Multivariate Cubic Spline: A Versatile DC Modeling Technique Suitable for Different Deep Submicron Transistors

Md Sakib Hasan, Sherif Amer, Syed K. Islam, and Garrett S. Rose

IEEE SoutheastCon 2019, Huntsville, AL, Apr 2019.

©2019 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

The online home for this paper may be found at: http://web.eecs.utk.edu/~grose4/

Citation Information (BibTex):

@INPROCEEDINGS{hasan2019evaluation,
  author="Md Sakib Hasan and Sherif Amer and Syed K. Islam and Garrett S. Rose",
  title="Multivariate Cubic Spline: A Versatile DC Modeling Technique Suitable for Different Deep Submicron Transistors",
  booktitle="IEEE SoutheastCon 2019",
  month="April",
  year="2019",
  address="Huntsville, AL, USA"}
Multivariate Cubic Spline: A Versatile DC Modeling Technique Suitable for Different Deep Submicron Transistors

Md Sakib Hasan†, Sherif Amer*, Syed K. Islam†, and Garrett S. Rose*

*Department of Electrical Engineering & Computer Science, University of Tennessee, Knoxville, TN 37996 USA
†Department of Electrical Engineering and Computer Science, University of Missouri, Columbia MO 65211 USA

Corresponding Author, mhasan4@vols.utk.edu

Abstract—This work presents multivariate cubic spline polynomial as a versatile and efficient method for DC modeling of modern transistors with very different underlying physics including MOSFET (metal-oxide-semiconductor field-effect transistor), MESFET (Metal-Semiconductor-Field-Effect-Transistor), HBT (heterojunction bipolar transistor), HEMT (High-electron-mobility transistor) and a novel silicon-on-insulator four-gate transistor (G4FET). A set of available training data from TCAD simulation, analytical expression and experimental measurements is used to determine the coefficients of the spline model and then the model is validated using another set of test data. The developed model expresses the drain current as a multivariate cubic spline and it is shown to be valid across a wide range of bias conditions provided the independent variables are inside the range of data set used for training. The formulation of the cubic spline ensures its continuity along with the continuity of its first and second order derivatives which is highly desirable for implementation in a SPICE simulator. The model shows excellent predictive capability for different kinds of devices. This can be very useful for modeling deep-submicron emerging devices for which any closed-form analytical solution is not yet available.

I. INTRODUCTION

Numerical modeling is an alternative to physics-based modeling which can be used to develop simple, fast and precise device models. The methodology consists of taking empirical data (measured/TCAD/ analytical) as input, using the data to develop a model and then reproducing the complex nonlinear behavior of the semiconductor devices with reasonable accuracy. They can be used to model different types of transistors with very different operating principles and fabricated using various process technologies. There is an abundance of emerging deep-submicron devices with extremely complex underlying mechanism. It is quite difficult and time consuming to develop closed-form analytical solutions suitable for SPICE implementation for these devices. A model completely based on semiconductor device physics necessitates the solution of coupled nonlinear differential equations. This is not feasible for modern circuits containing a very large number of transistors. Numerical modeling techniques can be useful in these cases by providing a simpler alternative.

A plethora of numerical modeling approaches has been reported in literature. The model developers have often used quadratic or higher order polynomials in their model formulation since their unique properties render them suitable for implementation in circuit simulator. This work reports a multivariate cubic spline polynomial approach to model DC I-V characteristics of several different transistors with widely different working mechanisms. The modeled devices include deep submicron MOSFETs for different channel lengths (22, 32 and 45 nm), Metal-Semiconductor-Field-Effect-Transistor (MESFET) [1], Heterojunction bipolar transistors (HBT) [2], high-electron-mobility transistors (HEMT) [3] and a novel silicon-on-insulator four-gate transistors (G4FET) [4].

First, a set of training data, obtained from TCAD Sentaurus simulation/ analytical expression/ experimental measurements, has been used to determine the parameters of the model. Then the model accuracy has been verified against a different set of test data. The model expresses the output current as a multivariate cubic spline polynomial where the independent variables can be terminal voltages, terminal currents (base current for BJT), geometric dimensions (W, L, tox) etc. The model formation ensures that it has zeroth ($C^0$), first ($C^1$) and second ($C^2$) order continuity which makes the model suitable for most circuit simulators employing Newton-Raphson method for solving circuit equations. In this work, we only show DC I-V modeling but the same approach can be used to model C-V characteristics to enable transient simulation. The excellent predictive capability of the model for different kinds of devices bears testimony to its versatility and effectiveness. To demonstrate this approach, we have implemented it without any prior assumption of knowledge about underlying physical mechanisms. This approach can aid model developers to develop simple, fast and accurate SPICE models for deep submicron emerging devices with no closed-form analytical solution yet available which, in turn, will enable circuit designers to explore innovative applications using those devices.

II. PRIOR WORKS IN NUMERICAL MODELING

Numerical methods usually employ table look-ups and polynomials to interpolate between data points. This is an empirical approach towards modeling. Even physics based models usually have some fitting parameters which are empirically determined to accurately match experimental results. A table lookup based approach was used in [5] to address the specific
requirements of analog circuit design, such as accurately reproducing small-signal parameters, large signal nonlinearities, subthreshold characteristics, short-channel effects, and voltage dependent capacitances. In order to accurately capture short-channel effects for efficient circuit simulation, a table lookup model consisting of a main table complemented with a coarse three dimensional sub-table (capturing substrate effects) and another table to interpolate among channel lengths was developed and implemented in circuit simulator to get rid of the limitations of analytical models [6]. An approach to dynamic MOSFET modeling, based on the interpolation of terminal charges and conductive currents with physically motivated functions has been reported [7]. This kind of modeling is particularly useful for simulation of low-voltage mixed signal circuits. A general n-dimensional first order continuous table model has been proposed in [8] where each table model has been shown to reproduce the exact behavior of the DC current expressions of two basic physical device models; the Ebers-Moll model for bipolar junction transistor and the GLASMOST model for MOSFETs with high accuracy and less evaluation times. Simple interpolation methods have been developed in [9] to construct any current table from a small basis set of tables for variation of width, length, and temperature. Quadratic B-splines with not-a-knot boundary conditions were used for length and temperature interpolation, whereas simple scaling along with decomposition of channel has been done for width variation to take narrow width non-idealities into consideration. A blending function combining exponential and polynomial interpolation for the accurate evaluation of the drain current in moderate inversion was developed in [10]. This model is flexible with options for several interpolation methods and the user can choose one depending on the speed, memory and accuracy requirement. The model performed well for DC, AC and transient analyses.

In [11], Bernstein approximation technique has been extended to multidimensional variation diminishing interpolation and applied to DC current and intrinsic charge modeling of the MOSFET to increase simulation efficiency. Lagrange and Bernstein interpolation polynomials have also been used to model multi-gate transistors [12]. Monotonic Piecewise Cubic Interpolation has been used in [13] to determine the MOSFET operating point using stored table value generated by a 2-D device simulator. A table-style spline formulation has been presented in [14] using quadratic splines ensuring continuity of the function and its derivative and presented a new data-compression technique to efficiently store the coefficients of spline polynomial. Multidimensional linear and cubic spline interpolation have also been used for modeling multi-gate SOI devices [15]. In another work on table method [16], a methodology of generating compact and accurate first order table model for highly nonlinear multidimensional behavior has been demonstrated. A multi-variate regression polynomial model was reported in [17] to model DC current-voltage characteristics of SOI multiple independent gate transistors.

III. MODEL DEVELOPMENT

To develop a interpolating polynomial that goes through a set of data points such as \((x_0, y_0), (x_1, y_1), (x_m, y_m)\), a polynomial \(f(x)\) is defined as,
\[
y_i = f(x_i); \quad i = 1, 2, \ldots, m
\]

There are many different ways to formulate interpolating polynomials. One special case is using piecewise polynomial and this is known as Spline interpolation. This method has several advantages over a single high degree polynomial. One major problem with single high degree polynomial interpolant is Runge’s phenomenon i.e. occurrence of oscillation between points, especially at the boundary [18]. If the data set is regular and monotonic, low degree spline polynomials can show high predictive accuracy while avoiding Runge’s phenomenon. Cubic spline, as its name implies, has third degree piece-wise polynomials. It is not too complex yet is able to satisfy important continuity requirements and has all the aforementioned advantages unique to spline interpolation. For these reasons, multivariate cubic spline polynomials have been used in this work to develop models for multiple transistors [19].

Given \(n\) points in the plane, \((x_k, y_k), k = 1, 2, \ldots, n\) with distinct \(x_k\)’s, there is a unique polynomial in \(x\) of degree less than \(n\) whose graph passes through the points. If two successive points are \((x_k, y_k)\) and \((x_{k+1}, y_{k+1})\) then the \(k\)th interval between these two points can be interpolated using splines. The interval index \(k\) is such that, \(x_k \leq x < x_{k+1}\). The local variable, \(s\) is \(s = x - x_k\). The first divided difference is \(\delta_k = (y_{k+1} - y_k)/(x_{k+1} - x_k)\). Let, \(h_k\) denote the length of \(k\)th subinterval i.e. \(h_k = x_{k+1} - x_k\); then, \(\delta_k = (y_{k+1} - y_k)/h_k\). Let, \(d_k\) denote the slope of the interpolant at \(x_k\) i.e. \(d_k = P'(x_k)\).

The cubic spline polynomial will be the following function on the interval \(x_k \leq x \leq x_{k+1}\), written in terms of local variables \(s\) and \(h = h_k\).

\[
P(x) = \frac{3hs^2 - 2s^3}{h^3} y_k + \frac{h^3 - 3hs^2 + 2s^3}{h^3} y_k + \frac{s^2(s - h)}{h^2} d_{k+1} + \frac{s(s - h)^2}{h^2} d_k
\]

The advantage of this cubic polynomial formulation is that it can satisfy four interpolation conditions; two on the function values and two on the derivative values so that, \(P(x_k) = y_k, P(x_{k+1}) = y_{k+1}\). \(P'(x_k) = d_k, P'(x_{k+1}) = d_{k+1}\).

Now the values of \(d_k\)’s are determined in a way as to satisfy the second order continuity of the spline polynomial. Using this added constraint we get,
\[
h_k d_{k-1} + 2(h_k - 1 + h_k) d_k + h_{k-1} d_{k+1} = 3(\delta_k - 1 + h_k \delta_{k-1})
\]

If knots are equally spaced, equation 2 becomes
\[
d_{k-1} + 4d_k + d_{k+1} = 3\delta_{k-1} + 3d_k
\]

If we apply the procedure described above at each interior knot \(x_k, k = 2, \ldots, n - 1\), we will get \(n - 2\) equations with \(n\)
unknowns. To get the required $n$ equations we have used not-a-knot approach at the boundary of the interval. For the first and last two subintervals, a single cubic polynomial is used. These two end conditions provide us with two more required equations to uniquely determine the values of $d_k$’s.

Tensor product formulation can be used to extend the same analysis for multiple independent variables. To determine the value of an arbitrary desired point, we use cubic interpolation of the values of closest knot points in each dimension. There is no limitation on the number of independent variables in multivariate cubic spline formulation. In this paper, we have used terminal gate and drain voltages as independent variables for spline interpolation. Variables related to device geometry such as $W$ (width) and $L$ (length) may also be used as independent variables to include the effect of device geometry on I-V characteristics. Though we have restricted ourselves to DC I-V characteristics, the same methodology can also be employed to model C-V characteristics to enable transient simulation.

IV. MODEL VALIDATION

DC I-V characteristics of five different transistors with very different underlying physics including MOSFET, MESFET, HBT, HEMT and a novel silicon-on-insulator four-gate transistors ($G^4FET$) have been modeled using multivariate cubic spline polynomial. Analytical expression, TCAD simulation and experimental results have been used for the development and validation of the model. Mean model error, $E$ has been measured as,

$$E = \left(\frac{\sum_{i=1}^{n} |I_M - I_0|}{I_0 + \epsilon}\right)/n \times 100\%$$

Here, $I_0$= test data, $I_M$= model prediction, $n$= number of data points and $\epsilon$ is a small number below which errors are progressively scaled to zero. $\epsilon$ is used to make sure that the error estimate is not unduly influenced by lower currents and is able to reject outliers in data, especially noisy data at low bias levels [20].

A. Short-channel (deep sub-micron) MOSFET

To validate the performance of the model for short-channel MOSFETs, three sets of data for three very short channel lengths 22 nm, 32 nm and 45 nm have been generated using TCAD Sentaurus. The substrate background doping level and the peak concentration of the halo implant have been defined as $5 \times 10^{15}/cm^3$ and $4 \times 10^{18}/cm^3$, respectively. The junction depth for the halo, the extension and the source/drain implants are set as 0.05 $\mu$m, 0.008 $\mu$m and 0.04 $\mu$m, respectively. The nitride spacer length, the gate oxide thickness and the high-k oxide thickness have been defined as 0.014 $\mu$m, 0.0006 $\mu$m and 0.0020 $\mu$m, respectively. All I-V simulations are performed using hydrodynamic transport model, where the carrier temperature equation for the dominant carriers is solved together with the electrostatic Poisson equation and the carrier continuity equations.

Fig. 1 shows the surface plot of $I_D’$ as a function of $V_{DS}$ and $V_{GS}$. The model parameters are extracted based on these training data following the procedure outlined in section III. Then drain currents, $I_D’$ for new drain and gate voltages are calculated using the multivariate cubic spline. A comparison between the isolines predicted by the cubic spline model and a new set of test data obtained from TCAD Sentaurus is shown in Fig. 2. It shows reasonably good agreement between TCAD data and model prediction.

The same measurements have also been done for 32 nm NMOS in TCAD Sentaurus. The training data set is shown in Fig. 3 and the test results for different set of gate voltages is shown Fig. 4. The relatively small value of corresponding mean model error shows good matching between the test data and the model prediction.

To demonstrate the range of validity of the model, simulations have also been performed for 45 nm channel length NMOS device. The training data set is shown in Fig. 5 and the results for a different test set of gate voltage is shown in Fig. 6. As seen from these figures, the model is capable of quite accurately predicting the MOSFET DC characteristics of sub-micron to deep sub-micron devices with severe short-channel effects.
B. Metal-Semiconductor-Field-Effect-Transistor (MESFET)

The Metal Semiconductor Field Effect Transistor (MESFET) consists of a conducting channel between a source and drain contact region. A Schottky metal gate controls the carrier flow by varying the depletion layer width underneath the metal contact which modulates the thickness of the conducting channel to modulate the current. The normalized drain current, \( I'_{DS} \) for the \( n \)-channel MESFET can be written as:

\[
I'_{DS} = \frac{I_{DS}}{(W/L)} = q\mu_n N_d d [V_{DS} - \frac{2}{3} (\phi_i - V_{GS} + V_{DS})^{2/3} \frac{\sqrt{V_p}}{V_{DS}} - \frac{(\phi_i - V_{GS})^{2/3}}{\sqrt{V_p}}] (1 + \lambda V_{DS}); V_{GS} - V_{TH} \geq V_{DS}
\]

\[
= q\mu_n N_d d [V_{GS} - V_{TH} - \frac{2}{3} (V_p - \frac{(\phi_i - V_{GS})^3}{\sqrt{V_p}})] (1 + \lambda V_{DS}); V_{GS} - V_{TH} \leq V_{DS}
\]

Here, \( V_{DS} \) = drain-source voltage, \( V_{GS} \) = gate-source voltage, \( V_{TH} \) = threshold voltage, \( W \) = channel width, \( L \) = channel length, \( \mu_n \) = mobility of electrons, \( N_d \) = channel doping density, \( q \) = electron charge, \( 1.6 \times 10^{-19} \), \( d \) = thickness of channel, \( \epsilon \) = electric permittivity of the channel material, \( \phi_i \) = built-in Schottky potential between gate metal and channel semiconductor and \( \lambda \) = channel length modulation parameter. The values of the parameters used in this simulation are, \( \mu_n = 1000 cm^2/(Vs) \), \( V_{TH} = -1.0579 V \), \( \phi_i = 0.8564 V \), \( V_p = 1.9143 V \), \( d = 49.79 nm \), \( \lambda = 0.05 V^{-1} \).

Fig. 7 shows the surface plot of \( I'_{DS} \) as a function of \( V_{DS} \) and \( V_{GS} \). Here, the dots are the data points calculated using equation 4. These are used to train the cubic spline model. Then the normalized drain current, \( I'_{DS} \) for unknown drain and gate biasing condition are calculated using the multivariate cubic spline polynomial extended for two variables. equation 4 is then used to produce another set of test data to validate model prediction accuracy. These new values are compared to the values predicted by the model. A comparison between the model predictions and results obtained from equation 4 is shown in Fig. 8. With relatively low mean error, the matching is quite good between model and test data. As evident from this figure, the matching gradually improves with the increase in gate bias.

C. Heterojunction Bipolar Transistor (HBT)

The heterojunction bipolar transistor (HBT) is a type of bipolar junction transistor (BJT) which uses different semicon-
The HBT has several improvements over the BJT such as it can operate at very high frequencies going up to a few hundreds of GHz. HBT is in demand for very fast modern RF (radio frequency) circuits. It is also used in RF power amplifiers in cellular phones to ensure high power efficiency. Detailed theory of heterojunction bipolar transistor was developed by Herbert Kroemer in 1957 [2]. Here, the results for \( n-p-n \) and \( p-n-p \) HBT are shown and the results for conventional homojunction \( n-p-n \) BJT are also included. All simulations are performed in Sentaurus Device using the hydrodynamic transport model, where the carrier temperature equation for the dominant carriers is solved together with coupled Poisson and carrier continuity equations. Doping-dependent mobility model is used, high-field saturation effects are accounted for and Shockley-Read-Hall and Auger generation-recombination models are activated.

First, we show the results for an analytic \( Si_{0.84}Ge_{0.16} \) n-p-n HBT device structure, which is created with the Sentaurus Structure Editor. The peak concentration of the base, emitter, and collector implant are defined as \( 5 \times 10^{18} \), \( 5 \times 10^{19} \) and \( 1 \times 10^{18} \text{ cm}^{-3} \), respectively. The training data set has been generated for six different base currents equally spaced between 1 \( \mu \text{A}/\mu \text{m} \) to 5 \( \mu \text{A}/\mu \text{m} \) and it consists of normalized collector current for collector-emitter voltage ranging from 0 to 1.5 V.

Then the developed model is tested against a test data set for four different base currents equally spaced from 1.4 \( \mu \text{A}/\mu \text{m} \) to 3.8 \( \mu \text{A}/\mu \text{m} \). The training data and test results for isolines are shown in Fig. 9 and Fig. 10, respectively.

A conventional homojunction \( Si \) n-p-n BJT has also been modelled using cubic spline polynomial. The training data set for this device is shown in Fig. 11 and the comparison...
between model prediction and a new set of test data is shown in Fig. 12. As shown in these figures, the model works almost equally well for both Si BJT and SiGe HBT transistors.

**D. High Electron Mobility Transistor (HEMT)**

The multivariate cubic spline polynomial model is used to model the DC $I-V$ characteristics of InGaAs HEMTs [21][22]. TCAD Sentaurus is used to generate both the training and the test data sets. A schematic of the HEMT device structure used for TCAD simulation is shown in Fig. 13. The parameters used in HEMT layer structure are as follows: Substrate = 0.8 µm GaAs, channel = 10 nm InGaAs, spacer = 34.5 nm AlGaAs, cap layer = 30 nm GaAs, passivation layer = 50 nm nitride, location of delta-doping layer = 31 nm, thickness of doping layer = 2 nm. The Schottky gate contact is etched into the top spacer layer. This gate recess is 15 nm deep and gate length is 0.25 µm. At each side of the gate, a 40 nm wide oxide layer isolates the gate from the cap layer. The Schottky contact is defined as a Schottky contact with a Schottky barrier height of 0.9 eV and electron and hole recombination velocities of 107 cm/s.

The training data set has been generated for six different gate voltages ranging from -0.8 V to 0.8 V in 0.32 V increment. It consists for normalized drain current for drain voltage ranging from 0 to 1.5 V. Fig. 14 shows the surface plot of the training data. Then the developed model is tested against a test data set for five different gate voltages ranging from -0.64 V to 0.64 V in 0.32 V increment. The test results for isolines are shown in Fig. 14. The small relative errors for all five cases show excellent match between TCAD data and multivariate cubic spline model.

**E. Four gate silicon-on-insulator transistor ($G^4$FET)**

This model can be used to include multi-gates into the model development with relative ease while still being capable of predicting DC response quite accurately. As an example, a...
four gate silicon-on-insulator transistor known as $G^4FET$ has been chosen. $G^4FET$ is a relatively recent SOI (silicon-on-insulator) device [4] with four independent gates. Two of these gates are lateral junction-gates similar to a double gate JFET and two are vertical oxide gates similar to SOI MOS devices. The oxide gate on top plays the role of traditional MOS gate. The substrate bias acts as the bottom gate through FET action through buried oxide. The top and bottom gates are asymmetric unlike double gate SOI devices. These vertical gates induce accumulation/depletion/inversion of free carriers close to oxide interfaces. A conventional $p$-channel SOI MOSFET can be transformed into a $n$-channel $G^4FET$ using two body contacts on the opposite sides of the channel. The $p+$ doped source and drain of a $p$-channel MOSFET act as lateral junction-gates and provide FET action by modulating the width of the channel. The heavily $n+$ doped body contacts are used as the source and the drain for the $n$-channel $G^4FET$. This is how an inversion-mode, $p$-channel MOSFET can be transformed into an accumulation/depletion-mode $n$-channel $G^4FET$.

Few examples of experimentally demonstrated innovative circuit applications using this device are analog multipliers [23] [24], temperature compensated voltage references [25], digital applications such as multi-threshold inverter, universal and programmable gate, three transistor full adder [26] etc. A simple two transistor tunable negative resistance circuit built with G4FETs has been used to realize LC oscillators and Schmitt trigger circuits [27] and its tent map shape characteristics can be leveraged to build chaotic logic gates [28] suitable for hardware security applications [29]. $G^4FET$ has been credited with the inspiration behind the development of electrostatically formed nanowires employed for threshold logic functions [30] [31] and high sensitivity gas sensing and femtomolar bio-marker detection [32] [33].

To capture the current-voltage (I-V) characteristic of a $G^4FET$, cubic spline polynomial was used to derive the numerical model from the available data set. Current-voltage data for different bias values have been gathered from both experiment and TCAD Sentaurus for both $p$-channel and $n$-channel $G^4FET$ transistors. The current-voltage characteristics predicted by the model are then tested against another set of test data. Here, we briefly show the results for two cases. The first case consists of an $n$-channel $G^4FET$ simulated with TCAD Sentaurus. The training data set has been generated for different drain-source ($V_{DS}$), top gate ($V_{TG}$), bottom-gate ($V_{BG}$) and junction gate ($V_{JG}$) voltages. For simplicity, both junction gates are tied together and then based on these data, drain current ($I_{DS}$) regression model is developed as a function of four voltages ($V_{DS}, V_{TG}, V_{BG}, V_{JG}$). The test results for different top gate voltages is shown in Fig. 16. To demonstrate the predictive power of the model, we have also used experimental measurement from a PDSOI $p$-channel $G^4FET$ with width and length of 0.4 $\mu$m and 0.9 $\mu$m, respectively to develop a cubic spline model. Then the model is tested against a new set of test data is shown in Fig. 17. The corresponding mean error for different operating conditions show quite accurate match between the experimental data and the model. A more detailed exposition on this work has been reported in [15].

![Fig. 16: Comparison between model and TCAD data for $n$-channel $G^4FET$](image1)

![Fig. 17: Comparison between model and experimental data for $p$-channel $G^4FET$](image2)

## V. Conclusion

In this work, a multivariate cubic spline polynomial model is used to develop DC model of four different modern transistors and subsequently validated using TCAD data and experimental results illustrating its accuracy and versatility. The model with its first and second order continuity and validity in all operating regimes is suitable for circuit simulator implementation. Although the independent variables considered in this work include terminal voltages/currents, the same procedure can be used to extend the model by incorporating more variables such as geometric dimensions, terminal capacitances etc. for scalable transient and ac simulation. As demonstrated with five transistors with very different physical conduction mechanisms, the developed model is quite general and can be used in modeling any emerging single/multi-gate device for which accurate physics based simplified compact model is either too complex to be efficiently evaluated or has not yet been developed.

## References
