floor below $10^{-15}$ for the entire circuit. The data fit to the error function extrapolate to a minimum BER of $10^{-480}$ at the optimal bias of 1.82 mA. Additional curves correspond to the BER scaled for reduced device size and power.

The BER of the AnotB gate at 6 GHz is shown as function of its flux bias $I_{flux}$. A 32-bit input pattern generated with an Anritsu MP1763C was split and applied to the inputs with a 15 bit relative shift, and the XOR output was compared to the correct pattern with an Anritsu MP1764C error detector. Error bars on the lowest points correspond to counting statistics of 4 errors (left) and 5 errors (right). Near the center, no errors detected for a period of 30 hours gives an error floor below $10^{-15}$ for the entire circuit. The data fit to the error function extrapolate to a minimum BER of $10^{-480}$ at the optimal bias of 1.82 mA. Additional curves correspond to the BER scaled for reduced device size and power.

The BER of the AnotB gate at 6 GHz is shown as function of its flux bias $I_{flux}$. A 32-bit input pattern generated with an Anritsu MP1763C was split and applied to the inputs with a 15 bit relative shift, and the XOR output was compared to the correct pattern with an Anritsu MP1764C error detector. Error bars on the lowest points correspond to counting statistics of 4 errors (left) and 5 errors (right). Near the center, no errors detected for a period of 30 hours gives an error floor below $10^{-15}$ for the entire circuit. The data fit to the error function extrapolate to a minimum BER of $10^{-480}$ at the optimal bias of 1.82 mA. Additional curves correspond to the BER scaled for reduced device size and power.

The BER of the AnotB gate at 6 GHz is shown as function of its flux bias $I_{flux}$. A 32-bit input pattern generated with an Anritsu MP1763C was split and applied to the inputs with a 15 bit relative shift, and the XOR output was compared to the correct pattern with an Anritsu MP1764C error detector. Error bars on the lowest points correspond to counting statistics of 4 errors (left) and 5 errors (right). Near the center, no errors detected for a period of 30 hours gives an error floor below $10^{-15}$ for the entire circuit. The data fit to the error function extrapolate to a minimum BER of $10^{-480}$ at the optimal bias of 1.82 mA. Additional curves correspond to the BER scaled for reduced device size and power.
error rate
\[ p = \frac{1}{4} \text{erfc} \left( \frac{\pm (I - I_c) / 20}{\sqrt{2} \delta I} \right), \tag{2} \]
where \( I - I_c \) is the distance of the flux bias current \( I \) from the error threshold \( I_c \), and \( \delta I \) is the root-mean-square noise current. The prefactor takes into account that for either error mode only half of the bits in the pattern contribute to the error rate. The factor of twenty represents the transfer function between applied flux bias and current induced in the loop containing the Josephson junctions. Numerical fits were obtained with less than 1% asymptotic standard error. On the left, \( I_c = 0.66 \) mA and \( \delta I = 1.02 \) \( \mu \)A; on the right, \( I_c = 3.04 \) mA and \( \delta I = 1.56 \) \( \mu \)A. At the optimal bias point of 1.82 mA, the data extrapolate to negligible BER for the gate under test. The lowest rate actually measured is below \( 10^{-15} \) for the entire circuit, including the output data link.

Discussion

The extrapolated minimum BER in our test indicates that device size and power could be scaled down still further. The error mechanism may involve either storage errors \([17]\), decision errors \([18]\), or even timing errors in an over-clocked circuit \([19]\). In all of these cases, noise current scales as the square root of the Josephson critical current, while the current scale of the error threshold goes linearly. Applying this scaling to Eq. 2 indicates that device size could be reduced by a factor of ten and still extrapolate to a minimum error rate of \( 10^{-44} \), which is negligible even for the most demanding applications including high-end computing. As a practical matter, BER below \( 10^{-44} \) could be maintained over a wide flux bias margin of \( \pm 30\% \) by scaling device size down by a factor of four. Measured noise current in our test is consistent with previous results for the gray zone of the RSFQ comparator \([20, 21]\), but with much lower bit error rates relative to RSFQ circuits \([22, 24]\) due to the larger operating margins in RQL.

The chip power scales linearly with number of junctions and frequency. We measured that with a 12.6 \( \mu \)W power supply, 800 junctions clocked at 6 GHz have a worst-case data-dependent power ratio of 0.91, corresponding to a variation in bias current amplitude of \( \pm 2\% \). Such a circuit scaled to \( 10^6 \) junctions and with a manageable maximum bias current variation of \( \pm 10\% \) would require a 6 mW power supply, amounting to only 15 mA of current on a 50 \( \Omega \) line. On the same circuit we measured a 1.4 ps worst-case variable clock propagation delay, independent of frequency. The given \( 10^6 \) junction circuit would correspond to twenty times less coupling to the clock line. Because variable clock delay scales as \( k^2 \) this circuit would have a timing variation of only 5 ps, or only \( \pm 1\% \) of the clock period at 6 GHz. The AC clock provides a stable clock reference that suppresses accumulative clock jitter, so ultimately clock frequency is limited by the switching time of the Josephson junction, which scales linearly with feature size. At the 0.8 \( \mu \)m lithography node, we can expect a 70 GHz maximum clock frequency \([23]\), or alternately a 6 GHz operation with twelve levels of logic per pipeline stage.

AC power distribution on-chip will benefit from the exceptional microwave properties of superconducting materials that have found applications ranging from single-photon qubit resonators with Q of \( 10^4 \) to THz dark matter detectors \([26, 27]\). Monolithic integration of RQL gates with microwave components, including power splitters, matching networks, and phase shifters, is a strength of the technology. Power dissipation in these passive components is as low as \( 1\% \) per wavelength \([28, 29]\), which would correspond to only 2.3% of the applied power in our shift register experiment. In the cryopackage, a 15 mA amplitude for the clock is consistent with low heat transport on the wires \([31]\). Very high clock rates up to 71 GHz have already been demonstrated for Josephson voltage standards using waveguides in the cryopackage \([31, 32]\).

Computational efficiency of the measured circuits are approaching \( 1000 k_B T \) with further reductions expected using smaller devices, giving unmatched efficiency in terms of operations per Joule. This means the technology offers a low energy solution for high end computing even after taking into account the overhead of the cryocooler, on the order of 1000 W/W at 4.2 K \([33]\). Because the 700 GHz energy gap in Nb makes superconductors inherently radiation-hard \([34]\), the technology may be useful for computationally intensive applications in space. Since device size and power can be scaled with temperature to remain in the noise-limited regime, the technology would be ideal for classical control, readout, and error-correction feedback for solid state qubits \([33]\) operating at millikelvin.

This work was supported in part by the Defense Microelectronics Activity under the Advanced Technology Support Program. The authors thank Marc Manheimer for discussion, John Fusco for administration, and acknowledge assistance from Isaac Carruthers with the software design environment, from Donald Miller and Steve Shauck with the design.

\[ \text{quentin.herr@ngc.com} \]

[1] Data Center Report to Congress, U.S. Environmental Protection Agency, www.energystar.gov, (2007).
[2] ITRS 2004 Update: Emerging Research Devices, Intern. Technol. Roadmap for Semicond., www.itrs.net/links/2004Update/2004_05_ERD.pdf (2004).
[3] Q.P. Herr, A.D. Smith and M.S. Wire, High speed data
