

## Scaling limit of digital circuits due to thermal noise

Kenji Natori<sup>a)</sup> and Nobuyuki Sano

*Institute of Applied Physics, University of Tsukuba, Tsukuba, Ibaraki 305, Japan*

(Received 24 November 1997; accepted for publication 16 February 1998)

The error probability at a node of a digital circuit exposed to thermal noise agitation is investigated and the minimal dissipation–reliability relation for practical electronic circuits is derived. The digital circuit is modeled by an inverter chain with ideal transfer characteristics, and the error probability due to spurious data transfer caused by the thermal noise fluctuation is evaluated as a function of the node switching energy. The maximal error probability at each node allowed by the reliability requirement of the total system leads us to the minimal node energy dissipated per logical switching, which amounts to around 12 eV in the future  $10^{10}$  gate system operated at a 10 GHz clock rate with a  $10^4$  FIT level reliability. In view of the device size-scaling trend of large-scale integrated circuits, the minimal node energy is expected to be reached at a feature size of 10–20 nm. © 1998 American Institute of Physics. [S0021-8979(98)07810-4]

### I. INTRODUCTION

The downsizing of devices in metal–oxide–semiconductor (MOS) large-scale integrated circuits (LSIs) has continually evolved so that a gigascale integration of the tens-of-nanometer-size device will be achieved at the beginning of the next century. The scaling limit of the device size is discussed in various aspects of technology,<sup>1,2</sup> but no clear boundary has so far been pointed out.

What is the minimal energy for an information processing system to carry out a logic operation? This is an old, and yet a new question. A large quantity of discussion has been carried out, mainly, from the fundamental point of view with the use of conceptual devices.<sup>3,1</sup> But the discussion based on realistic circuit systems has been little presented, although the device size is approaching the scaling limit.

In principle, dissipation-less computing is possible as is shown by Bennett.<sup>4</sup> In a computer that consists of a “reversible logic gate,”<sup>5</sup> where information is not discarded in the process of logical switching and the input information is generated from the output information by reversely driving the gate, the forward flow of logic operation and the backward one can be made in a state of equilibrium without dissipation when not driven. Introduction of a potential slope will drive the system forward to yield the output, at the same time causing dissipation of energy corresponding to the potential difference between the input and the output of the system. The slope can be made arbitrarily small by use of a small driving force, at the expense of slow computation. The required dissipation can be made infinitesimal, but not zero. The logic operation in a realistic computing system shows completely different aspects. We need to obtain the correct result at the output terminal of the system when we check there after a passage of definite time. The error probability of the result, which should be extremely small, may depend on the force to drive the system toward the correct direction resisting the thermal agitation. An infinitesimal dissipation is

enough to drive the system to the direction at all, but a finite quantity of energy may be necessary to get a solution within a certain period with the required reliability.

A realistic computing system usually consists of irreversible logic gates; a part of information unnecessary for succeeding progression is discarded at the logic gate and only the necessary part is transferred. These information reduction processes are shown to bring about an energy dissipation of order  $kT$  per logical step,<sup>6</sup> where  $k$  is the Boltzmann constant and  $T$  the temperature. The entropy increase accompanied by the irreversible process brings about that amount of dissipation. Keyes and Landauer<sup>7</sup> showed that the dissipation is  $kT \ln 2$ , assuming an infinitely slow operation in their time-modulated potential-well model. The dissipation–reliability problem has also been discussed. In his negentropy discussion, Brillouin<sup>8</sup> showed that with use of a specific harmonic oscillator model the entropy increase  $\Delta S = k \ln r$  is inevitable in a measurement operation resisting thermal agitation.  $r$  is the reliability of the result defined by the inverse of the error probability. The minimal entropy increase corresponding to the maximal error probability  $1/2$  is given by  $k \ln 2$ . Neyman<sup>9</sup> argued that the elementary logic operation may be considered as the elementary measurement of a binary quantity and extended the above result to  $\Delta E \geq kT \ln r \Delta I$ , where the energy dissipation of the digital information processing per unit information quantity is expressed by  $\Delta E / \Delta I$ . A more reliable result requires a larger dissipation. Landauer and Woo<sup>10</sup> also recognized a similar trade-off suggested by Neyman under restricted conditions in their time-modulated viscous potential-well model operated at arbitrary velocity.

This paper discusses the dissipation–reliability problem of an electronic digital system. The discussion does not deal with the fundamental problem in terms of the conceptual device but applies to the practical electronic circuits, and the minimal dissipation–error probability relation for the system is presented. In Sec. II, we first discuss the model for the unit of logic operation in electronic circuits, then analyze the thermal noise at the circuit node, and finally show the mini-

<sup>a)</sup>Electronic mail: natori@hermes.esys.tsukuba.ac.jp

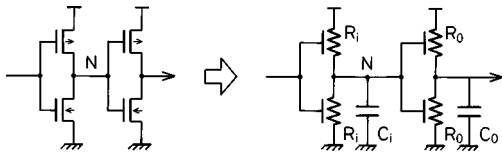


FIG. 1. A part of a CMOS inverter chain that is a representative of the digital circuit. Further, the MOS transistors are regarded as effective resistors  $R_i$  and  $R_o$  whose resistance is controlled by the gate potential. The circuit nodes have node capacitances  $C_i$  and  $C_o$ .

imum dissipation per logic operation required by the reliability of the whole system. Section III is the discussion and Sec. IV is the conclusion.

## II. ANALYSIS

### A. Modeling of a digital circuit

As an example of the digital circuit, we take up a complementary metal-oxide-semiconductor (CMOS) inverter chain shown in Fig. 1. Any logic gate in a digital circuit can be represented by an inverter in the present discussion without loss of generality. The circuit node is charged up and discharged by metal-oxide-semiconductor field-effect transistors (MOSFETs) and is clamped to an external potential level outside the transition time. MOSFETs are regarded as effective resistors whose resistance is controlled by the input signal fed to the gate electrode. We take up an inverter whose input is tied to the node  $N$ , output has a capacitance  $C_o$ , and both the load and the driver have a common effective on-resistance  $R_o$ . The input node  $N$  has the capacitance  $C_i$ , and is connected to the other pair of the load and driver with the common effective on-resistance  $R_i$ . We assume an ideal transfer characteristic curve for these inverters as is shown in Fig. 2. When the input potential level is less than  $V/2$ , where  $V$  is the supply voltage, the output is  $V$ , while the output changes to 0 if the input is raised over  $V/2$ . The potential level of the node  $N$  changes in response to the transition of the previous node, and is clamped to either the supply voltage or ground (GND) level for most of the time except for a short switching time. All the while, the level suffers from the perpetual level fluctuation denoted by  $v$  due to the thermal noise. First, let us take up a circuit node statically retaining the data being clamped to a potential level. When a large fluctuation with  $v > V/2$  ( $v < -V/2$ ) continues in the node clamped to GND (supply voltage) for a

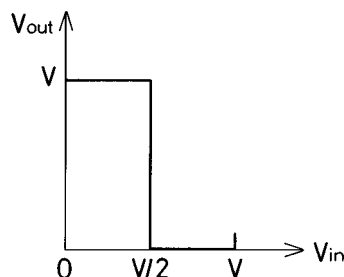


FIG. 2. The ideal transfer characteristics of a CMOS inverter. The threshold voltage is at one-half of the supply voltage.

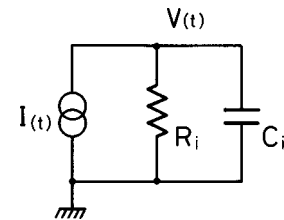


FIG. 3. An equivalent circuit of circuit node  $N$ , whose potential is clamped to the GND level.  $I(t)$  is the thermal noise expressed by a current source and  $v(t)$  is the level fluctuation due to it.

time longer than the time constant  $R_o C_o$  of the inverter, a spurious level change is transferred to the next node. The return to the correct level with  $|v| < V/2$  in the next phase may sometimes succeed in suppression of the error spreading, but a circuit error results if the preceding spurious data transfer slips through. We define, here, the spurious data transfer to the next node as the generation of a circuit error. Next, we consider a circuit node at the moment of logical switching. The node is expected to perform the operation at a given instant within a period comparable to the time constant of the gate, which is tightly designed along the critical path of the signal flow to achieve high-speed operation. In this case also, a large fluctuation with  $v > V/2$  ( $v < -V/2$ ) that is superimposed on the normal level transition, and that continues longer than the time constant, causes a false high (low) level and transmits a spurious signal to the next node. A correct level restored at the shifted timing will confuse the subsequent logic operation and possibly cause a circuit error.

Thus, we assume two points. One is that the circuit node is charged up or discharged through effective resistors, and is clamped to a fixed potential level except for the short transition time. The other is that a circuit error is generated when a level fluctuation larger than  $V/2$  continues longer than the time constant of the node. The discussion is not restricted to the CMOS circuit but applies to any digital circuits where these two points are satisfied.

### B. Thermal noise

Microscopically, the effective resistor that constitutes the circuit includes quite a large number of carriers inside, and the velocity  $\mathbf{u}$  of these carriers at a given instant  $t$  is distributed according to Maxwell's law of velocity distribution (Gaussian distribution). The current  $I$  through the resistor is expressed as

$$I = \frac{1}{L} \sum q u_x, \quad (1)$$

where the  $x$  direction is along the current flow,  $L$  is the length of the resistor,  $q$  is the carrier charge, and the summation is over all the carriers in the resistor.  $I$  is a stochastic quantity because it depends on the distribution of  $u_x$ . When node  $N$  is clamped to a potential level, the mean value of  $I$  is 0 but it is distributed around the value due to the stochastic distribution of  $u_x$ . Since the velocity of each carrier perpetually changes due to scattering,  $I$  fluctuates within the distribution and forms a time series  $I(t)$ , the thermal noise. Since a linear combination of quantities obeying the Gaussian dis-

tribution also obeys the Gaussian distribution,  $I(t)$  at each  $t$  is distributed according to the Gaussian distribution.

$I(t)$  brings about the potential fluctuation  $v(t)$  at node  $N$ . The magnitude of  $v(t)$  is evaluated by use of the equivalent circuit in Fig. 3.<sup>11</sup> Assuming a sufficiently long observation time  $T$ , we expand the time series  $I(t)$  in a Fourier series,

$$I(t) = \sum_{n=1}^{\infty} [I_c(\omega_n)\cos \omega_n t + I_s(\omega_n)\sin \omega_n t], \quad 0 \leq t < T, \quad (2)$$

where

$$I_c(\omega_n) = \frac{2}{T} \int_0^T I(t)\cos \omega_n t dt, \quad (3)$$

$$I_s(\omega_n) = \frac{2}{T} \int_0^T I(t)\sin \omega_n t dt, \quad (4)$$

and  $\omega_n = 2\pi n/T$ . The direct current component with  $\omega_n = 0$  vanishes. Then, we obtain the expression of  $v(t)$  after a straightforward manipulation,

$$v(t) = \sum_{n=1}^{\infty} \frac{R_i\{[I_c(\omega_n) - R_i C_i \omega_n I_s(\omega_n)]\cos \omega_n t + [I_s(\omega_n) + R_i C_i \omega_n I_c(\omega_n)]\sin \omega_n t\}}{1 + R_i^2 C_i^2 \omega_n^2}, \quad 0 \leq t < T. \quad (5)$$

The Fourier components of  $v(t)$  within  $0 \leq \omega_n \leq 2\pi/R_o C_o$  constitute the potential fluctuation  $\tilde{v}(t)$  that lasts longer than  $R_o C_o$ , and hence, may cause a circuit error. Those with  $\omega_n > 2\pi/R_o C_o$  are innocent and can be ignored. More precisely, a circuit error results when

$$\tilde{v}(t) = \sum_{0 < \omega_n \leq 2\pi/R_o C_o} \frac{R_i\{[I_c(\omega_n) - R_i C_i \omega_n I_s(\omega_n)]\cos \omega_n t + [I_s(\omega_n) + R_i C_i \omega_n I_c(\omega_n)]\sin \omega_n t\}}{1 + R_i^2 C_i^2 \omega_n^2}, \quad 0 \leq t < T, \quad (6)$$

exceeds  $V/2$  in a node clamped to GND, or  $-\tilde{v}(t)$  exceeds  $V/2$  in a node clamped to  $V$ . These two cases are equivalent due to the symmetrical distribution of  $\tilde{v}(t)$ , and the circuit error probability is given by the probability that  $\tilde{v}(t)$  exceeds  $V/2$ . Since  $I(t)$  obeys the Gaussian distribution, we can conclude that  $I_c(\omega_n)$  and  $I_s(\omega_n)$ , and hence,  $v(t)$  and  $\tilde{v}(t)$  obey the Gaussian distribution. The mean value  $\langle \tilde{v}(t) \rangle$  vanishes, and we need to evaluate the variance  $\langle \tilde{v}(t)^2 \rangle$ . If we designate the power spectrum of  $I(t)$  by

$$S_I(f) = \lim_{T \rightarrow \infty} \frac{T}{2} \langle |I_c(\omega_n)|^2 + |I_s(\omega_n)|^2 \rangle, \quad (7)$$

where  $f$  is the frequency, then the power spectrum of  $v(t)$  denoted by  $S_v(f)$  is given by

$$S_v(f) = \frac{R_i^2}{1 + (2\pi f R_i C_i)^2} S_I(f), \quad (8)$$

as is suggested by Eq. (5). Hence, that of  $\tilde{v}(t)$ , denoted by  $S_{\tilde{v}}(f)$ , is expressed by

$$S_{\tilde{v}}(f) = S_v(f), \quad 0 \leq f \leq \frac{1}{R_o C_o}, \\ = 0, \quad \frac{1}{R_o C_o} < f. \quad (9)$$

According to the general theory of thermal noise, we have<sup>12</sup>

$$S_I(f) = \frac{4}{R_i} hf \left( \frac{1}{2} + \frac{1}{\exp(hf/kT) - 1} \right), \quad (10)$$

where  $h$  is the Planck constant. Even if the circuit operation frequency is as high as  $f \approx 10^{12}$  Hz, Eq. (10) can be reduced

to the Nyquist's formula  $S_I(f) = 4kT/R_i$  with sufficient accuracy. The variance of  $\tilde{v}(t)$ , denoted by  $\sigma^2$ , is evaluated by the Wiener-Khinchin's theorem as

$$\sigma^2 = \langle \tilde{v}(t)^2 \rangle = \int_0^{\infty} S_{\tilde{v}}(f) df = \frac{2kT}{\pi C_i} \tan^{-1} \left( 2\pi \frac{R_i C_i}{R_o C_o} \right). \quad (11)$$

Let us introduce a normal distribution of  $\tilde{v}(t)$  with the mean value 0 and variance  $\sigma^2$ ,  $f(\tilde{v}, \sigma)$ . Then, the circuit error probability  $p$  at node  $N$  is given by

$$p = \int_{\frac{1}{2}V}^{\infty} f(\tilde{v}, \sigma) d\tilde{v} \\ = \frac{1}{2} \operatorname{erfc} \left( \sqrt{\frac{\pi C_i V^2}{16kT} \left[ \tan^{-1} \left( 2\pi \frac{R_i C_i}{R_o C_o} \right) \right]^{-1}} \right), \quad (12)$$

where  $\operatorname{erfc}(x)$  is the complementary error function. The quantity  $C_i V^2$  is equal to  $\oint R_i i^2 dt$ , the energy consumed at node  $N$  during a cycle of charging and discharging, and will be referred to as the node energy hereafter. The directly flowing current from the supply source to GND is neglected considering the ideal transfer characteristics in Fig. 2. Notice that this quantity is nothing but the dissipation per logical switching, and is equal to the well-known figure of merit called the power-delay product. We can assume  $R_i C_i = R_o C_o$  considering the optimization of the circuit delay without loss of generality, and the argument of the complementary error function is reduced to  $0.373\sqrt{C_i V^2/(kT)}$ . Equation (12) allows us to derive the relation between the node energy  $C_i V^2$  and the circuit error probability  $p$  as is plotted in Fig. 4. The parameter is the temperature. We need  $C_i V^2$  to be around 10

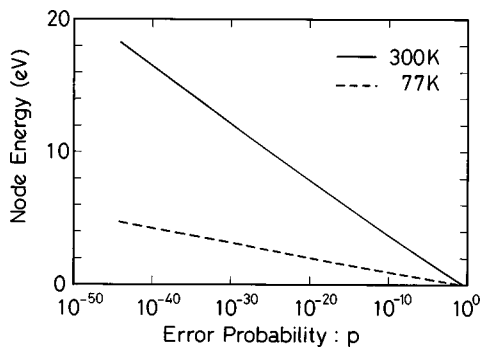


FIG. 4. The minimal node energy (the energy dissipated per logic switching at a node) necessary to suppress the error probability to less than  $p$ . The parameter is temperature.

eV for a realistic small value of  $p$ . Since  $\operatorname{erfc}(x) \approx \sqrt{\pi}^{-1} [\exp(-x^2)/x]$  for  $x \gg 1$ , the  $C_i V^2 - p$  relation is approximated by an asymptotic expression

$$\begin{aligned} C_i V^2 &= 7.2kT [\operatorname{erfc}^{-1}(2p)]^2 \\ &\approx 7.2kT \ln\left(\frac{1}{p}\right) - 3.6kT \ln\left[\ln\left(\frac{1}{p}\right)\right] \\ &\quad - 3.6kT \ln(4\pi), \end{aligned} \quad (13)$$

for a sufficiently small value of  $p$ . In the three terms on the right-hand side, the first one is dominant.

Note that the quantity  $p$  mentioned above is the probability of a circuit error occurrence at the node during the time constant, and the influence of the pileup of past errors is not included. The past error will be cleared in succeeding periods after the noise pulse is removed.

### C. System reliability and device scaling limit

Generally, requests to the error probability of each circuit node come from the specification about the reliability of the total system. We assume that an information processing system for field use is required to operate without an error for more than ten years ( $= 87\,600$  hours), i.e., less than  $10^4$  FIT level. Let us take up a large system that includes about  $10^{10}$  gates and is operated with a high clock frequency of 10 GHz. These values are practical as a future system in view of the proposed semiconductor development trend.<sup>13</sup> The error occurrence at each gate has to be less than 1 per  $3 \times 10^{28}$  clock periods. By recalling that each gate is equivalent to the inverter circuit in the previous subsection, and regarding the clock period as the time constant of the node, the requirement to the error probability of a circuit node amounts to  $p < 3 \times 10^{-29}$ . It is unnecessary to discuss the working ratio because all of the circuit node, both data-retaining and the actively operating, need to be considered. According to Fig. 4, this demands the node energy to be larger than 11.6 eV (3.0 eV) in the room-temperature (77 K) operation of the system. Or in other words, these values are 1.86 aJ (0.48 aJ) for room-temperature (77 K) operation in the usual expression of the power-delay product. Notice that a far larger value is required compared with the bare  $kT$ . This is because the small  $p$  means to endure an accidental voltage fluctuation

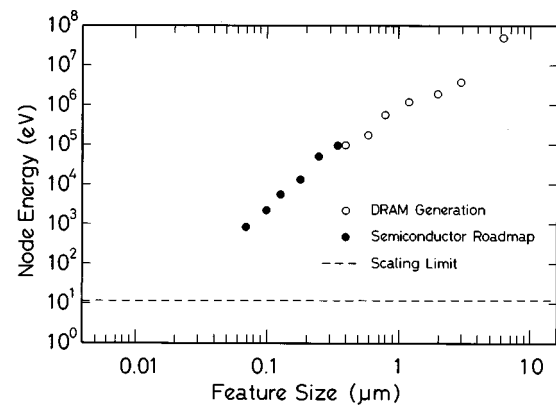


FIG. 5. The node energy reduction as a function of device feature size. The open circle is the past DRAM development trend from 16 kbit to 64 Mbit size. The filled circle is the device size trend from 1995 to 2010 proposed by the U.S. Semiconductor Roadmap (Ref. 13). The dashed line is the device scaling limit proposed by the present theory.

as large as  $11\sigma$ , and the node energy is proportional to the square of the supply voltage that overcomes the fluctuation. One may doubt that such an extremely rare case of  $11\sigma$  has a practical physical meaning that has to be considered. A fluctuation as large as  $11\sigma$  is visualized in the following example. Suppose that a small effective resistor includes 500 electrons. On the average, the number of electrons running in the right direction and that in the left direction are the same, and 250 each. The  $11\sigma$  deviation means that three-quarter of 500 electrons are running in one direction and one-quarter in the opposite direction. Such a situation is extremely rare, but by no means impossible.

These values set a lower bound to the power consumption of the total system. The power consumption of the above example is inevitably larger than 1.9 W (0.5 W) in room temperature (at 77 K), with a working ratio of 1%. The required value of  $p$  strongly depends on the size of the system. In addition, the use of redundant circuits for error correction greatly relaxes the requirement. In actuality, a majority of circuit nodes in a system may inevitably have far larger node energies due to undesired parasitic capacitance, and only a fraction has a critical node energy effectively reducing the system size to be considered. However, Fig. 4 shows that a node energy of several to 10 eV per node is still necessary even if  $p$  is relaxed by many orders of magnitude. Rather, we had better say that some 10 eV is required irrespective of the system size.

### III. DISCUSSION

The lower limit of the node energy necessarily influences the downsizing of devices. Let us compare the above result with the trend of device size scaling. Figure 5 shows the node energy reduction as a function of the device feature size. The node energy is estimated as the fan-out times  $C_{tr} V^2$ , where  $C_{tr}$  is the capacitance of the minimum width transistor of the feature size and a mean fan-out of 3 is assumed here. The open circles show dynamic random access memory (DRAM) generations from 16 kbit to 64 Mbit, and the filled circles show the United States Semiconductor Roadmap.<sup>13</sup> The dashed line shows the minimal node energy

derived above. Figure 5 suggests that the scaling limit of the MOS LSI device size lies around 10–20 nm.

In contrast to the static circuitry where the node level is clamped to external potentials, such a node level fluctuation is absent in dynamic circuitry where an internal circuit node is charged on switching and then isolated. A circuit error due to the thermal noise fluctuation may creep in only during the logical switching, and all other nodes passively retaining information are protected from error occurrence. It is possible to reduce the number of circuit nodes statically retaining information with the use of a dynamic circuitry technique, although some other troublesome aspects like the refresh operation have to be taken up. If the whole circuitry could be replaced by the dynamic one, the effective system size exposed to the thermal noise would be reduced by a factor of the working ratio, and hence, the required magnitude of  $p$  would be multiplied by the inverse of the factor. However, the working ratio of 1% assumed in the above example reduces the required node energy by only 0.9 eV or less. Notice that the relaxation of the scaling limit by the use of dynamic circuitry is far from remarkable.

The maximal value of  $p$  in a binary digital circuit is  $1/2$ , and the corresponding value of  $C_i V^2$  is estimated to be 0 in Eq. (12). Notice that the entropy increase  $k \ln 2$  due to logical switching is not included in the present theory.

The supply voltage will be gradually lowered from the present 5 V to around 0.1 V in the far future. The relation  $CV^2 = QV$  where  $Q$  is the necessary charge for switching a node, shows that minimal  $Q$  requires a few electrons for the present to around 100 electrons for the future low supply voltage. We can expect that single electronics will be improbable and multielectronics will be inevitable in the future, as long as we use the circuitry employed in the present integrated electronics.

It is interesting to compare our result with the fundamental theories introduced in Sec. I. Our result claims that a larger dissipation is required as the reliability of the total system is improved, and the dissipation–reliability relation is approximated by Eq. (13), where  $1/p$  is the reliability. These points have a similarity to the Brillouin–Neyman theory.<sup>8,9</sup> However, the Brillouin theory expects that the minimal dissipation for the case  $p = 1/2$  is  $kT \ln 2$ , whereas our result is 0. Brillouin assumes that the distribution of the noise fluctuation is the Boltzmann distribution on non-negative energy levels of a harmonic oscillator, whereas we assume it is the Gaussian distribution of the voltage with a mean value 0. The numerical factor 7.2 in Eq. (13), which stands for  $(16/\pi) \tan^{-1}(2\pi R_i C_i / R_o C_o)$ , is also different. The factor controls the coupling between the unit circuits, and properly influences the error probability of the operation. There is an essential difference in spite of the apparent similarity. Our result expects far larger dissipation compared with fundamental theories. The electronic circuits used in the present-day integrated systems perform the irreversible information processing with considerable high speed. It is not surprising that the maximal performance anticipated by ideal fundamental theories cannot be attained due to lots of practical restrictions.

## IV. CONCLUSION

Circuit nodes that are either clamped to a potential level or under logical switching operation suffer from level fluctuations due to thermal noise. The probability distribution of the fluctuation amplitude obeys the Gaussian distribution. If a sufficiently large fluctuation lasts for a longer period than the circuit time constant, the accidental level change due to fluctuation is transferred to the next node as a spurious signal and causes a circuit error at the output. The probability of such an error occurrence strongly depends on the value of the node energy  $CV^2$ , the energy dissipated per switching. This value is equal to the well-known figure of merit of the circuit, the power-delay product. In order to reduce the probability of error occurrence at a circuit node less than a certain level, the level fluctuation should be suppressed accordingly and the node energy must be larger than a certain value. A request on reliability of the total system naturally introduces the maximal error probability allowed in internal circuit nodes, and hence, leads to the lower bound of the node energy  $CV^2$ . In ultralarge systems in the future, this lower bound will amount to around 12 eV (or 2 aJ in terms of the power-delay product), which is far larger than bare  $kT$ . The lower bound of the node energy at the same time gives the lower bound of the power consumption of the total system. The downsizing of devices in LSI brings about a rapid decrease of the node energy of the circuit. The request on reliability of the total logic system will establish the lower limit to the downsizing of devices. The above result, as well as the MOS device scaling trend, indicate that the lower limit of device size in MOS LSI lies around 10–20 nm. Note that the scaling limit of a device is not a constant but depends on the required reliability of the total system. The lower bound of the node energy also implies that single electronics will be improbable and multielectronics will be inevitable in future ultrahigh integrated systems as long as the present style digital circuitry is employed. Single-electron transistor circuits based on the Coulomb blockade should be discussed separately.

## ACKNOWLEDGMENTS

One of the authors wishes to thank Professor A. Natori of the University of Electro-Communications for valuable discussions. This work was supported by the Ministry of Education, Science, Sports, and Culture under a Grant-in-Aid for Scientific Research on Priority Areas, “Ultimate Integration of Intelligence on Silicon Electronic Systems.”

<sup>1</sup>R. W. Keyes, in *Proceedings of the International Symposium on Foundations of Quantum Mechanics in the Light of New Technology* (ISQM-Tokyo'83), Tokyo, 1983 (The Physical Society of Japan, Tokyo, 1984), pp. 111–126.

<sup>2</sup>K. Natori, *J. Appl. Phys.* **76**, 4879 (1994).

<sup>3</sup>For a review, see R. Landauer, in *Der Informationsbegriff in Technik und Wissenschaft*, edited by O. G. Forberth and C. Hackl (Oldenbourg, München, 1986), pp. 139–158, and references therein.

<sup>4</sup>C. H. Bennett, *IBM J. Res. Dev.* **17**, 525 (1973).

<sup>5</sup>E. Fredkin and T. Toffoli, *Int. J. Theor. Phys.* **21**, 219 (1982).

<sup>6</sup>R. Landauer, *IBM J. Res. Dev.* **5**, 183 (1961).

- <sup>7</sup>R. W. Keyes and R. Landauer, *IBM J. Res. Dev.* **14**, 152 (1970).
- <sup>8</sup>L. Brillouin, in *Science and Information Theory*, 2nd ed. (Academic, New York, 1962), pp. 195.
- <sup>9</sup>M. S. Neyman, *Telecommun. Radio Eng. (USSR)*, Part 2 **21** (11), 68 (1966).
- <sup>10</sup>R. Landauer and J. W. F. Woo, *J. Appl. Phys.* **42**, 2301 (1971).
- <sup>11</sup>A. van der Ziel, *Noise* (Prentice-Hall, Englewood Cliffs, NJ, 1954).
- <sup>12</sup>H. Ekstein and N. Rostoker, *Phys. Rev.* **100**, 1023 (1955).
- <sup>13</sup>*The National Technology Roadmap for Semiconductors, 1994 Edition*, by Semiconductor Industry Association, U.S.A., December 1994 (SEMATECH, Inc., Austin, 1996).