All-Silicon Quantum Computer

T. D. Ladd,* J. R. Goldman, F. Yamaguchi, and Y. Yamamoto[†]

Quantum Entanglement Project, ICORP, JST, Edward L. Ginzton Laboratory, Stanford University, Stanford, California 94305-4085

E. Abe and K. M. Itoh[‡]

Department of Applied Physics and Physico-Informatics, Keio University, Yokohama 223-8522, Japan

A solid-state implementation of a quantum computer composed entirely of silicon is proposed. Qubits are ^{29}Si nuclear spins arranged as chains in a ^{28}Si (spin-0) matrix with Larmor frequencies separated by a large magnetic field gradient. No impurity dopants or electrical contacts are needed. Initialization is accomplished by optical pumping, algorithmic cooling, and pseudo-pure state techniques. Magnetic resonance force microscopy is used for ensemble measurement.

The primary difficulty in the construction of quantum computers is the need to isolate the qubits from the environment to prevent decoherence, while still allowing initialization, control, and measurement. To date, the most successful experimental realizations of multi-qubit, many-gate quantum computers have used room-temperature, liquid NMR with "pseudo-pure" states [1]. These computers are able to maintain isolation from the control and measurement circuitry by employing weak measurement on a large ensemble of $\sim 10^{18}$ uncoupled, identical molecules. Although such a large, highly mixed ensemble may bring the existence of entanglement into question [2], the arbitrary unitary evolution afforded by the RF-controlled quantum gates assures that these computers behave nonclassically [3]. Their principal limitation results from their small initial nuclear polarization. The size of the effective subensemble of nuclei contributing to the pseudo-pure state, and hence the effective signal-to-noise ratio (SNR), decreases exponentially with each added qubit, leaving liquid NMR unlikely to exceed the 10-qubit level without substantial modification [4].

The proposals of Kane [5] and others [6] to use single nuclear spins in a low-temperature solid solve the scalability problem of solution NMR, but they introduce the problem of single-nuclear-spin measurement. It remains an experimental challenge to fabricate a structure in which individual nuclei are sufficiently coupled to an electronic system for single-spin measurement, but also sufficiently isolated for long coherence times.

In this Letter, we propose a different solid-state NMR implementation of quantum computation which introduces electron-mediated cooling, but maintains the weak ensemble measurement that has made solution NMR quantum computers so successful. The device is made entirely of silicon, with no electrical gates or impurities. As will be discussed below, the qubits are spin-$1/2$ nuclei that are located in relatively isolated atomic chains, as shown in Fig. 1. The nuclei within each chain are distinguished by a large magnetic field gradient created with a nearby microfabricated ferromagnet [7]. Each nucleus has about 10^5

ensemble copies in a plane orthogonal to its chain. This structure is embedded in a thin bridge whose oscillations provide readout via magnetic resonance force microscopy (MRFM) [8,9].

The advantages of an all-silicon implementation are many. Foremost, the crystal growth and processing technology for silicon are highly matured. In particular, the most sensitive structures for force detection have been made from pure silicon. Also, it is good fortune that the family of stable nuclear isotopes of silicon is quite simple: 95.33% of natural silicon is ^{28}Si or ^{30}Si, which are both spin 0, and 4.67% is ^{29}Si, which is spin $1/2$, perfect for the qubit. Thus, silicon is well suited for nuclear-spin isotope engineering [10]. Another crucial motivation for the choice of silicon is the observation that nuclei in a
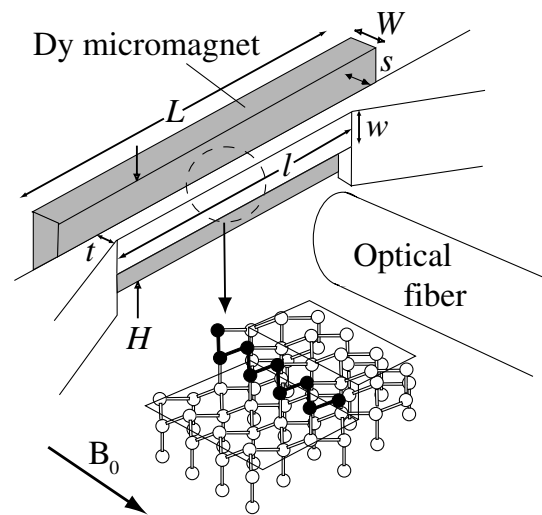

FIG. 1. The integrated micromagnet and bridge structure. The bridge has length $l = 300$ μm, width $w = 4$ μm, and thickness $t = 0.25$ μm. The micromagnet, a distance $s = 2.1$ μm from the bridge, has length $L = 400$ μm, width $W = 4$ μm, and height $H = 10$ μm. The inset shows the structure of the silicon matrix and the terrace edge. The darkened spheres represent the ^{29}Si nuclei, which preferentially bind at the edge of the Si step.

semiconductor may be polarized by cross relaxation with optically excited, spin-polarized conduction electrons [11]. Although there are many other means for dynamic nuclear polarization, optical pumping in semiconductors has one important feature: the electrons whose hyperfine couplings make the polarization possible recombine shortly thereafter and, hence, do not contribute to decoherence during the computation. The absence of any spurious spins, nuclear or electronic, should leave the ^{29}Si nuclei well decoupled from their environment.

In the following, we describe an example procedure for fabricating the structures of Fig. 1 [12]. We start with an isotopically depleted ^{28}Si(111) ($<0.05\%$ ^{29}Si) wafer which has been miscut $1°$ towards ($\bar{1}\bar{1}2$) and placed in a multichamber molecular beam epitaxy machine equipped with a scanning tunneling microscope (STM) and a preheating stage. Reference [13] demonstrates that highly regular arrays of surface steps can be produced on vicinal Si(111)-(7×7) using a simple annealing sequence. Because of the high energy of interrupting or misplacing 7×7 domains, this procedure leads to atomically straight step edges along the ($1\bar{1}0$) direction for up to 2×10^4 lattice sites. The average terrace width is ~15 nm. Once we verify the straight step edges using the STM, we transfer the wafer to the growth chamber to form atomic chains of ^{29}Si using the step-flow mode at temperatures $T \sim 600$ °C [14]. Arrival of ^{29}Si isotopes at the substrate surface induces a surface transition from 7×7 to 1×1 [15]; i.e., ^{29}Si isotopes travel to the step edges and form atomically straight lines. We terminate the evaporation of ^{29}Si when the atomic chains are one atom wide and run completely along the step [16]. A ^{28}Si capping layer of 15 nm is grown on top of ^{29}Si chains, and we repeat the same multistep annealing, ^{29}Si chain growth, and capping sequence to produce replicas of parallel ^{29}Si chains [17]. The last step in the growth process is the capping of ensembles of ^{29}Si chains by a thick ^{28}Si layer.

An oxide layer is buried beneath the ^{29}Si wire block by layer-transfer techniques. The bridge is created with e-beam lithography, plasma etching, and selective chemical etching of the oxide layer. The dysprosium (Dy) micromagnet is fabricated using evaporative deposition and lift-off.

Calculations similar to those in Ref. [7] show that the magnetic field gradient due to the micromagnet is $\partial B^z/\partial z = 1.4$ T/μm, which is persistent over the thickness of the bridge and is superposed with a large homogeneous field B_0 of ~7 T. The distance in the z direction between two ^{29}Si nuclei in an atomic chain, which we notate a, is 1.9 Å, so the gradient leads to a qubit-qubit frequency difference of $\Delta\omega = a\gamma\partial B_z/\partial z = 2\pi \times 2$ kHz. The active region is a 100 μm by 0.2 μm area in the center of the bridge, containing $N = 10^5$ chains persisting over the bridge thickness. This active region is sufficiently small and the magnetic field sufficiently homogeneous that all equivalent qubits in an atomic plane lie within a bandwidth of 0.6 kHz.

For initialization of the quantum computer, we propose to employ optical pumping, algorithmic cooling, and pseudo-pure state techniques. The premise of optical pumping is that nuclei exchange Zéeman energy with a bath of electrons which have been preferentially excited into a single spin state by circularly polarized light. The nuclei thereby relax thermally to an effective spin temperature corresponding to the nonequilibrium electron-spin polarization. Once those electrons recombine, the nuclei retain their spin polarization for the "dark" T_1 time, which is extremely long (200 h in Ref. [11]). In low-field (~1 G) experiments at 77 K [11,18], nuclear polarizations have not exceeded 0.1% due to limited electron spin polarizations and long recombination times, a result of silicon's indirect band gap. Improved nuclear polarization in silicon may be observable in higher magnetic fields (~10 T) and lower temperatures (~4 K), and in silicon nanostructures where rapid recombination of electrons via surface states can help maintain the electron-spin polarization [19].

The physical cooling afforded by optical pumping is well suited to complement an algorithmic cooling technique introduced by Schulman and Vazirani [20,21]. This technique redistributes the entropy among a register of qubits to a subregister that is then discarded (decoupled and ignored) [22]. This method is expensive; it takes time to perform the (classical) logic operations and, worse, it sacrifices many qubits. An initial register of size n_0 will, for small initial polarization p_0, shrink to size $n_0 p_0^2/2\ln2$ if the procedure is taken to the entropy limit. However, the very long T_1 affords ample time for the procedure, and the number of available qubits in our configuration before initialization can be thousands. Moreover, the algorithm need not be taken to the entropy limit, since large but still incomplete polarizations can be handled with pseudo-pure state techniques, the consequences of which will be discussed below in the context of SNR.

The secular component of the dipolar Hamiltonian which couples the ith spin to the jth spin within one chain is written [23]

$$\hat{\mathcal{H}}_{ij} = \frac{\mu_0}{4\pi} \gamma^2 \hbar^2 \frac{1 - 3\cos^2\theta_{ij}}{r_{ij}^3} \hat{I}_i^z \hat{I}_j^z$$
$$\equiv -\hbar\delta\omega_{ij}\hat{I}_i^z\hat{I}_j^z, \tag{1}$$

where r_{ij} is the length of the vector connecting the spins and θ_{ij} is its angle with the applied field. Nearest neighbors along the proposed atomic chains are not exactly parallel to ($1\bar{1}0$), but rather zigzag with angle $\theta_{i,i+1}$ satisfying $\cos^2\theta_{i,i+1} = 2/3$, leaving $\delta\omega \equiv \delta\omega_{i,i+1} = 2\pi \times 0.4$ kHz. Other terms of the dipolar Hamiltonian require the exchange of energy in the amount of $\hbar\Delta\omega$ or more. The long T_1 in silicon indicates the inefficiency by which this energy may be exchanged with degrees of freedom external to the nuclei. In strongly coupled dipolar systems, this energy can be compensated for by the dipolar bath of identical nuclei [24]. In the present scheme, the members of the nuclear ensemble are so far removed that this

energy bath is absent. Hence, we assume these nonsecular terms are well suppressed. The Hamiltonian of Eq. (1) may be "switched off" by applying a periodic succession of narrow band π pulses at, for instance, the ith resonant frequency [25]. Simultaneous decoupling of more than two qubits may be accomplished by timing the selective π pulses according to the entries of an appropriately sized Hadamard matrix; a pair of qubits may be selectively recoupled in order to implement two-bit gates [26]. Note that this Hadamard pulse scheme serves the second purpose of refocusing inhomogeneous broadening caused by the in-plane nonuniformity of the field gradient and any small bulk susceptibility effects.

The presence of a large magnetic field gradient provides a natural means for performing MRFM on a magnetization M^z, since this technique is sensitive to the gradient force given by $F^z = M^z \partial B^z / \partial z$. The experiment is performed in high vacuum ($<10^{-5}$ torr) and at low temperatures (~ 4 K). A coil is used to generate the RF pulses for logic operations and decoupling sequences; it also generates the continuous-wave radiation for readout. An optical fiber-based displacement sensor is used to monitor deflection of the bridge using interferometry. Sub-Ångstrom oscillations can be detected; larger oscillations can be damped with active feedback which avoids additional broadening while maintaining high sensitivity [27].

Readout is performed using cyclic adiabatic inversion [23], which modulates the magnetization of a plane of nuclei at a frequency near or on resonance with the bridge. The spins of resonant frequency ω_i are irradiated with the RF field $B^x = 2B_1 \cos\{\omega_i t - (\Omega/\omega_m)\cos(\omega_m t)\}$, where ω_m is the modulation frequency chosen to be near the resonance of bridge oscillations, and Ω is the frequency excursion, which should be much smaller than $\Delta\omega$ [9]. The z component of the ith plane's magnetization is deduced from the phase of the resulting bridge oscillation. Simultaneous detection of signals from multiple planes is possible if the different planes to be measured are driven at distinct modulation frequencies ω_m.

The force resolution for MRFM is limited by thermal fluctuations of the mechanical oscillator [28]. Force resolutions of 5.6×10^{-18} N/$\sqrt{\text{Hz}}$ have been reported for single crystal silicon cantilevers at 4 K [29]. The thermal noise threshold of the bridge structure in Fig. 1 is estimated to be $\sim 1.2 \times 10^{-17}$ N/$\sqrt{\text{Hz}}$ based on a lumped harmonic oscillator model and assuming a modest quality factor Q of 10^4 [30]. This model yields an estimated spring constant of $k \approx 0.0042$ N/m and a resonance frequency $\omega_c / 2\pi \approx 23$ kHz. The detectable signal will depend upon the initial polarization p after optical pumping and algorithmic cooling. The force from the subensemble magnetization corresponding to a pseudo-pure state is estimated [4] as

$$ F^z = \frac{\hbar\Delta\omega}{2a} N \left[\left(\frac{1 + p}{2} \right)^n - \left(\frac{1 - p}{2} \right)^n \right]. \quad (2) $$

The number of qubits available in this scheme may be found by maximizing n in Eq. (2), constrained by the criterion that the force exceeds the thermal noise threshold; the results of such maximization are shown in Fig. 2. At low p, exponential improvements in p are needed to increase the number of measurable qubits n. Once p exceeds about 60%, however, we find $n \sim (1 + p)/(1 - p)$, escaping the exponential downscaling which plagues solution NMR.

There are several possible sources of decoherence in this proposal: magnetic fluctuations in the dysprosium, thermal currents in the dysprosium, fluctuations of paramagnetic impurities in the silicon, thermal motion of the bridge, and uncontrolled dipolar couplings between nuclei. We estimate that the latter two sources are the most important, so we limit our present discussion to them.

A calculation using the thermal noise statistics of a high-sensitivity mechanical oscillator [28] estimates the T_2 time scale due to the bridge's thermal drift as $T_2^c = k\omega_c Q a^2 / \Delta\omega^2 k_B T \approx 25$ s. Active feedback stabilization is expected to increase this timescale by as many as 4 orders of magnitude, bringing it close to the lengthy T_1 time scale.

There are two sorts of uncontrolled dipolar couplings to consider. There are the in-plane couplings, in which the participating nuclei have equal Larmor frequencies and can thus participate in spin flip-flop processes. Since the average distance between chains is ~ 15 nm, these processes cause decoherence on a time scale of order $T_2^h \sim 4\pi(15 \text{ nm})^3 / \gamma^2 \hbar \mu_0 \sim 100$ s. This already lengthy time scale may be increased by the addition of dipolar refocusing sequences such as WAHUHA [25].

There will also be spurious couplings between nuclei in one homogeneous plane and copies of nuclei in the next. Such couplings occur when two qubits are recoupled for a logic gate, and they cause a small error in each logic gate. To estimate this error, suppose we recouple one qubit to another qubit m planes away. The strongest coupling seen by a nucleus is the in-chain coupling, and its rate is $\sim \delta\omega/m^3$. The couplings to all of the neighbors may be treated as a T_2 decoherence process, with
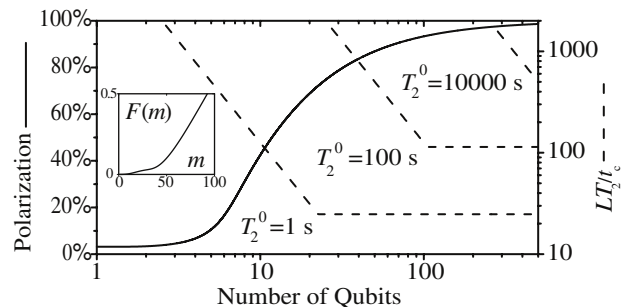

FIG. 2. A plot showing the scalability of the present scheme. The solid curve, corresponding to the left axis, shows the polarization p needed in order for a number of qubits n to be measurable. The dashed lines, corresponding to the right axis, plot the number of logic gates (decoherence time T_2 divided by the pulse sequence cycle time t_c) times the length of a "block" of the decoupling sequence L against n for several values of T_2^0.

$$\left(\frac{1}{T_{2m}^{\text{r}}}\right)^2 = \frac{1}{16}\left(\frac{\delta\omega}{m^3}\right)^2 \sum_i \frac{(\lambda_i^2/m^2 - 2)^2}{(\lambda_i^2/m^2 + 1)^5}. \qquad (3)$$

Here T_2 has been estimated as the inverse square root of the second moment [25] and λ_i is the lateral distance to the ith chain normalized by a. The ratio of the approximate gate time $m^3/\delta\omega$ to T_{2m}^{r} is plotted in the inset of Fig. 2. This function, $F(m)$, represents the approximate error in the gate when attempting to couple qubits m planes apart. To keep gate errors low, the computer should couple only nearby neighbors and handle more distant couplings by bit swapping. In this way, gate errors due to unrefocused nuclear couplings between chains is limited to approximately $F(1) \approx 10^{-6}$.

These cross couplings also influence the clock speed of this computer, as we now explain. When qubits are decoupled by the Hadamard scheme mentioned above, the resulting pulse sequence has a clock time $t_c = Ln^2/\Delta\omega$, where n is the number of qubits being decoupled and $L/\Delta\omega$ is the amount of time devoted to one π pulse [26]. As qubits are added, the amount of time needed to recouple them grows. For very large n, however, some qubits in the chain become so distant that the r_{ij}^3 factor in Eq. (1) renders their interaction negligibly small. In this case the Hadamard pulse scheme can be truncated to decouple qubits only in sets of l, and a new decoherence process is introduced with $T_{2l}^{\text{t}}\delta\omega = l^3[1 + F^2(l)]^{-1/2}$. If we compare the total decoherence time $T_2 = (1/T_2^0 + 1/T_{2l}^{\text{t}})^{-1}$, where T_2^0 combines all other decoherence processes, to the clock speed t_c, we find that there is an optimum l at which to truncate the Hadamard pulse scheme. The effective number of logic gates T_2/t_c, therefore, at first decreases as n^2 and then flattens once n reaches this optimum l, as shown in Fig. 2.

By consulting Fig. 2, we see that this scheme can, for sufficiently large T_2^0 and sufficiently high polarizations, allow substantially more qubits than solution NMR, without the need for single-spin measurement or unrealistic advances in fabrication, measurement, or control technologies.

*Electronic address: tladd@stanford.edu
†Also at NTT Basic Research Laboratories, 3-1 Morinosato-Wakamiya Atsugi, Kanagawa, 243-0198, Japan.

‡Also at PRESTO-JST.
[1] N. A. Gershenfeld and I. Chuang, Science **275**, 350 (1997); D. G. Cory *et al.,* Proc. Natl. Acad. Sci. U.S.A. **94**, 1634 (1997); E. Knill *et al.,* Phys. Rev. A **57**, 3348 (1998).
[2] S. L. Braunstein *et al.,* Phys. Rev. Lett. **83**, 1054 (1999).
[3] R. Schack and C. M. Caves, Phys. Rev. A **60**, 4354 (1999).
[4] W. S. Warren, Science **277**, 1688 (1997).
[5] B. E. Kane, Nature (London) **393**, 133 (1998).
[6] For example, G. P. Berman *et al.,* Phys. Rev. Lett. **87**, 097902 (2001), and references therein.
[7] J. R. Goldman *et al.,* Appl. Phys. A **71**, 11 (2000).
[8] J. A. Sidles, Appl. Phys. Lett. **58**, 2854 (1991).
[9] D. Rugar *et al.,* Nature (London) **360**, 563 (1992).
[10] K. M. Itoh and E. E. Haller, Physica (Amsterdam) **10E**, 463 (2001).
[11] G. Lampel, Phys. Rev. Lett. **20**, 491 (1968).
[12] A similar fabrication scheme has been proposed for a Kane-type [5] quantum computer: I. Shlimak *et al.,* J. Phys. Condens. Matter **13**, 6059 (2001).
[13] J. Viernow *et al.,* Appl. Phys. Lett. **72**, 948 (1998); J.-L. Lin *et al.,* J. Appl. Phys. **84**, 255 (1998).
[14] C. K. Ming *et al.,* Acta. Phys. Sin. **39**, 1945 (1990).
[15] A. V. Latyshev *et al.,* Phys. Status Solidi A **113**, 421 (1989).
[16] Accidental ^{28}Si or ^{30}Si isotopes in the ^{29}Si source will lead to chain interruptions; this will reduce the effective ensemble size unless the impurity level of the source is much less than $1/n$, where n is the number of qubits.
[17] The diffusion of the buried silicon chains is negligible at the $\sim 600\,°$C growth temperature, according to the measurement of the silicon self-diffusion constant in H. Bracht *et al.,* Phys. Rev. Lett. **81**, 393 (1998).
[18] N. T. Bagraev *et al.,* JETP Lett. **25**, 190 (1977), and references therein.
[19] C. Delerue *et al., Light Emission in Silicon,* edited by D. J. Lockwood (Academic, New York, 1998), Chap. 7.
[20] L. J. Schulman and U. V. Vazirani, in *Proceedings of the 31st ACM Symposium on Theory of Computing* (Association for Computing Machinery, New York, 2001), p. 322; D. E. Chang *et al.,* Chem. Phys. Lett. **338**, 337 (2001).
[21] P. O. Boykin *et al.,* e-print quant-ph/0106093.
[22] Alternatively, the "hot" qubits could be recooled if their T_1 can be made substantially less than the other qubits. This allows the restrictive entropy limit to be escaped [21]. Such rapid cooling may be possible in this proposal if the optical pumping field can be limited to a very small (~ 100 nm) depth from the silicon surface.
[23] A. Abragam, *Principles of Nuclear Magnetism* (Oxford University Press, Oxford, U.K., 1961).
[24] A. Z. Genack and A. G. Redfield, Phys. Rev. Lett. **31**, 1204 (1973).
[25] M. Mehring, *High Resolution NMR in Solids* (Springer-Verlag, Berlin, 1983).
[26] D. W. Leung *et al.,* Phys. Rev. A **61**, 042310 (1999).
[27] U. Dürig *et al.,* J. Appl. Phys. **82**, 3641 (1997).
[28] T. B. Gabrielson, IEEE Trans. Electron Devices **40**, 903 (1993).
[29] T. D. Stowe *et al.,* Appl. Phys. Lett. **71**, 288 (1997).
[30] F. R. Blom *et al.,* J. Vac. Sci. Technol. B **10**, 19 (1992).