An Approach for Dense Superconducting Memories with Column Sense*

Paul Bradley and Perng-Fei Yuh
Hypres, Inc.
175 Clearbrook Rd.
Elmsford, N. Y., 10523

Summary

More than any other obstacle, the lack of a fast, dense random access memory (RAM) has hindered the development of Josephson digital electronics. The principal limitation has been the physical size of the superconductive memory cell which is almost always a loop of superconductor which stores a persistent circulating current as the data. This inductor together with the necessary switches and readout circuitry have typically been between one and two orders of magnitude larger than a semiconductor RAM using the same linewidths. This much lower density has meant that superconductive memories of a given capacity are much larger than the semiconductor alternative; this extra distance and time-of-flight for a memory access cancels the order of magnitude speed advantange the superconductive circuits enjoy.

Hypres has proposed, designed, and successfully tested a much smaller memory cell than any other in superconductive electronics. It is extremely simple and requires only one inductor and one Josephson junction per cell. A way to eliminate the switches that steer current and sharing the sense electronics make this possible. The circuit is described in figure 1. When the selection currents are applied to a cell, the flux state of this cell may be made to change to 0 or 1 fluxon. When the cell's state changes from 0 to 1 a Single Flux Quantum (SFQ) pulse is produced that is detected by the sense circuitry shared among several cells. The sense circuitry in turn regenerates the pulse and passes it to the next sense gate in the column and so on, until it arrives at the edge of the array without attenuation or distortion. This constitutes the read operation. No pulse would be produced if a 1 was written into a cell that already contained a 1 since the flux state would not change. The size of this cell implemented in 2 μm design rules is 35 x 17 μm^2, between half and a fourth the area of most superconductive designs. There is potential to shrink the cell even further, particularly in NbN at 10 Kelvins.

If this memory approach works well in large arrays, superconducting digital processing may enjoy a significant speed advantage over semiconductors. There exists a broad class of problems, one of which is real-time signal processing, where parallel computing is not effective and a much faster processor with faster memory could be used to great advantage.

*Research sponsored by SDIO/IST and managed by NRL

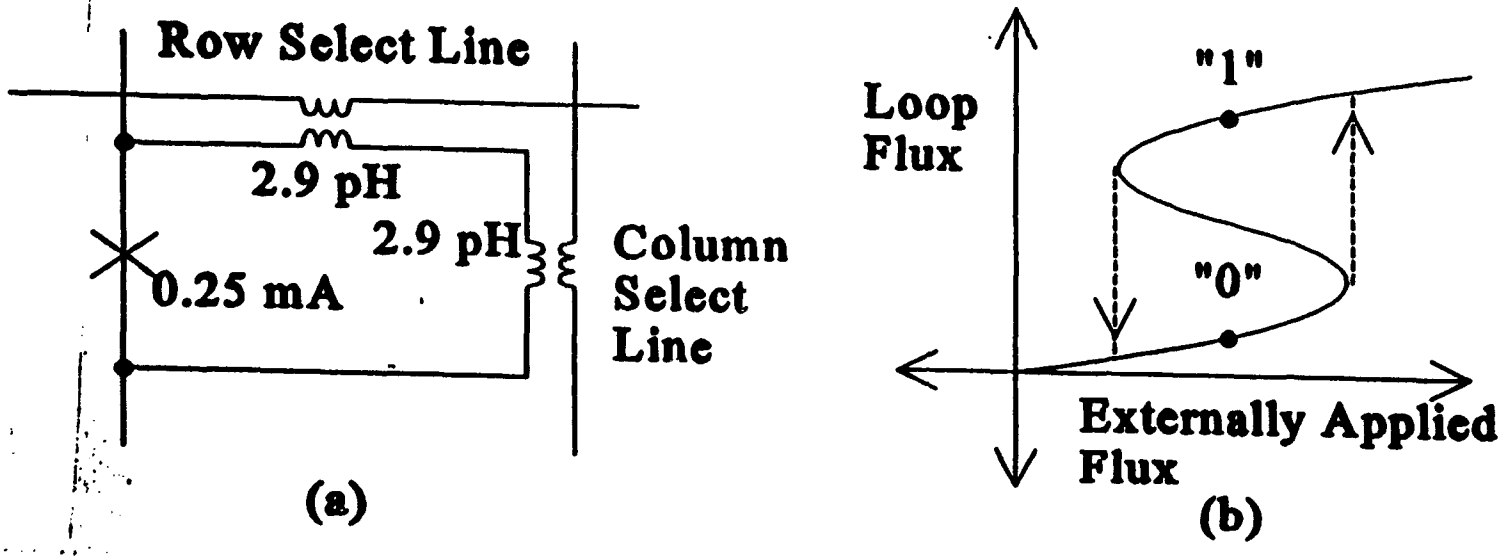

Fig. 1 (a) The single cell schematic circuit and (b) the relationship between the flux applied to the storage cell and the total flux in the loop (applied flux minus the flux due to the screening current induced in the loop). When the induced loop current reaches the critical current of the Josephson junction, a sudden jump between flux states occurs.

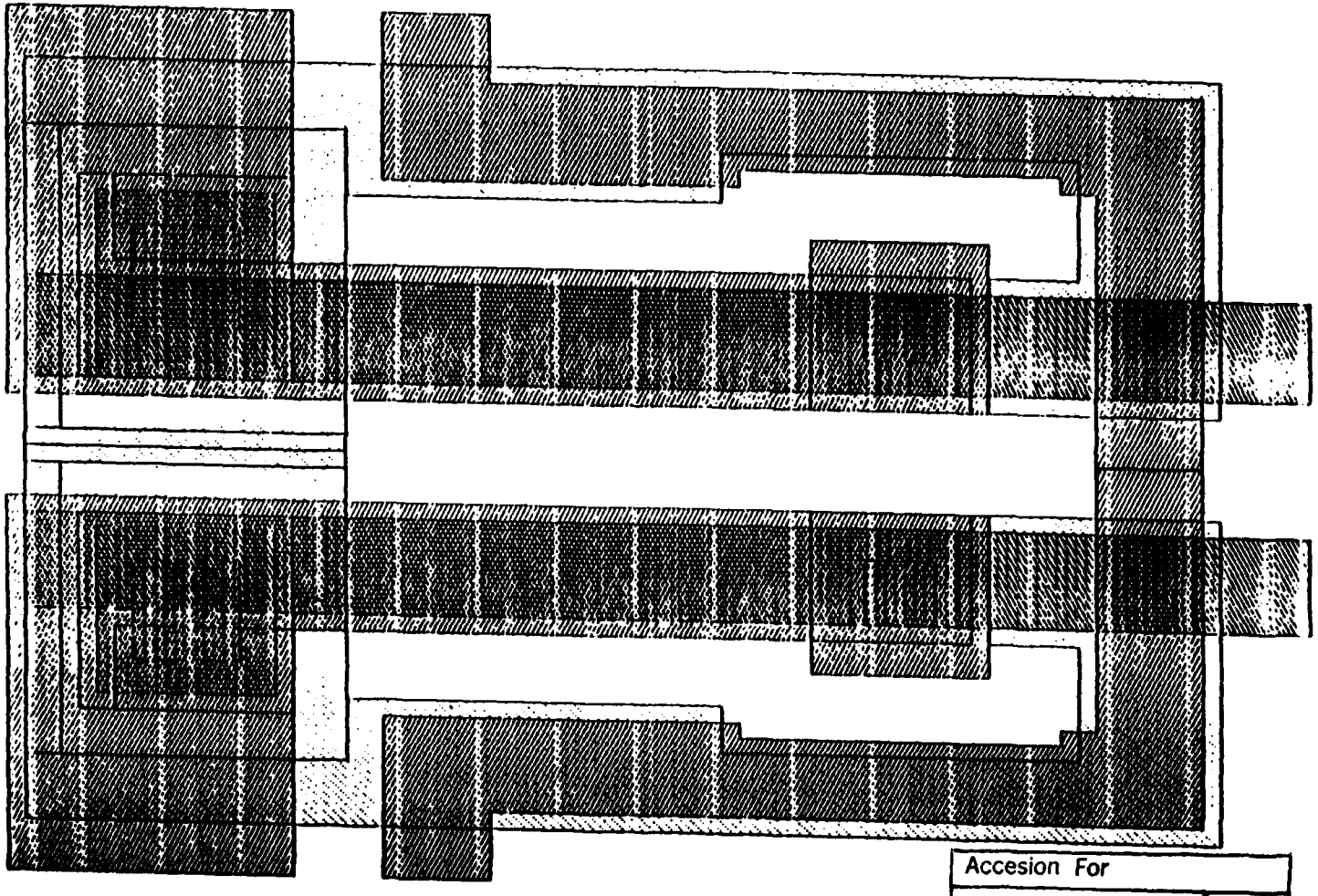

Fig. 2 The layout of two single junction memory cells. The area is dominated by the storage loop and selection lines transformer coupled to it.

DOD SBIR Phase I Program Final Report (Topic SDIO93-015)

An Approach for Dense Superconducting Memories with Column Sense

The primary goal of the Phase I proposal was to demonstrate feasibility of the single junction memory cell and get a good idea of the expected density and speed possible by actually building and testing the key circuits. The conclusion of the study is that it is a feasible design and that it overcomes the biggest problem of previous superconductive memory designs: circuit density. The tasks described in the Phase I work plan are addressed in order below (these tasks are included as the last page in this report for the reader's convenience).

The simplest superconducting circuit capable of storing information is a superconducting loop of wire with a persistent circulating current. It would as a consequence enclose a fixed amount of magnetic flux, but most memory must also be able to change its data or the amount of magnetic flux in the loop and this is not easily done without other circuit elements. Large currents or magnetic fields sufficient to drive the wire normal are necessary to change the state of the wire. However, the addition of a Josephson junction to the loop allows magnetic flux to pass in or out of the loop as its critical current is reached with relative ease.

A superconducting loop broken by one or more Josephson junctions is called a Superconducting Quantum Interference Device (SQUID) in analogy with optical interferometers because it exhibits periodic behavior as a function of the applied magnetic flux. The Meissner effect prevents magnetic fields from penetrating bulk superconductors under most conditions but does not prevent magnetic flux from passing through an area of weak superconductivity (the Josephson junction). This allows writing and erasing of digital data as flux in the loop. From simulations we have chosen the product of the loop inductance and the junction critical current for the cell to be about 3/4 of a flux quantum ($\Phi_0 = h/2e = 2.07 \times 10^{-15}$ Webers). In this case the single junction SQUID exhibits hysteretic behavior and may be used as a memory element.

Figure 1a shows the schematic circuit of a single cell including the row and column select lines used to choose which cell of an array is accessed. Figure 1b relates the magnetic flux contained in the SQUID loop to the magnetic flux applied externally by the selection line inductors. Note that it is not single-valued and that sudden transitions occur from one part of the curve to another because it is hysteretic; these transitions correspond to the read and write operations. For example, if we increase the applied flux enough to force the SQUID into the one fluxon state, we have written

a "1". If the cell already had one fluxon stored in it, then no fluxon slips through the junction and no voltage appears across the junction. If the cell had no fluxon stored in it, then one would be added and it would produce a Single Flux Quantum (SFQ) voltage pulse across the junction as it entered. The change of flux state from zero to one is detected when this pulse is sensed. Note that this is a destructive readout.

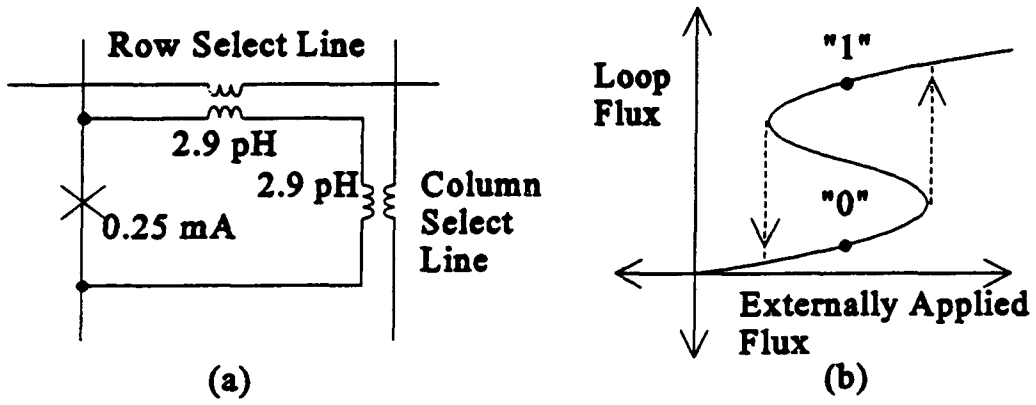

Fig. 1 (a) The single cell schematic circuit and (b) the relationship between the flux applied to the storage cell and the total flux in the loop (applied flux minus the flux due to the screening current induced in the loop). When the induced loop current reaches the critical current of the Josephson junction, a sudden jump between flux states occurs.

This new method of reading the contents of the cell has allowed us to reduce the number of circuit elements in the memory cell by more than a factor of two over earlier designs. A typical approach uses a multiple junction SQUID as part of the storage loop and another one to perform the read operation (e.g., see the IBM memory cell in Fig. 2).

The approach proposed by HYPRES allows using a single readout circuit to sense the contents of a number of memory cells rather than only one as in most schemes. The cells initiate the SFQ pulses that are sensed by the common sense gate and then regenerated periodically down a column of cells to the edge of the array in the same way that a Josephson Transmission Line (JTL) propagates SFQ pulses without attenuation or

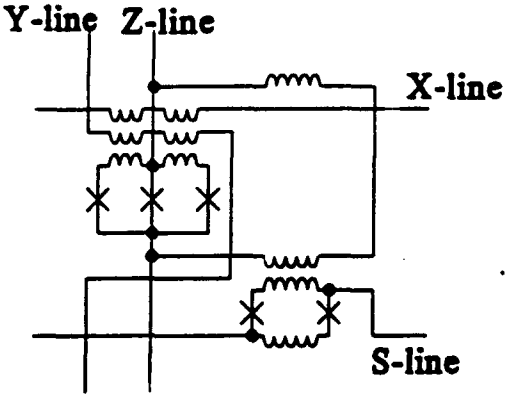

Fig. 2 The IBM cell with two SQUIDs, three selection lines (X, Y, and Z), a storage loop, and a sense line (S).

distortion. In fact, a number of memory cells in series are used as the inductor to form a JTL and the sense gate is nothing more than a shunted Josephson junction of the JTL. The maximum number of single junction memory cells per sense gate is set by the requirement that the JTL inductance times the junction critical current not be more than half of a flux quantum ($\Phi_0 = 2.07$ mA-pH). Since the critical current must be about 100 μA to avoid noise-induced switching, the total inductance must not be more than about 10 pH. Simulations using inductor values estimated from the layout of a memory cell indicate that eight cells per sense gate should be feasible(Fig. 3).

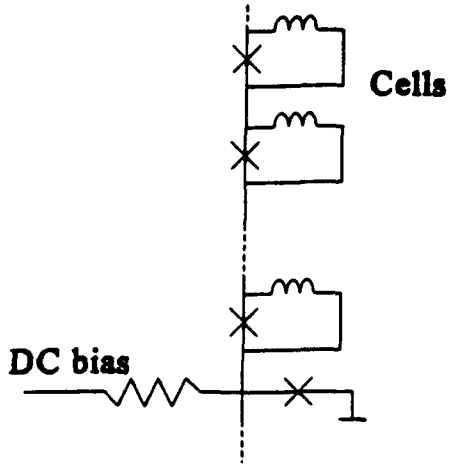

Fig. 3 An array of single junction memory cells shown connected to a sense gate. The parasitic inductances are not shown for clarity, although they do contribute significantly to the inductance of the array and ultimately limit the number of cells possible for each sense gate.

In order to minimize the loss of signal coupled to the readout at the edge of the array we used a direct injected Rapid Single Flux Quantum (RSFQ)[1] toggle flip-flop. When actual layouts were considered the transformer coupled circuits' coupling efficiency was seen to be low. The requirement that very small inductances be used in a transformer (over a ground plane) circuit means that we would couple only about half of the energy from the pulse into the sense circuit reducing noise margins. The rest is wasted on generating magnetic flux not coupled to the readout circuit. Among the various direct coupled circuits, signal to noise performance is expected to be similar to the RSFQ circuits, but the RSFQ circuits have the advantage of dc power.

The write, read, and rewrite control logic has been chosen to be RSFQ circuits, primarily because of the high speed, dc power, and compatibility with the memory cell array. Bias margins have been simulated to be about +/- 30% for most logic gates which should lead to high yield for this part of the circuit. It would be desirable to use RSFQ for all the circuits if possible, but for the row and column drivers we require more voltage than they can easily generate so some latching logic appears desirable to drive these lines quickly.

The other peripheral circuits would be comparable to those used in other memory designs and therefore would occupy about the same area. The most complex would be

the decoder circuits for interpreting the address to activate the proper row and column drivers of the array. A good candidate for this function is the so-called "loop decoder" developed originally at IBM for use in cache RAM and later at HYPRES. A 5-bit decoder has a simulated decode time of less than 300 ps, a power dissipation of 40 microwatts, and an area of about 0.5 mm x 1.2 mm making the successfully tested HYPRES design especially attractive (Fig. 4). A six bit decoder would have a little over twice this area. The decoder also has the property of being dc powered as does the RSFQ family of logic that is desirable for implementing the memory logic circuitry.

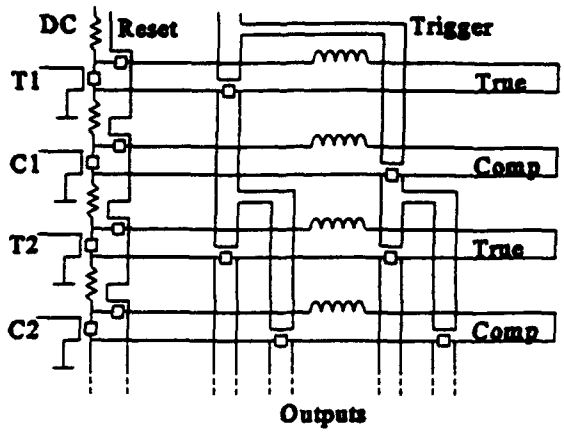

Figure 4 The first two stages of the loop decoder. The boxes are SQUIDs that act as switches triggered by current in the wires nearby. True and complement address bits are applied at left to establish circulating currents in one of each pair of the major loops. The trigger pulse is applied and just one SQUID in each pair of loops switches to yield current in only one of the output loops at the bottom.

An N-bit loop decoder is essentially N pairs of large superconducting loops connected to their neighbors in a binary tree. Each bit of the address is represented by a circulating current in either of two loops that are called the true and complement loops for that bit. A "1" for that bit of the address would be represented by a current stored in the true loop and a "0" by a current in the complement loop. Both true and complement loops may have their current switched into smaller loops, each of which controls two other switches (SQUIDs), one in each of the next pair of true and complement loops in the tree. Each of these switches shunt current into the smaller loops if and only if the major loop containing it has a circulating current as well as the minor loop from the previous stage that controls that switch.

In essence, once the address currents are established in the major loops, a trigger event starts a binary cascade of decision-making in the minor loops that switches only the output corresponding to the chosen address of 2^N possible outputs. The switch acts as an AND gate that can produce an output only if the large loop containing it has a circulating current as well as the smaller loop that controls it. Once this has switched a particular row or column driver of the memory array, the currents in the loops are turned off by a pulse in the reset line.

We may calculate the access time from the sum of the decode time, the propagation time of selection lines to the cell, the cell switching time, and the readout

time. It takes about 100 ps to generate the complement address bits and setup the circulating current in the major loops of the decoder and about 30 ps per stage for the decoding process once the trigger has initiated the decode. For a 4K RAM, we would need two six bit decoders (X and Y) which would simultaneously decode addresses in about 300ps. The extra time is required for propagation delay in the earlier stages of the decoder since it must extend across the width of the array (about 120 μm/ps). If we assume the array is about 2.5 mm in the longest direction (see next section) this would add 20ps (6 x 30ps + 20ps = 200ps). The driver circuit would add about 20ps of delay because we would prefer two gap voltages to drive the transmission lines to the cell. This reduces the width of the selection lines because a higher impedance is tolerable. The worst case time-of-flight from the driver to the cell would be another 20ps assuming we are able to use transmission lines of the proper impedance. The cell switching takes about 10ps and the data propagates out of the array in about 50ps on the JTL. Another 20ps is required to reach the corner of the array where the data would be buffered and sent off-chip. This buffer would require about 20ps as well. This brings the access time to a total of 430ps, possibly somewhat more depending on the size of the layout of the peripheral circuitry. Since this is a destructive readout memory the cycle time is just the access time added to the time necessary to write back the data just read. This should be about twice the access time except that the decode need not be repeated and a worst case propagation delay from the output to the drivers is added. This would give a cycle time of about 600ps. For larger arrays this would increase somewhat.

The memory cell layout is shown in Fig. 5 and is 17 μm x 35 μm. The area is dominated by the inductor of the memory storage loop which is also used for the transformers for the x and y selection lines. Three micron minimum width wires were used to design the cell and two micron vias were used as well as two micron minimum spacing between wires of the same layer. The area would be reduced approximately by the square of any improvement in linewidth, although driver voltages would need to be increased to maintain the proper selection currents. The sense circuit is approximately twice the cell area, but is not needed more than once for every 4 or 8 memory cells so most of the improvement in

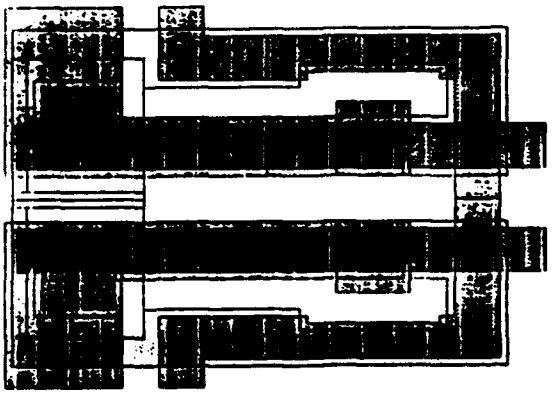

Figure 5 The layout of two single junction memory cells. The area is dominated by the storage loop and selection lines transformer coupled to it.

density of the cell is retained at the array level.

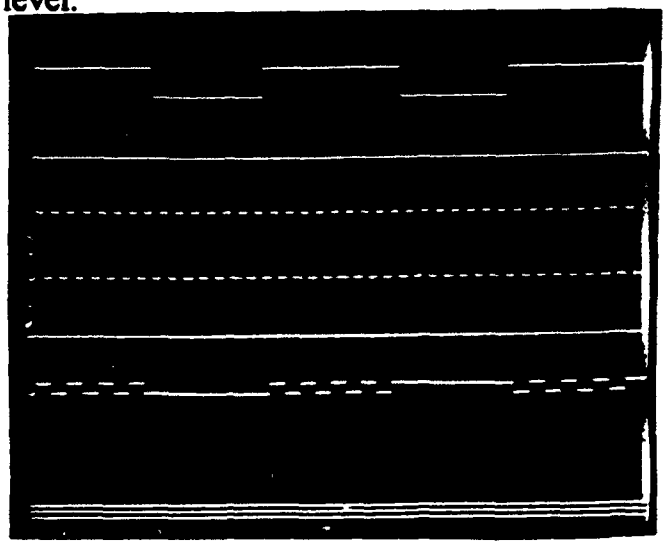

Figure 6 The correct operation of a single memory cell. The top trace is the sense gate bias, the second trace is the dc bias for the flip-flop, the third trace is the selection line current, and the fourth trace is the output of the flip-flop that receives the output from the sense gate. Note that every rising edge of the selection line current produces an SFQ pulse and toggles the flip-flop if the sense gate bias is high.

The memory cells were tested individually with the sense circuits and in arrays of 2, 4, and 8 cells with the sense circuitry. The individual cells worked well as shown in fig. 6. The top trace in the photo is the combination of the x and y selection lines. When a fluxon is stored in the cell during the rising edge of this waveform it generates a pulse that propagates through the sense circuits to the edge of the array and arrives at the flip-flop. The T flip-flop toggles if the pulse is received (the bottom trace). An advantage of the flip-flop is that it is possible to detect whether any errors occur during the entire time of testing since the flip-flop output will invert if a pulse is lost or spuriously generated. We detected no errors during the time of testing whenever input and bias currents were within the margins of operation. The allowable variations in each control or bias current were measured when all other currents were nominal and found to be +/- 18 % for the flip-flop dc bias and the combined x and y selection line. The allowable variation of the sense bias current was +/- 55 %. The two and four cell arrays were tested and also found to work. When the 8 cell array was tested it appeared to be working correctly part of the time, but the error rate was so high that not even a single sweep of the oscilloscope could be seen without errors. The failure mode was that the flip-flop failed to toggle part of the time it should have rather than that it toggled at times it should not have. This implies that the SFQ pulses were too small to be reliably detected in the eight cell array. Apparently, the current pulse was smaller than expected by the time it arrived at the sense gate, possibly due to the inductance of the 8 cells in series being too large. Another possible cause is that the pulse may have had the higher frequency components filtered out by propagation through the array yielding a wider, lower amplitude pulse.

In conclusion, we have demonstrated that the single junction memory cell works as expected, can be placed in arrays in which the sense circuitry is shared among several cells to reduce area and that the estimated performance of a 4K bit array is competitive with other superconductive designs. The principal advantage is that in

even larger arrays the small cell size prevents propagation delays from dominating the access and cycle times. This could have applications in supercomputing, digital signal processing, or any other application not well satisfied by slower parallel processing. A particularly intriguing possibility is the application of this idea to focal plane arrays at 10 Kelvins as in the Superconducting Infra-Red Focal Plane Array Sensor (SIRFPAS) program funded by BMDO. If implemented in niobium nitride technology, the memory cell would become even smaller because the larger penetration depth allows us to shorten the length of the inductor used as the storage element. In combination with superconducting analog-to-digital converters, the digital signal processing capability would allow calibration of infra-red detector arrays. Background subtraction or responsivity corrections to account for differences between detector elements are possible as well as more sophisticated signal processing to eliminate errors caused by gamma ray hits. The advantage of all this is that the signals would be digitized before leaving the sensor chip thereby avoiding interference and improving the signal-to-noise ratio of the detector. The number of wires necessary to communicate with the detector may be reduced by increasing the amount of signal processing on the detector chip.

This final report is the only deliverable of the Phase I program on dense superconducting memory. Each of the 8 tasks described in the work plan of the Phase I proposal are addressed in order in this report. For convenience the description of these tasks is repeated on the following page. The results will be published by P. F. Yuh in the Transactions on Applied Superconductivity.

E. PHASE I WORK PLAN

The work to be performed will consist of the following tasks:

Task 1: Design and simulation cells for column sense. The 1-SQUID and 1-JJ cell described in this proposal are to be carefully simulated. Additional improvement or new designs will be implemented. The cells will be simulated together with the proposed sense circuits. The number of cells one sense gate can sense will be determined.

Task 2: Design and simulation of sense circuits. The sense circuits described in this proposal are to be carefully simulated. These include a magnetically coupled sense gate, a direct injection gate, and an RSFQ sense gate. Additional improvement or new designs will be implemented. The sense circuits will be simulated together with the proposed cells.

Task 3: Design and simulate the write, read, and rewrite control logic circuits. A cache memory architecture will be chosen and designed. The control logic circuits will be designed and simulated.

Task 4: Design and simulate some important peripheral circuits. A decoder will be chosen and simulated. The results will be useful to determine the access time and cycle time for Task 5.

Task 5: Simulate the access time and cycle time of the chosen memory architecture. Important parameters of cache memory performance such as access time, cycle time, and power dissipation will be evaluated based on the designs of Task 1-4. Further designs of Task 1-4 will be iterated based on the results obtained in Task 5.

Task 6: Layout the final cell and/or sense circuit and determine the size of the cell. One or more cells will be laid out and its size determined. More simulation will be performed to include the parasitic of the layout.

Task 7: Fabricate and test the designed cell and/or sense circuits. The chips will be fabricated by the Hypres Nb process on part of a whole wafer to reduce the processing cost. The fabricated chip will be evaluated at low speed. HYPRES will provide all the necessary test equipment.

Task 8: Final Report and publication. Following completion of the test and evaluation phase, a final report will be written and submitted to the sponsor. The results will be also presented and/or published in a conference or Journal.