Cambridge University Press
978-0-521-85727-7 - Ultra Low Power Bioelectronics: Fundamentals, Biomedical Applications, and Bio-inspired Systems
Rahul Sarpeshkar
Excerpt
More information

Section I

Foundations

Cambridge University Press
978-0-521-85727-7 - Ultra Low Power Bioelectronics: Fundamentals, Biomedical Applications, and Bio-inspired Systems
Rahul Sarpeshkar
Excerpt
More information

1 The big picture

It is the harmony of the diverse parts, their symmetry, their happy balance; in a word it is all that introduces order, all that gives unity, that permits us to see clearly and to comprehend at once both the ensemble and the details.

It is through science that we prove, but through intuition that we discover.

Henri Poincaré

This book, *Ultra Low Power Bioelectronics*, is about ultra-low-power electronics, bioelectronics, and the synergy between these two fields. On the one hand it discusses how to architect robust ultra-low-power electronics with applications in implantable, noninvasive, wireless, sensing, and stimulating biomedical systems. On the other hand, it discusses how bio-inspired architectures from neurobiology and cell biology can revolutionize low-power, mixed-signal, and radio-frequency (RF) electronics design. The first ten chapters span feedback systems, transistor device physics, noise, and circuit-analysis techniques to provide a foundation upon which the book builds. Chapters that describe ultra-low-power building-block circuits that are useful in biomedical electronics expand on this foundational material, followed by chapters that describe the utilization of these circuits in implantable (invasive) and noninvasive medical systems. Some of these systems include cochlear implants for the deaf, brain implants for the blind and paralyzed, cardiac devices for non-invasive medical monitoring, and biomolecular sensing systems. Chapters that discuss fundamental principles for ultra-low-power digital, analog, and mixed-signal design unify and integrate common themes woven throughout the book. These principles for ultra-low-power design naturally progress to a discussion of systems that exemplify these principles most strongly, namely biological systems. Biological architectures contain many noisy, imprecise, and unreliable analog devices that collectively interact through analog and digital signals to solve complex tasks in real time, with precision, and with astoundingly low power. We provide examples of how bio-inspired systems, which mimic architectures in neurobiology and cell biology, lead to novel systems that operate at high speed and with low power. Finally, chapters on batteries, energy harvesting, and the future of energy discuss tradeoffs between energy density and power density, which are essential in architecting an overall low-power system, both at small scales and at large scales.

The book can serve as a text for senior or graduate students or as a reference for practicing engineers in the fields of

- Ultra-low-power Electronics: Chapters 1 through 22, 25, and 26.
- Biomedical Electronics: Chapters 1 through 22, 25, and 26.
- Bio-inspired Electronics: Chapters 1 through 18, 21 through 26.
- Analog and Mixed-Signal Electronics: Chapters 1 through 24.

In this busy day and age, many people with an interest in these fields may not have the time to read a whole book, especially one of this size. Therefore, the book has been written so that a reader interested in only a chapter or two can read the chapter and delve deeper if he/she would like. There is a slight amount of redundancy in each chapter to enable such sampling, with interconnections among the various chapters outlined throughout every chapter. The index should also be useful in this regard. Every reader should read Chapter 1 (this chapter). Chapter 2 on the fundamentals of feedback is also essential for a deeper understanding of many chapters. Chapters 1 through 10 provide a firm foundation, necessary for a deep understanding of the whole book.

Throughout this book, intuitive, geometric, and physical thinking are emphasized over formal, algebraic, and symbolic thinking. Physical intuition is extremely important in getting systems to work in the world since they do not always behave like they do in simulations or as the mathematical idealizations suggest they do. When the mathematics becomes intractable, usually the case in all but the simplest linear and idealized systems, intuitive and physical thinking can still yield powerful insights about a problem, insights that allow one to build high-performance circuits. Practice in physical thinking can lead to a lightning-fast understanding of a new circuit that lots of tedious algebra simply can never provide. Nevertheless, one must attempt to be as quantitative as possible for a deep understanding of any system and for theory and experiment to agree well. Thus, the book does not aim to substitute qualitative understanding for quantitative understanding; rather it attempts to maximize insight and minimize algebraic manipulations. We will always aim to look at problems in a physically insightful and original way such that the answer is intuitive and can be obtained exactly and quickly because the picture in our heads is clear.

Feedback is so fundamental to a deep understanding of how circuits work and how biology works that we shall begin this book with a review of feedback systems in Chapter 2. We shall see in this chapter that feedback is ubiquitous in physical, chemical, biological, and engineering systems even though the importance of feedback has been largely unappreciated. Throughout the book, we shall draw on our knowledge of feedback systems to derive or interpret results in a simple way that would not be possible without the use of this knowledge. For example, our discussion of physics in an MOS transistor will often use feedback analogies to understand the physics of their operation intuitively in Chapters 3 and 4. The equations of electron velocity saturation in an MOS transistor will be represented as a feedback loop in Chapter 6. We shall often avoid tedious Kirchoff's current law algebraic equations by simply drawing a feedback loop to provide all the answers for any transfer function, noise or offset analysis, robustness analysis, or dynamic analysis that we may need. We shall use feedback interpretations to understand how the noise in a transistor is affected by internal feedback within it. A deep understanding of feedback and circuits can enable a unified understanding of several systems in the world.

In both biomedical and bio-inspired electronics, it is important to deeply understand the biology. To understand and mimic biological systems in this book, we shall use circuits as a primary language rather than mathematics. Several nonlinear partial differential equations and structures in biology then translate into simple intuitive, lumped or distributed circuits. For example, we use such circuits to mimic the inner ear or cochlea, to understand the retina in the eye, to understand and mimic the heart, to mimic the vocal tract, to mimic spiking (pulsatile) neurons in the brain, and to understand and mimic biochemical gene–protein and protein–protein molecular networks within cells. Such circuits can help make engineers and physicists more comfortable with biology because it is described in a familiar language. Distributed circuits will help us understand Maxwell's equations and antennas intuitively. Circuits will help us quickly understand chemical reactions. Circuits will even help us understand the energy efficiency of cars.

In the rest of this chapter, we shall summarize some themes, ideas, principles, and biomedical and bio-inspired system examples that are discussed in depth elsewhere in the book. In this introductory chapter, the aim is to provide an intuitive 'big picture' without getting caught up in details, citations, proofs, mathematical equations and definitions, subtleties, and exceptions, which are addressed in the remaining chapters of the book. We shall start by discussing the importance of ultra-low-power electronics. We shall describe a power-efficient regime of transistor operation known as the subthreshold regime, which is enabling in low-power design. We shall then discuss important connections between information, energy, and power. We shall highlight some key themes for designing ultra-low-power mixed-signal systems that have analog and digital parts. We shall discuss examples of biomedical application contexts for low-power design, and fundamental principles of low-power design that are applicable to all systems, analog or digital, electronic or biological. After providing some numbers for the amazing energy efficiency of biological systems, we shall briefly discuss examples of systems inspired by neurobiology and by cell biology. Then, we provide a discussion of batteries and other energy sources, highly important components of low-power systems at small scales and at large scales. Finally, we shall conclude with a summary of the book's sections and some notes on conventions followed in the book.

1.1 Importance of ultra-low-power electronics

Ultra-low-power electronics in this book usually refers to systems that operate anywhere from a pico- to a milliwatt. However, the principles of ultra-low-power design are useful in all kinds of systems, even in low-power microprocessors, that dissipate 1 W, say, rather than 30 W, without compromising performance. In general, ultra-low-power electronic design is important in five different kinds of systems:

1. Systems that need to be portable or mobile and therefore operate with a battery or other energy source of reasonable size with a long lifetime or long time between recharges. The more miniature the system, the smaller the energy source, and the more stringent is the power constraint.
2. Systems that function by harvesting relatively small amounts of energy from their environment.
3. Systems that need to minimize heat dissipation.
4. Complex systems with many devices whose complexity simply will not scale unless the power and consequently heat dissipation of each of their components is kept within check.
5. Systems where the overall cost is a strong function of the size of the system or the cost of the battery in the system.

Biomedical systems are examples of systems where ultra-low-power electronics is paramount for multiple reasons. For example, biomedical systems that are implanted within the body need to be small and lightweight with minimal heat dissipation in the tissue that surrounds them. In some systems like cardiac pacemakers, the implanted units are often powered by a non-rechargeable battery. In others, like cochlear implants, the implants are traditionally constantly powered by rectified wireless energy provided by a unit outside the body. In either case, power dissipation dictates the size of the needed receiving coil, antenna, or battery, and therefore sets a minimal size constraint on the system. Size is important to ensure that there is space within the body for the implant and that surgical procedures are viable. Implant-grade batteries need conservative short-circuit-protection mechanisms to mitigate concerns about battery shorting and resultant tissue heating within the body. They are relatively expensive and the cost of implanted systems is strongly impacted by the costs of the battery. The costs of hermetic sealing and bio-compatibility of electronic implants are size dependent as well. A fully implanted system with a battery that has a limited number of wireless recharges must operate under stringent low-power constraints such that constant surgery is not needed to change the battery in a patient. The system must ideally function for 10 to 30 years without the need for a battery replacement. Implanted systems with ultra capacitors are capable of more recharges than batteries, but they have low energy densities such that more frequent recharging is necessary. Thus, ultra-low-power operation will always be paramount in implantable biomedical systems.

Noninvasive biomedical systems like cardiac medical tags are attached to the skin or to clothing for patient monitoring. They are powered by received RF energy or by a battery and also need to operate in an ultra-low-power fashion. In certain lab-on-a-chip or biomolecular-sensing instrumentation, ultra-low-noise and precision electronics is often more important than ultra-low-power operation. Fortunately, a good ultra-low-power designer is also a good ultra-low-noise designer, since both kinds of designers need to be skilled in the management of circuit noise. Therefore, many of the techniques for ultra-low-power design that we discuss in this book will also enable the reader to become a skilled ultra-low-noise designer. For example, we will discuss how to architect electronics for a micro-electro-mechanical-system (MEMS) vibration sensor with 0.125 parts-per-million sensitivity (23-bit precision), and for ultra-low-noise and micropower neural and cardiac amplifiers in the book.

This book will, in large part, use biomedical systems as examples. However, the principles, techniques, and circuits that we describe in this book are general and are applicable in several other portable systems such as in cell phones, next-generation radios, sensor networks, space, and military hardware. In fact, many of the principles of ultra-low-power design that we shall explicitly and implicitly discuss, e.g., energy recycling, apply to non-electrical systems and at much larger power scales as well. They can and already are being exploited in the design of low-power mechanical systems such as next-generation cars. Cars, which operate on average at nearly 42 kW when going at 30 mph today, cannot afford to operate with such power consumption in our planet's future.

1.2 The power-efficient subthreshold regime of transistor operation

Since power is the product of voltage and current, low-power systems must necessarily operate at low voltages and/or low currents[1]. Electronic systems that run on 0.5 V power-supply voltages already exist. It is hard to scale the power-supply voltage of electronic circuits below 0.25 V and preserve reliable performance in digital circuits, as we discuss in Chapter 21. The performance of analog circuits significantly degrades at such low power-supply voltages. Thus, power savings via voltage reduction is inherently limited in both the digital and the analog domains. For ultra-low-power systems, it is more promising to focus on currents, which can be scaled from pA to mA levels in modern-day MOS transistors with appropriate use of transistor geometries and bias voltages in the subthreshold regime of transistor operation.

The subthreshold region of operation is present in a transistor when it is operated below its *threshold voltage*, a region where digital circuits are sometimes approximated as being 'turned off'. In reality, the threshold voltage is merely a useful approximation to describe transistor operation. Just as the current in a diode decreases exponentially as the voltage across it is decreased, the current in a subthreshold transistor decreases exponentially as the magnitude of the transistor's gate-to-source voltage is decreased. Nevertheless, just as it is sometimes useful to

[1] In non-dissipative devices like ideal inductors or capacitors, current and voltage variables are always orthogonal to each other, such that power is never dissipated even at high voltages and/or high currents. However, even in such devices, in practice, finite dissipative losses are minimized with low voltages and/or low currents.

view a diode as having a 'turn-on threshold' of 0.6 V due to the steep nature of its exponential current–voltage characteristics, it is sometimes useful to view a transistor as having an abrupt threshold voltage at which it turns on.

The 'leakage current' in a transistor that is turned off in a digital system is dominated by the transistor's subthreshold current. Such leakage current can be considerable in large digital systems. For example, 100 million transistors × 1 nA of leakage current per transistor yields 100 mA of standby leakage current. Due to the lowering of the threshold voltage of the transistor in advanced transistor processes, the absolute value of the subthreshold leakage current increases as MOS technologies progress towards ever-smaller dimensions. Subthreshold operation also occupies an increasingly larger fraction of the range of power-supply operation as transistor sizes get progressively smaller. For all of these reasons, there has been a great renewal of interest in the subthreshold regime of operation. Subthreshold operation in both analog and digital circuits has almost become a necessity.

The maximal frequency of operation possible in diffusion-current-determined subthreshold operation scales inversely with the square of the transistor's channel length. In contrast, the maximal speed of velocity-saturated above-threshold operation only scales inversely with the channel length. Thus, subthreshold operation is rapidly allowing faster speeds of operation and may no longer be viewed as a 'slow regime' of operation of the transistor. For example, 1 GHz analog preamplifiers with all-subthreshold operation can now be built in a 0.18 μm process and digital circuits can be made to operate at such speeds as well.

The subthreshold region of operation is a region where the bandwidth available per ampere of current consumed in the transistor is maximal. The power-supply voltage needed for subthreshold operation is also minimal since the saturation voltage of a transistor in subthreshold is only 0.1 V. Due to the high bandwidth-per-ampere ratio and the ability to use small power-supply voltages, the bandwidth per watt of power consumed in the transistor is maximized in its subthreshold regime. Consequently, subthreshold operation is the most power-efficient regime of operation in a transistor.

For all these reasons, this book focuses heavily on the use of subthreshold circuits for ultra-low-power electronic design in the analog and the digital domains. In the subthreshold region of operation, often also referred to as the weak-inversion region of operation, it is important to ensure that systems are robust to transistor mismatch, power-supply-voltage noise, and temperature variations. Hence, circuit biasing and feedback techniques for ensuring robustness are important. We shall discuss them in various contexts in the book, but particularly in Chapter 19, where we discuss the design of ultra-low-power biomedical system chips for implantable applications, and in Chapter 22 on ultra-low-power digital design. Furthermore, the subthreshold regime is characterized by relatively high levels of noise, since there are few electrons per unit time to average over. Thus, throughout the book, we shall discuss device noise, how to mitigate it, how to analyze it, how to design around it, and, in some cases, how to even exploit it.

Since ultra-energy-efficient biological systems also operate with Boltzmann exponential devices, subthreshold operation is highly useful in mimicking their operation. Thus, subthreshold operation is enabling in bio-inspired systems as well.

1.3 Information, energy, and power

Information is always represented by the states of variables in a physical system, whether that system is a sensing, actuating, communicating, controlling, or computing system or a combination of all types. It costs energy to change or to maintain the states of physical variables. These states can be in the voltage of a piezoelectric sensor, in the mechanical displacement of a robot arm, in the current of an antenna, in the chemical concentration of a regulating enzyme in a cell, or in the voltage on a capacitor in a digital processor. Hence, it costs energy to process information, whether that energy is used by enzymes in biology to copy a strand of DNA or in electronics to filter an input.[2] To save energy, one must then reduce the amount of information that one wants to process. The higher the output precision and the higher the temporal bandwidth or speed at which the information needs to be processed, the higher is the rate of energy consumption, i.e., power. To save power, one must then reduce the rate of information processing. The information may be represented by analog state variables, digital state variables, or by both. The information processing can use analog processing, digital processing, or both.

The art of low-power design consists of decomposing the task to be solved in an intelligent fashion such that the rate of information processing is reduced as far as is possible without compromising the performance of the system. Intelligent decomposition of the task involves good architectural system decomposition, a good choice of topological circuits needed to implement various functions in the architecture, and a good choice of technological devices for implementing the circuits. Thus, low-power design requires a deep knowledge of devices, circuits, and systems. This book shall discuss principles and examples of low-power design at all of these levels. Figure 1.1 shows the "low-power hand". The low-power hand reminds us that the power consumption of a system is always defined by five considerations, which are represented by the five fingers of the hand: 1) the task that it performs; 2) the technology (or technologies) that it is implemented in; 3) the topology or architecture used to solve the task; 4) the speed or temporal bandwidth of the task; and, 5) the output precision of the task. As the complexity, speed, and output precision of a task increase, the rate of information processing is increased, and the power consumption of the devices implementing that task increases.

[2] In Chapter 22, we shall see that, technically, if one operates infinitely slowly and in a manner that allows the states of physical variables to be recovered even after they have been transformed, energy need not be dissipated. In practice, in both natural and artificial systems, which cannot compute infinitely slowly, and which always have finite losses, there is always an energy cost to changing or maintaining the states of physical variables.

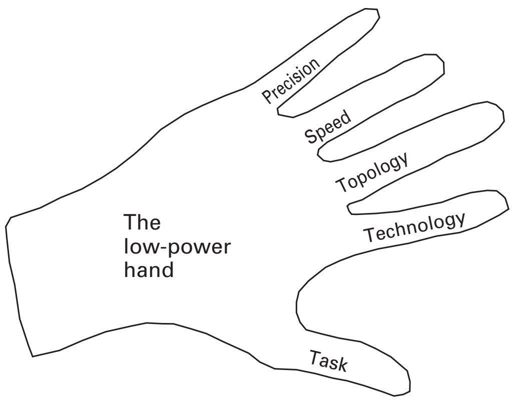

Figure 1.1. The low-power hand.

1.4 The optimum point for digitization in a mixed-signal system

The problem of low-power design may be formulated as follows: Suppose we are given an input **X**, an output function $\mathbf{Y}(t) = f(\mathbf{X}, t)$, basis functions $\{i_{out1} = f_1(\mathbf{v}_{in}), i_{out2} = f_2(d\mathbf{v}_{in}/dt), i_{out3} = f_3(\int \mathbf{v}_{in}), ..\}$ formed by the current-voltage curves of a set of technological devices, and noise-resource equations for devices in a technology that describe how their noise or error is reduced by an increase in their power dissipation for a given bandwidth; such noise-resource equations are described in Equation (22.4) in Chapter 22. Then, find a topological implementation of the desired function in terms of these devices that maximizes the mutual information between the actual output $\mathbf{Y}(t)$ and the desired output $f(\mathbf{X},t)$ for a fixed power-consumption constraint or per unit system power consumption. Area may or may not be a simultaneous constraint in this optimization problem. A high value of mutual information, measured in units of bits per second, implies that the output encodes a significant amount of desired information about the input, with higher mutual information values typically requiring higher amounts of power consumption.

Hence, low-power design is in essence an information-encoding problem. How do you encode the function you want to compute, whether it is just a simple linear amplification of a sensed signal or a complex function of its input, into transistors and other devices that have particular basis functions given by their current-voltage curves? Note that this formulation is also true if one is trying the minimize the power of an actuator or sensor, since information is represented by physical state variables in both, and we would like to sense or transform these state variables in a fashion that extracts or conveys information at a given speed and precision. In non-electrical systems, through (current) and across (voltage) variables play the roles of current and voltage, respectively. For example, in a fluid-mechanical system, pressure is analogous to voltage while volume velocity of fluid flow is analogous to current.

A default encoding for many sensory computations in man-made systems is a high-speed high-precision analog-to-digital conversion followed by digital signal processing that computes *f*( ) with lots of calculations. This solution is highly flexible and robust but is rarely the most power-efficient encoding. It does not exploit the fact that the actual meaningful output information that we are after is often orders-of-magnitude less than the raw information in the numbers of the signal, and the fact that the technology's basis functions are more powerful than just switches. It may be better to preprocess the information in an analog fashion before digitization, and then digitize and sample higher-level information at significantly lower speed and/or precision.

A familiar example in the field of radio engineering illustrates why analog preprocessing before digitization is important in lowering overall system power. Suppose a radio with a 1 GHz carrier frequency is built with a 16-bit analog-to-digital converter (ADC) such that all operations are done directly in the digital domain. Such a design affords maximum flexibility and robustness and allows one to build a 'software radio' that can be programmed to work at any carrier frequency up to 1 GHz as long as a broadband antenna is available and to function over a nearly 96 dB input dynamic range. The power consumption of the ADC in the software radio alone would be at least 1 pJ/(quantization level) $\times 2^{16}$ quantization levels $\times 4 \times 10^9$ Hz $= 256$ W, an unacceptable number for the few hundreds of milliwatt power budget of a cell phone! The figure of 1 pJ/(quantization level) represents a very optimistic estimate for the energy efficiencies of ADCs working at such simultaneously high speeds and precisions, thus far never reported. This figure also assumes that the precision of the ADC is not thermal noise limited, an optimistic assumption, and that the signal is sampled at four times the Nyquist rate, typical in many applications. The digital processor has to process 16-bit numbers at 10^9 Hz, which makes its power consumption high as well. It is no wonder that most radios are built with analog preprocessing to amplify and mix the high-bandwidth signal down to baseband frequencies where it can be digitized by a much lower-speed ADC and processed by a digital signal processor (DSP) operating at a low information rate. The power of the ADC and the power of the DSP are then significantly reduced. The analog circuits are built with power-efficient passive filters, power-efficient tuned low-noise amplifiers, power-efficient oscillators, and power-efficient active mixers. The extra degrees of freedom inherent in analog processing, i.e., every signal is continuous and not just a '1' or a '0', and the fact that every device has a technological input-output curve that can be exploited in the computation and is not just a switch, saves energy. Function approximation is not merely done with quantization and Boolean logic functions but with basis functions which are fewer and more efficient for the task. In radios, the basis functions inherent in inductors and capacitors allow for very energy-efficient, relatively noise-free filtering.

One of the themes of this book is that the delayed-digitization example in the preceding paragraph generalizes to other designs where power consumption needs to be optimized. In general, Figures 1.2 (a) and 1.2 (b) illustrate how the overall