1 Introduction

This chapter explains the basic motivations for developing MOS transistor models that can be used for the design of complementary MOS (CMOS) integrated circuits. It then gives a short history of the EKV MOS transistor model starting from the early development, motivated by the design of micropower circuits for watch applications, to the most recent developments. Finally, the structure of the book is highlighted in order to help the reader organizing his reading.

1.1 THE IMPORTANCE OF DEVICE MODELING FOR IC DESIGN

Modern large-scale integrated circuits are essentially composed of MOS transistors and their interconnections. Therefore, the design of such circuits requires some kind of a model for the transistors.

For noncritical digital circuits, this model may in principle be very simple. Indeed, modeling each transistor as an on–off switch would be sufficient to design purely logic circuits. However, as soon as there are critical races among transitions, the model must be extended to describe the dynamic behavior of the device, in order to obtain the rise and fall time of these transitions. This dynamic behavior is also needed when the frequency of operation approaches its maximum limit. With the reduction of supply voltage, more details must be introduced, such as the residual current of blocked transistors, the importance of which is increased.

Analog circuits contain usually a smaller number of transistors, but they are even more dependent on the exact behavior of each transistor. The design of high-performance analog circuits therefore requires a very detailed model of the transistor. This model must include a precise description of the voltage–current relationships, including the effect of the source that is often not grounded, and of the dynamic behavior of the device. Its behavior with respect to noise and to temperature variations must also be accounted for.

A transistor model intended for circuit design should serve two essential purposes:

It should first provide a good understanding of the various properties of the device to facilitate the synthesis of optimum circuit architectures. Indeed, in order to build robust circuits, the

Charge-Based MOS Transistor Modeling: The EKV Model for Low-Power and RF IC Design C. Enz and E. Vittoz

physical properties of the transistor must be exploited in a way that is minimally dependent on temperature and process variations. For this purpose, the model should be explicit. It should "speak to the mind," using no complicated or chained equations. Clarity should supersede precision and can be enhanced by means of graphical representations. This important aspect of a model is often underestimated and overlooked. It will be emphasized in this book, in particular in Part I, which essentially describes what can be called the core of the model.

Second, the model should be adapted to numerical simulations on a computer, embedded in a circuit simulator. For this purpose, precision supersedes clarity, and second-order effects must be accounted for. This can be obtained by predistorting variables, by chaining equations and/or by providing additional layers around a core model. The model does not need to be fully explicit, but it should be compact: it should use sufficiently simple expressions with minimum need for numerical iterations, in order to limit the computation time.

A transistor model should include a minimum number of process-dependent device parameters. This is to facilitate the very heavy task of extracting and following-up the value of these parameters, with their statistical distribution and temperature dependency.

Now, the correlation between these parameters (with process and temperature variations) must be known, in order to avoid designing circuits for irrelevant worst cases. For this reason, the device parameters should be explicitly based on independent and measurable process parameters. This is essential to be able to ascertain their amount of correlations while avoiding the almost impossible task of measuring all these correlations. It also makes the model predictive, allowing to foresee the characteristics of the transistor and hence the performance of the resulting circuits even before measuring the device.

The EKV model described in this book is believed to meet all the above expectations. It serves the two main purposes in a coherent manner. Its core requires just a few parameters to describe all the basic properties of the long-channel intrinsic device in an explicit manner. Layers are added to this core to account for short-channel and secondary effects.

1.2 A SHORT HISTORY OF THE EKV MOS TRANSISTOR MODEL

The model presented in this book results from a series of direct and indirect contributions along several decades. Its origins can be traced back to the early developments of electronic watches at CEH (French acronym for Watchmakers Electronic Center) in Switzerland [2].

The total power consumption had to be extremely low, less than 1μW, to ensure a few years of life to the single button-size cell battery. After the very first versions based on bipolar transistors [3], the CMOS technology was soon identified as the best approach to implement the digital electronic circuitry needed in a watch using a crystal resonator as the time reference. Supply voltage had to be very low, compatible with the 1.3 V delivered by the cell, so the development of low-threshold CMOS was a major challenge in the late 1960's [4].

The digital circuitry was essentially an asynchronous chain of divide-by-two stages. The main design problem was to minimize the number of node transitions in order to minimize the dynamic power. Another one was the elimination of logic hazards to improve the robustness against large local variations of the small gate voltage overhead, and this led to the first single clock circuits [5–7]. For these digital circuits, MOS transistors could be considered just as switches and hence no special model was required.

The problem was very different for the few analog subcircuits. Most important was the circuitry needed to sustain the oscillation of the quartz crystal resonator (the quartz oscillator). Each transistor had to be biased at a drain current much below $1\mu A$. Early measurements carried out in 1967 showed that the transistor behaved in a very strange manner at these very low current levels. Indeed, the well-known square-law transfer characteristics were replaced by an exponential over more than 5 order of magnitude of the drain current, very similar to bipolar transistors. This is how *weak inversion* popped out to the attention of micropower circuit designers in the late 1960s.

At that time, no transistor model was available for weak inversion, but they started coming out in subsequent years, mainly to account for what appears in digital circuits as a leakage current of blocked transistors. In 1972, M. B. Barron published a model for the grounded source device showing the exponential dependencies on drain voltage and on surface potential, with a rather complex expression relating the surface potential to the gate voltage [8]. The same year, R. M. Swanson and J. D. Meindl [9] showed that this relation could be accounted for by means of an almost constant factor, which became the *slope factor n* of our model. The following year, R. R. Troutman and S. N. Chakravarti [10] treated the case of nonzero source voltage. Then T. Mashuhara *et al.* [11] showed that the current depends on a difference of exponential functions of source and drain voltages. In the mean time, micropower analog circuit blocks were developed at CEH. They were first published in 1976 [12, 13], together with a model applicable for weak inversion circuit design, which was based on the previously mentioned work. This model already included two important features of the EKV model: *reference to the (local) substrate* (and not to the source) for all voltages and full *source–drain symmetry*. The related small-signal model including noise was also presented [14].

A symmetrical model of the MOS transistor in strong inversion was first published by P. Jespers in 1977 [15, 16]. Based on an idea of O. Memelink, this graphical model uses the approximately linear relationship between the local mobile charge density and the local "non-equilibrium" voltage in the channel. This charge-based approach has been adopted and generalized to all levels of current in the EKV model.

Another ingredient of EKV is the representation of the drain current as the difference between a *forward* and a *reverse* component. This idea was first introduced in 1979 by J.-D. Châtelain [17], by similarity with the Ebers–Moll model of bipolar transistors [18]. However, his definition of these two components was different from that adopted later, and was not applicable to weak inversion.

Even in micropower analog circuits, not all transistors should be biased in weak inversion. There was therefore a need for a good continuous model from weak to strong inversion. Such a model was developed at CEH by H. Oguey and S. Cserveny, and was first published in French in 1982 [19]. The only publication in English was at a Summer Course given in 1983 [20].

This model embodied most of the basic features that were retained later. It introduced a function of the gate voltage called control voltage, later renamed *pinch-off voltage* V_P. A single function of this control voltage and of either the source voltage or the drain voltage defined two components of the drain current (which became the forward and reverse components). This function was continuous from weak to strong inversion, using a mathematical interpolation to best fit moderate inversion.

In the mid-1980s, the model of Oguey and Cserveny was simplified by the second author for his undergraduate teaching at EPFL (Swiss federal Institute of Technology, Lausanne, Switzerland), and most further developments were carried out there. They started with the Ph.D. Thesis of the first author [21], in collaboration with F. Krummenacher. The model was

formulated more explicitly. Noise and dynamic behavior were introduced by exploiting the fundamental source–drain symmetry. The status of the model was presented at various Summer Courses [22–24] and a full paper was finally published in 1995 [25]. This publication gave its name to the model, but many important extensions were added later.

Probably the most important extension was the replacement of the current and transconductance interpolation functions between weak and strong inversion presented in [25] by a more physical based one, derived from an explicit linearization of the inversion charge versus the surface potential. The incremental linear relationship between inversion charge and surface potential was first considered by M. Bagheri and C. Turchetti [26], but the linearization of the inversion charge versus surface potential was originally proposed in 1987 by M. Maher and C. Mead [27, 28]. Several years later, different groups looked at this problem. B. Iñiguez and E. G. Moreno [29, 30] derived an approximate explicit relation between inversion charges and surface potential which included a fitting parameter. While their first linearization was done at the source [29], they later obtained a substrate referenced model based on the original EKV MOSFET model approach [25], which also included some short-channel effects. A similar approach was also proposed by Cunha *et al.* [31–34] who obtained an interpolated expression of the charges versus the potentials that used the basic EKV model definitions[1] [25] and was closely inspired from our approach. We also adopted the inversion charge linearization approach, since it offers physical expressions for both the transconductance-to-current ratio and the current that are valid from weak to strong inversion [35–38]. This gave rise to the *charge-based EKV model* which is discussed in this book. The inversion charge linearization principle was rediscovered once more in 2001 by H. K. Gummel and K. Singhal [39, 40]. Finally, a formal detailed analysis of the inversion charge linearization process and a rigorous derivation of the EKV model was finally published by J.-M. Sallese *et al.* in [41].

Note that this approach actually provides voltages versus currents expressions that cannot be explicitly inverted. It can nevertheless be easily inverted by using a straightforward Newton-Raphson technique or by an appropriate approximation. Both these techniques have been used in the final model implementation.

The basic long-channel charge-based EKV model was further developed by the EKV team to include the following additional effects:

Nonuniform doping: Nonuniform doping in the vertical direction was proposed by C. Lallement *et al.* in [42, 43].

Non-quasi-static model: A small-signal charge-based non-quasi-static model was presented by J.-M. Sallese and A.-S. Porret in [38, 44].

Polysilicon depletion and quantum effects: Polysilicon depletion and quantum effects were also added [45–47].

RF modeling: The EKV model was extended by the first author to also cover high-frequency operation for the design of RF CMOS integrated circuits [48–52].

Thermal noise: An accurate thermal noise model accounting for short-channel effects was developed by A. S. Roy and C. C. Enz [53–55].

[1] Unfortunately, Cunha *et al.* did not use the same definition of the specific current we have been using. Their specific current is actually four times smaller.

Extrinsic components: An accurate model of the parasitic capacitances was developed by F. Prégaldiny *et al.* [56].

EKV compact model: A model of the MOS transistor would be almost useless if it could not be used by circuit designers with a circuit simulator. To this purpose, the model has to be carefully implemented in the simulator so that it can run efficiently avoiding any convergence problems. The early EKV model (version 2.7) was implemented by M. Bucher as a compact model in many circuit simulators [37]. All the more recent developments were implemented in the version 3.0 of the EKV MOS transistor compact model [57, 58].

Parameter extraction: A compact model cannot be used without an efficient parameter extraction methodology. The EKV model uses an original parameter extraction methodology presented in [59–62]. (Reference [61] can be found on line at the EKV Web site [63].)

More recently, the research of the EKV team is more oriented toward the modeling of multigate MOS devices and more particularly on double-gate devices [64, 65].

Further parts of the model were derived by members of the team of researchers and Ph.D. students that developed its implementation as a CAD tool at EPFL [63].

1.3 THE BOOK STRUCTURE

This book is organized in three parts, which are briefly described below:

Part I describes the basic long-channel charge-based MOS transistor model. It is the core of the model around which all the other parts are built in a hierarchic manner following the basic structure of the EKV MOS transistor model. This part is self-contained and the reader can stop after it while still having a strong background in all the fundamental aspects of the EKV MOS transistor model. It includes all the most important aspects such as basic large-signal static model, small-signal dynamic model, noise model, and a discussion of temperature effects and matching properties. The other parts complete the basic model by adding more detailed descriptions of advanced aspects.

Part II presents more advanced aspects which are of utmost importance for understanding the operation of deep-submicron devices. It starts with the modeling of several nonideal effects that already affects long-channel devices before concentrating on short-channel effects. The model is then extended to also include the extrinsic part of the device.

Part III discusses additional aspects which become important when increasing the operating frequency. It presents a complete MOS transistor model, built on top of the two first parts, which is required for designing RF CMOS integrated circuits.