Parameter extraction - methods and status

J. Krause, A. Pawlak, M. Schroter
Chair for Electron Devices and Integrated Circuits (CEDIC)
University of Technology Dresden, Germany

krause@iee.et.tu-dresden.de
http://www.iee.et.tu-dresden.de/iee/eb/eb_homee.html

12th HICUM Workshop, Newport Beach, CA (USA), 2012

Outline

- Introduction
- Extraction flow
- Basic concepts
- List of test structures
- Joint extraction of RE and Rth
- Transit times
- Transfer current
- Summary

Introduction

- increasing number of physical effects and demand for highly accurate large- and small-signal compact models increases model complexity
  $\Rightarrow$ increased number of model parameters and correlation between physical effects
  $\Rightarrow$ parameter extraction becomes more difficult and effort increases

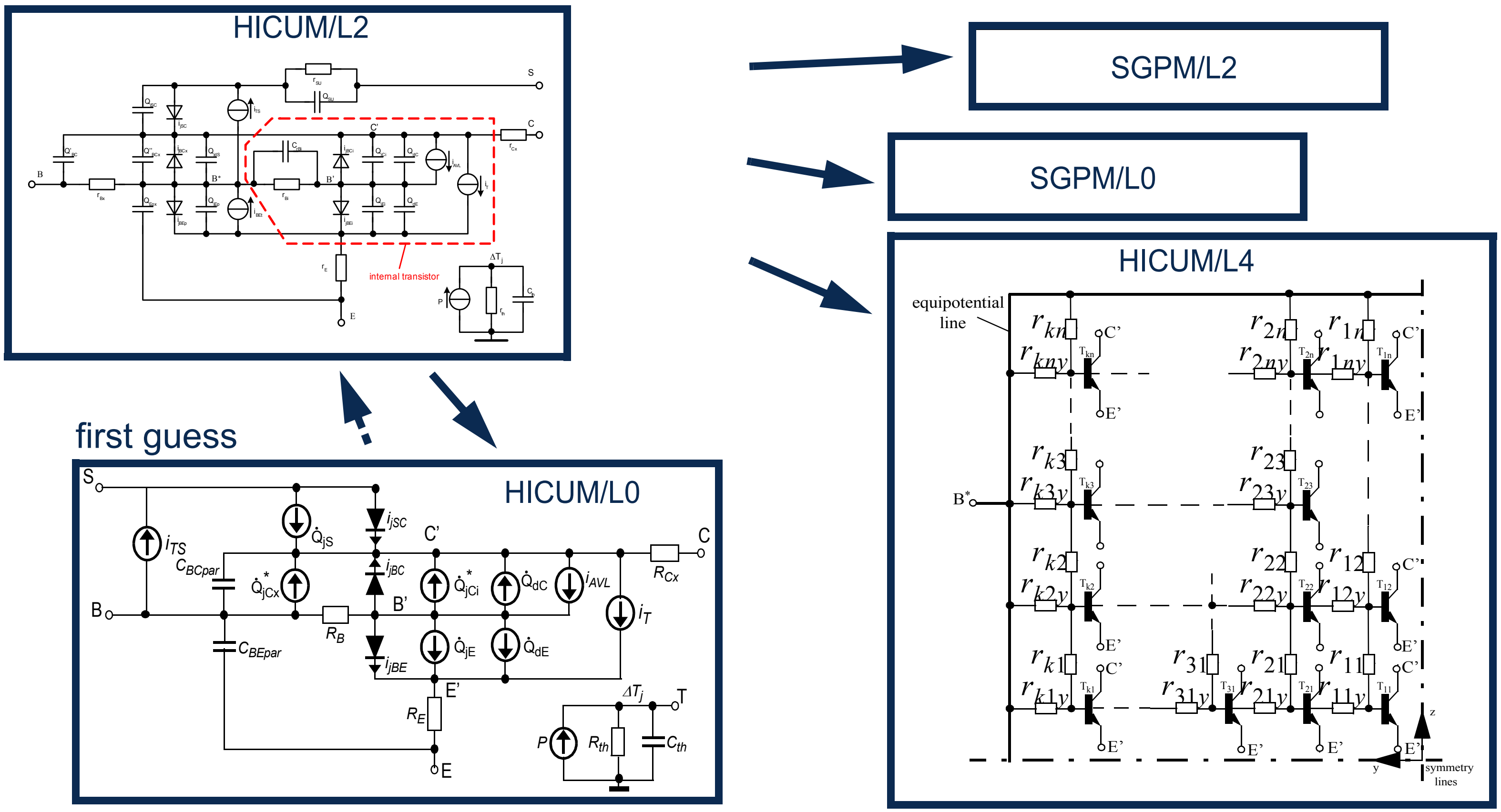
- model requirements from circuit design
  - accurate and valid over wide electrical, temperature, frequency, and geometry range
  - fast execution and numerically reliable

- requirements for parameter extraction
  - well-defined, sufficiently simple, fast and reliable procedures
  - standard measurement equipment should be sufficient for data acquisition
  - as small as possible model parameter correlation
  - clear and reliable procedures for parameter extraction
  - extraction procedues as automated as possible

- this presentation
  - review of general basic concepts and gerenal extraction flow
  - detailed discussion of extraction method for parameters of new HICUM/L2 version 2.31

Motivation

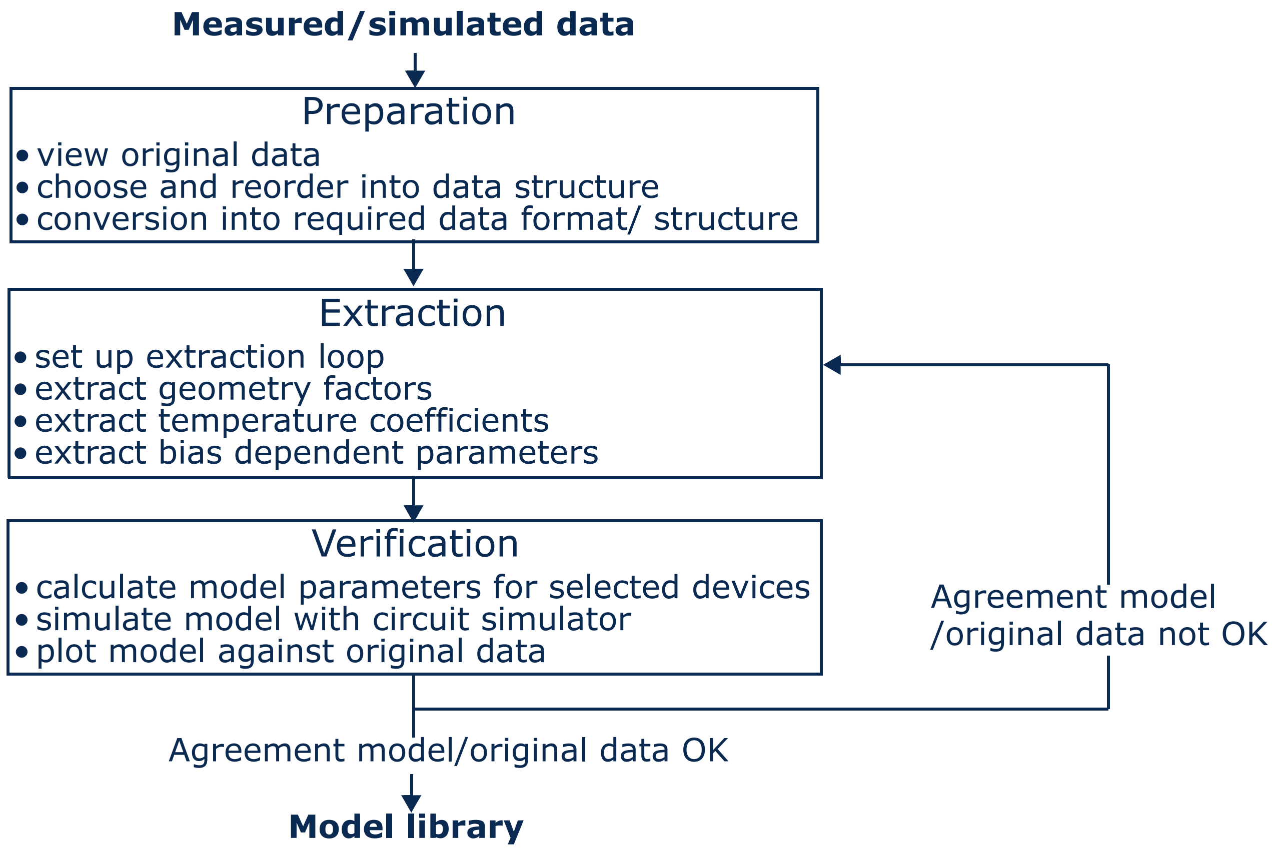
- Different model levels for finding a trade-off between calculation effort and accuracy
  - based on one compact model as physics-based and accurate as possible for generating the other model levels



Extraction flow

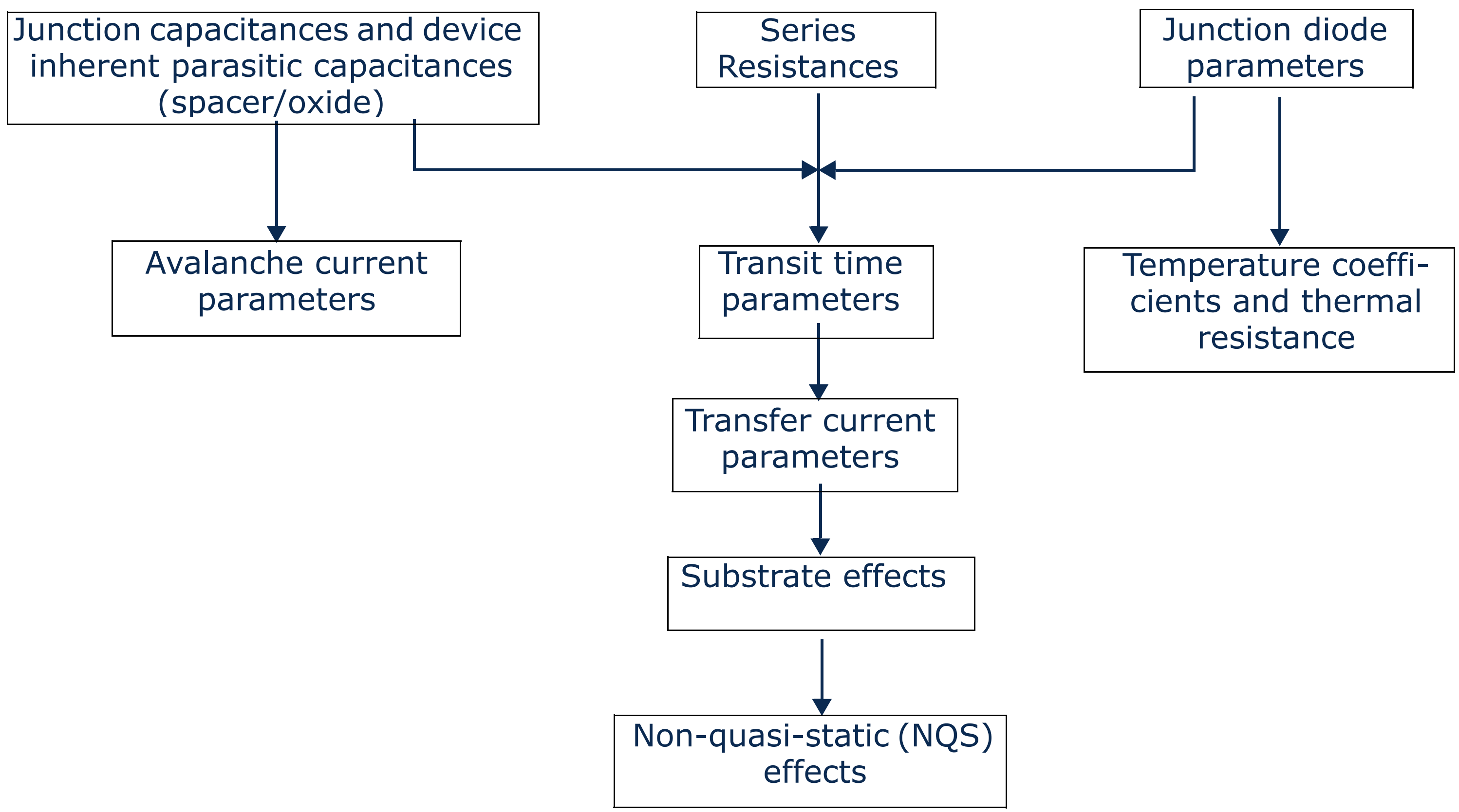
General methodology

- based on a set of data from measurements or simulation
- must be compared to results of compact model simulations



Extraction methodology for HICUM

- depending on availability of test structures and different geometries
- sequence for process-based geometry scalable model parameter extraction for HICUM/L2



Single geometry vs. scalable extraction

- conventional method (single geometry fitting)
  - no information on geometry effects, necessitating
    (a) simplified equivalent circuit and model $\Rightarrow$ loss of accuracy
    (b) non-physical model parameters (through optimization, merging …)
  - every transistor required (= anticipated) for circuit design has to be available on the wafer
  - requires "golden" wafer

=> no scaling and statistical modeling, no circuit optimization, large extraction effort

- process-based scalable approach
  - employs variety of transistors and special test structures => linear independent parameter extraction
  - requires somewhat higher initial investment: "few" devices have to be measured before the *first* parameter set can be generated => but very efficient model generation afterwards
  - physical values => enables statistcal modeling and shift to nominal parameters

=> enables scaling, statistical modeling, circuit optimization, significant reduction of overall extraction effort for foundries

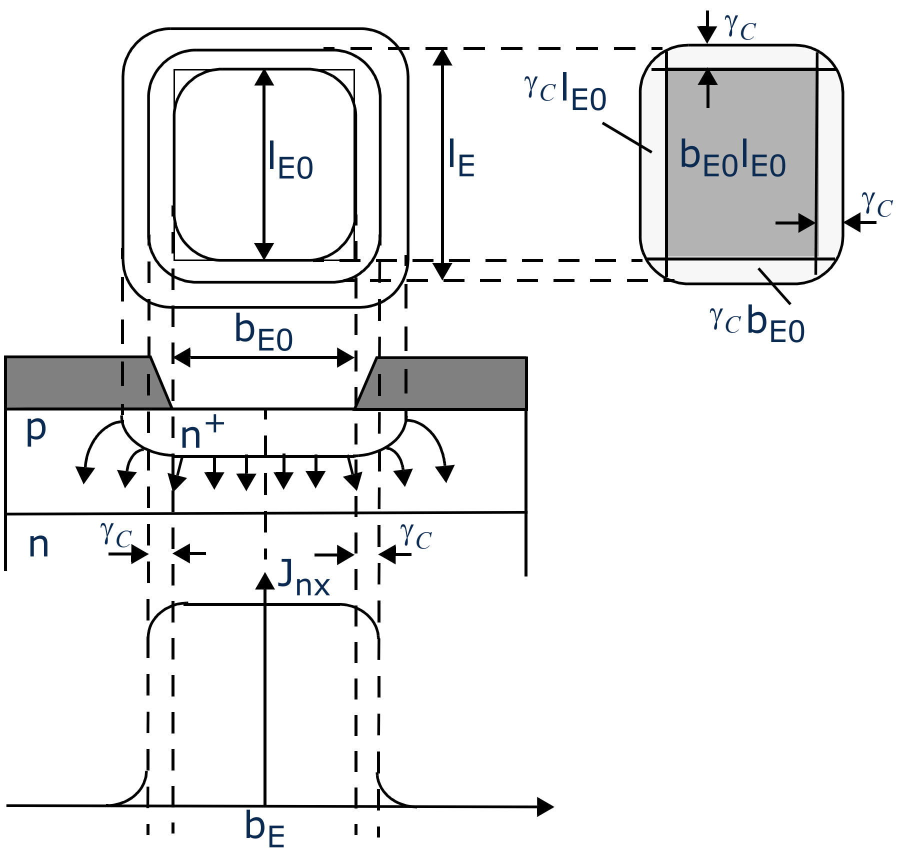
- scaling helps to understand physical effects and their geometry dependence

$\Rightarrow$ process-based scalable approach also aids process development

Basic concepts

Geometry Scaling

- γ_C, γ_B - factors for taking into account collector and emitter perimeter related current



- single transistor representation
- lumping area and perimeter related portions of collector current into a single component

  $\Rightarrow$ *effective* emitter area A_E:

$$A_E = A_{E0} + \gamma_C P_{E0} \text{ with } \gamma_C = \frac{J_A}{J_P}$$

- effective dimensions b_E, l_E:

$$b_E = b_{E0} + 2\gamma_C \text{ and } l_E = l_{E0} + 2\gamma_C$$

- include emitter corner rounding where required
- similar definition for γ_B, γ_{jE}

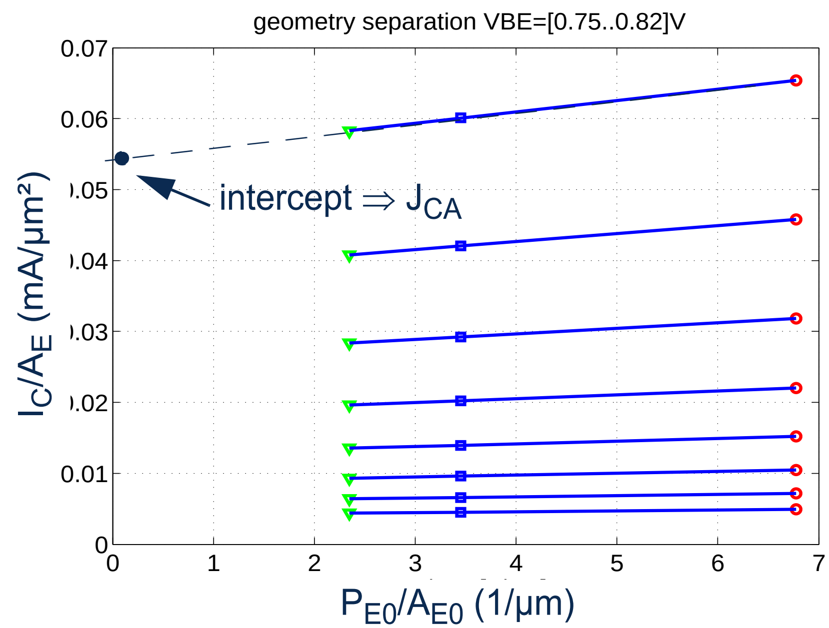
Standard geometry scaling

- Determination of process specific parameters
- Example: base current component from BE diode
  - Investigation of the characteristics, where only one junction is biased
    $\Rightarrow$ set V_{BC}=0V $\Rightarrow$ BE diode
  - use ***long*** structures for the extraction (i.e. l >> b)
  - scaling equations can be ***simplified*** to

$$I = J_A A + J_P P$$

$$\Rightarrow \text{ area normalization: } \frac{I}{A} = J_A + J_P \frac{P}{A}$$

  - plot I/A vs. P/A and extract
    - J_A from extrapolated intercept with y axis
    - J_P from slope



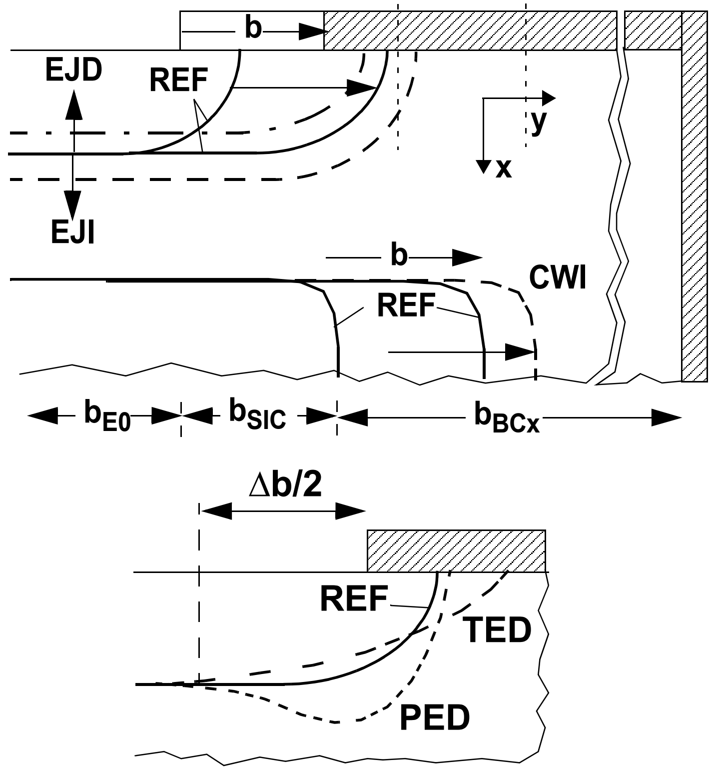
- apply same concept to charges and capacitances

$\Rightarrow$ determine process specific area and perimeter portion by simple linear regression

Geometry scaling pitfalls

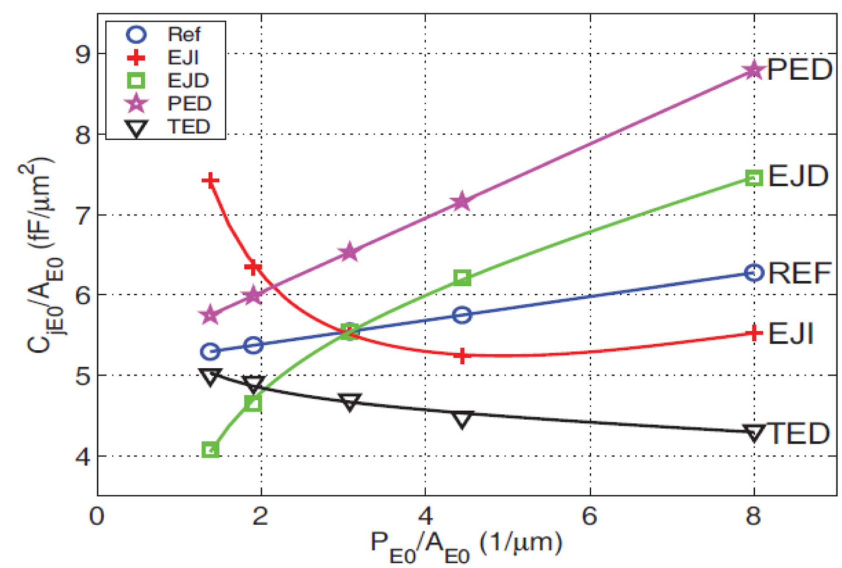
standard scaling

- linear scaling of electrical parameters of the internal transistor with emitter dimensions



non-standard scaling

- may lead to non-linear scaling with emitter dimensions or linear scaling with different (apparent) area and perimeter parameters
- caused by a variety of effects and fabrication conditions

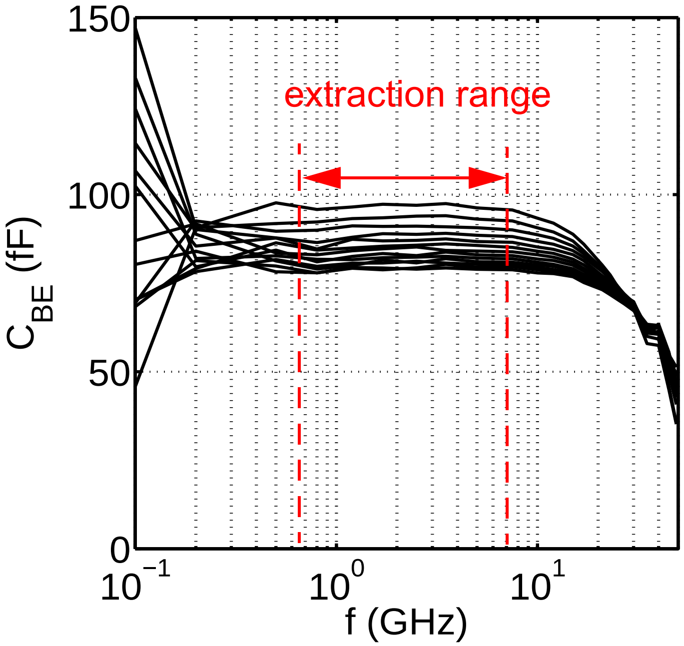


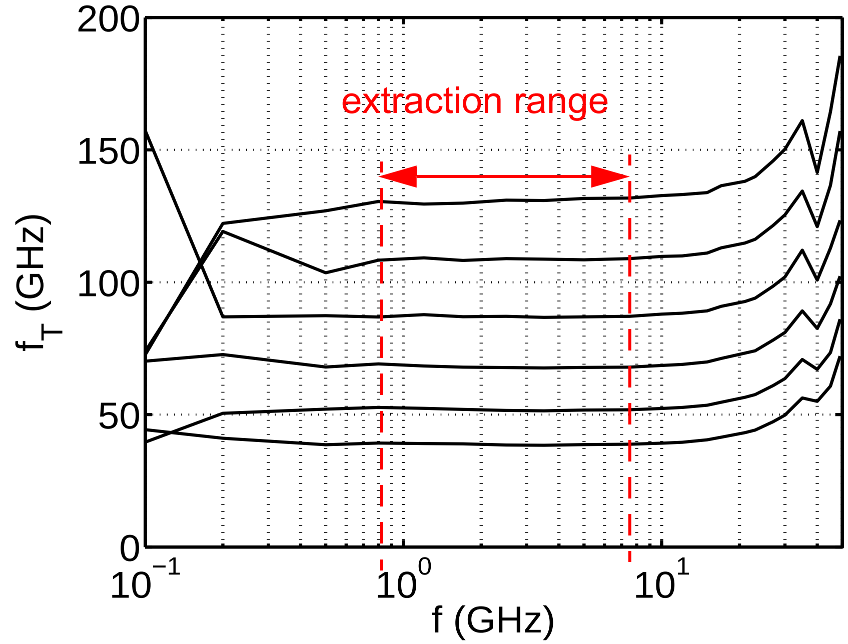
=> can cause significant additional effort for accurate scaling

Frequency range selection

... for determining dynamic quantities from measured data

$$C_{jE} = \frac{Im\{Y_{11} + Y_{12}\}}{\omega} \qquad f_T = \frac{f}{Im\{Y_{11}/Y_{21}\}}$$





- too low frequency: measurement inaccurate (noisy) and self-heating effects
- too high frequency: influence of RC time constants $\Rightarrow$ decrease in C_{BE}, increase in f_T
- Note: frequency range needs to be adapted to process performance

List of test structures

Special test structures

Type	Pad	Comments
Tetrodes - internal base sheet resistance - Q_{p0}	DC [AC]	- at least three (better four) widths: $b_{E0min} \geq 4{*}b_{E0min}$ - two different lengths for at least one width - structure with b_{E0}=0µm if possible $\Rightarrow$ base link
External base sheet resistance: - base contact resistance - silicided sheet resistance - poly-on-mono sheet resistance - poly-on-oxide sheet resistance [- base link resistance]	DC	necessary for extracting the external components of the base resistance and the scaling for different contact configurations and layouts later on - poly-on-mono and -oxide not required if fully silicided
Collector resistance components - buried layer - sinker and collector contact	DC	necessary for geometry scaling of collector resistance
Electro-thermal modeling	AC	verification of correct self-heating extraction and modeling of thermal coupling

- Note: these structures can also be used as process control monitors

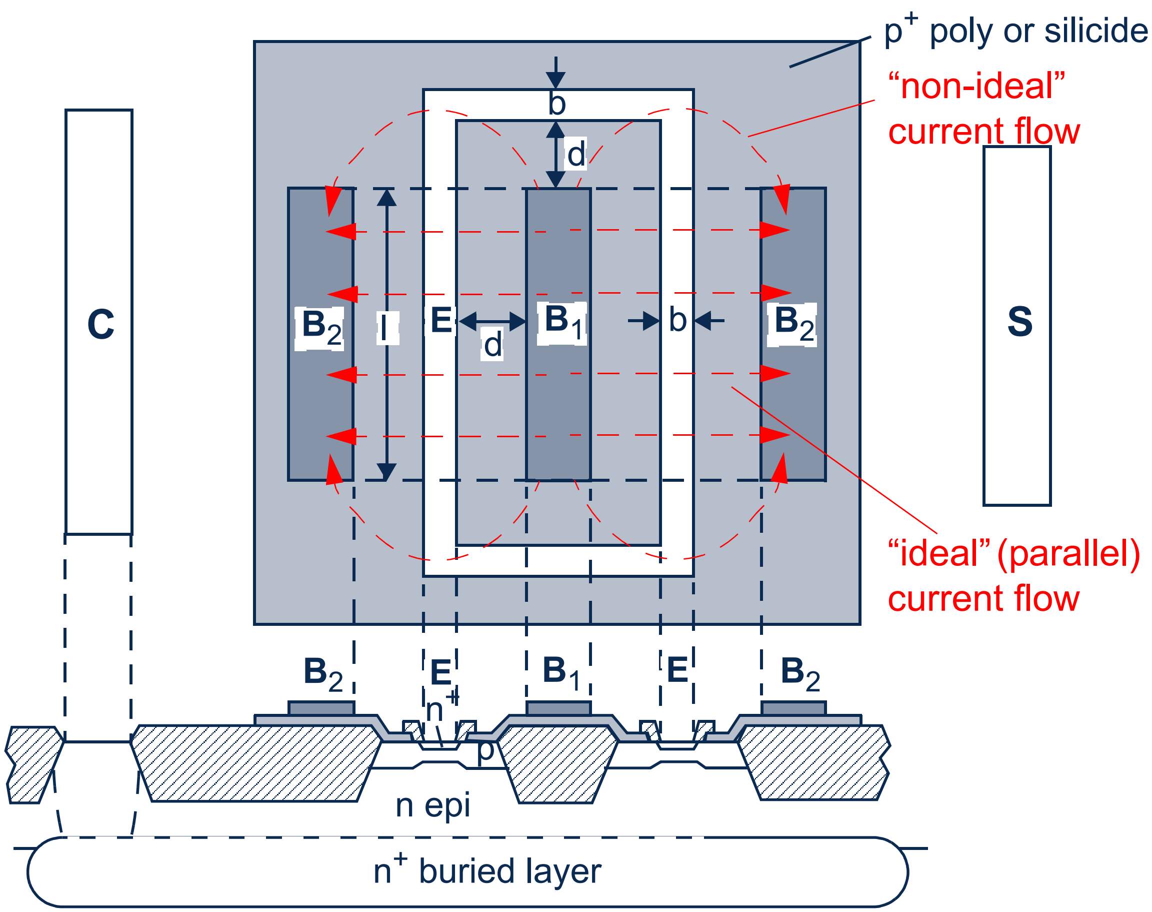
$\Rightarrow$ *linear independent* extraction of external/parasitic parameters and elements

List of test structures (cont'd)

npn transistors

Type	Pad	Comments
Long transistors with different emitter width in CBEBC configu-ration	AC	basic set for scalable parameter extraction ⇒ 3...5 widths with $b_{E0min} \geq 4*b_{E0min}$: - use for CV, DC and transit times - can also be used for R_E extraction
short transistors with typical dimensions for circuit design	AC	for verifying geometry scaling (e.g. BEC, ...)
CEB transistors	AC	single base transistors with b_{E0min} ⇒ geometry scaling of r_B
CBEBEBC transistors	AC	double-emitter transistors with b_{E0} = min, max ⇒ verification, finger and parameter scaling
Power cells: arrays of power transistors	AC	multi-emitter transistor arrays ⇒ for application and verification of distributed scalable modeling (HICUM/L4)
Deembedding structures: Open, Short, Through	AC	for transistors above, depend on pad layout and device size

Tetrode structure



- emitter encloses base contact B_1 $\Rightarrow$ no current flow through base poly between B_1 and B_1
- current between B_1 and B_2 is forced to flow underneath the emitter
- 2 different lengths $\Rightarrow$ correction of current spreading at the edges
- total resistance measured

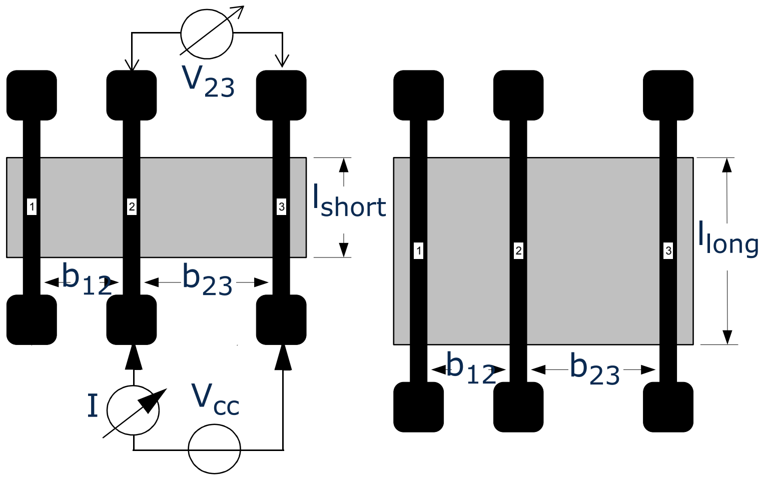
$$R = \frac{\Delta V_{B2B1}}{\Delta I} = r_x + \frac{r_{SBi}}{2}\frac{b_{E0}}{\Delta l}$$

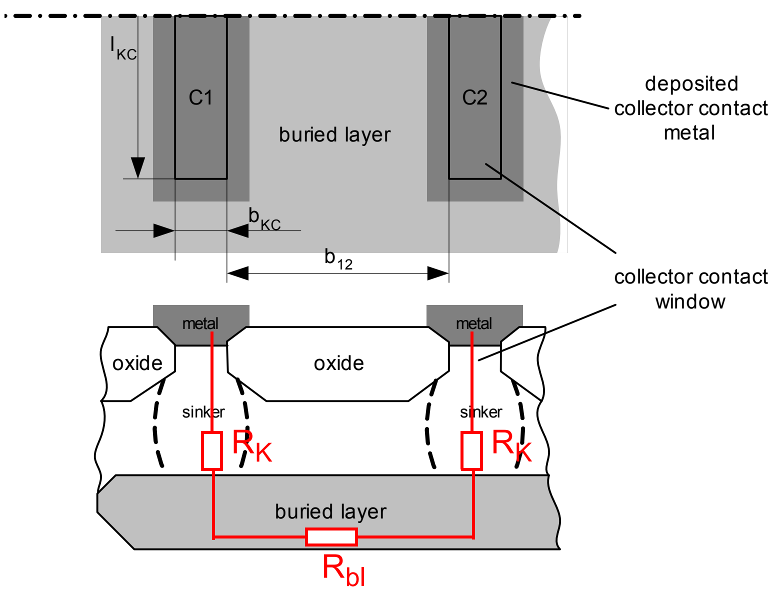
  - r_x: sum of all external resistance components

- application
  - extraction of base sheet resistance components (internal, link) and Q_{p0}
  - process control monitor for base region

Resistor structures

chains with 3...4 Kelvin contacts and two different lengths ($l >> b_{contact}$)





- apply voltage between two terminals i and j:

$$R_{ij} = \frac{V_{ij}}{I_{ij}} = 2R_K + R_{bl,\,ij} = 2\frac{r_{KC}}{l} + r_{Sbl}\frac{b_{ij}}{l}$$

- subtracting current of short from that of long structure removes non-ideal edge current

$$R_{\Delta ij} = \frac{V_{ij}}{I_{ij,\,long} - I_{ij,\,short}} = 2\frac{r_{KC}}{\Delta l} + r_{Sbl}\frac{b_{ij}}{\Delta l}$$

- edge corrected resistances for two pairs of terminals allow to determine sheet resistance...

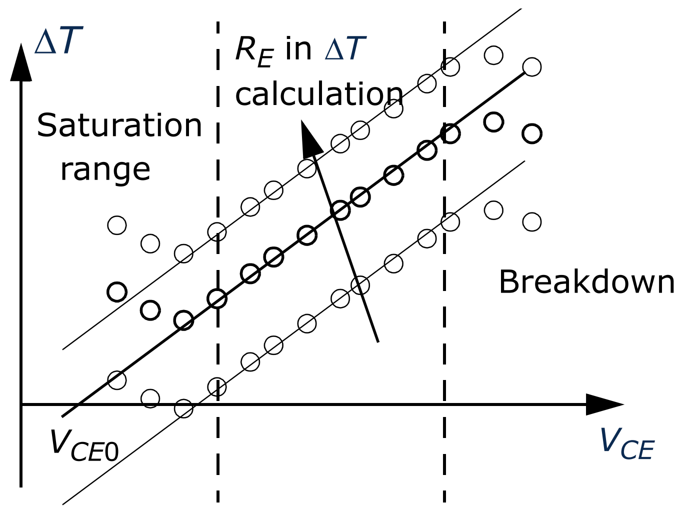
e.g., $r_S = (R_{\Delta 23} - R_{\Delta 12})\dfrac{\Delta l}{b_{23} - b_{12}}$

- ...and length specific contact resistance

Joint extraction of R_E and R_{th}

- Based on simplified self-heating model $\Delta T = R_{th} I_C (V_{CE} - R_E I_E - R_{Cx} I_C)$
- Variables to be known:
  - From measurement: I_C, I_E, V_{CE}, also needed: I_B, V_{BE}
  - From test structures or assumption: R_{Cx}
- Method:
  - Measurement of output characteristics with fixed I_C (or forced I_B)
  - Calculation of ΔT from measured V_{BE} and I_B and known dependence $I_B = I_{BES}(T) \exp\left(\frac{V_{BE} - R_E I_E}{V_T(T)}\right)$
  - In forward active operating range, almost linear dependence $\Delta T(V_{CE})$
  - Linear extrapolation to $\Delta T=0 \Rightarrow$ from equation above it follows from $R_{th}>0$ and $I_C>0$

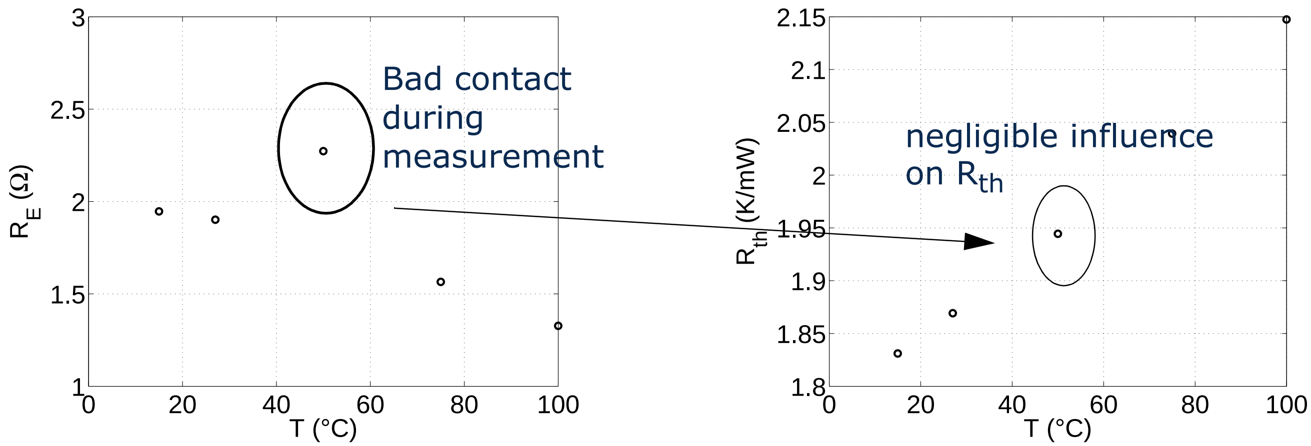
$$V_{CE0} - R_E I_E - R_{Cx} I_C = 0 \Rightarrow R_E = \frac{V_{CE0} - R_{Cx} I_C}{I_E}$$



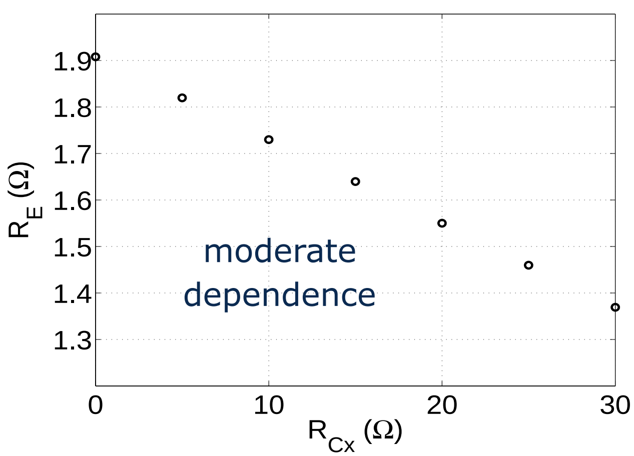
  - BUT: R_E already used for calculation of $\Delta T \Rightarrow$ Iteration for R_E
  - Calculation of R_{th} from slope of $\Delta T(V_{CE})$

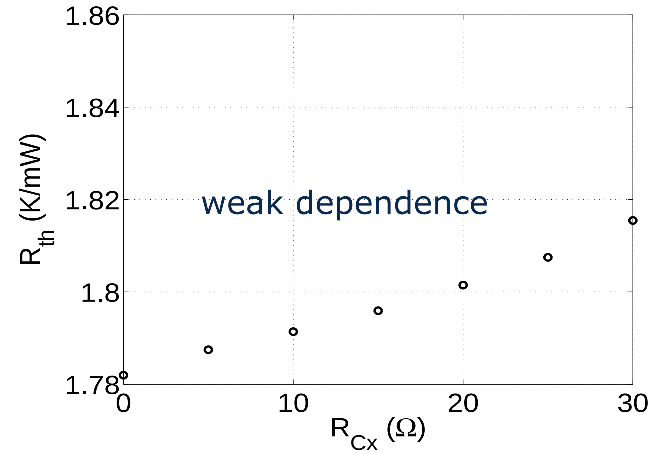
Results of combined R_E and R_{th} extraction

- Extraction for all temperatures separately



- Influence of uncertainties of previously extracted R_{Cx}



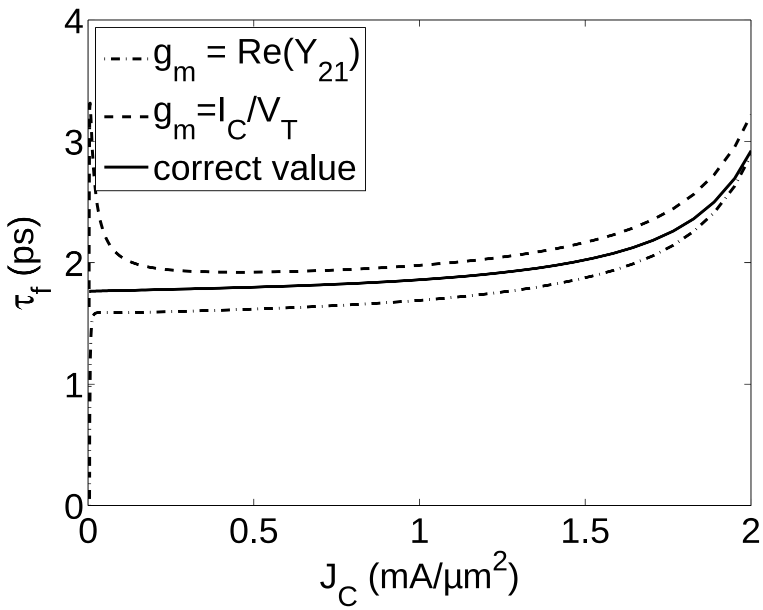


Transit times

- Determination of transit time τ_f from measurements:

$$\tau_f(V_{BCi}, i_{Tf}) = \frac{1}{2\pi f_T} - (R_E + R_{Cx})C_{BC} - \sum \frac{C_{BB}}{g_m}$$, C_{BB} = sum of all capacitances at B node

  - transconductance: $g_m = I_T / V_T$ (classical method) or $g_m = \lim_{f \to 0} Y_{21}(f)$.



- However: both g_m approaches lead to errors due to inconsistency of equation itself

  $\Rightarrow$ need to use complete deembedding of internal transistor according to compact model equivalent circuit

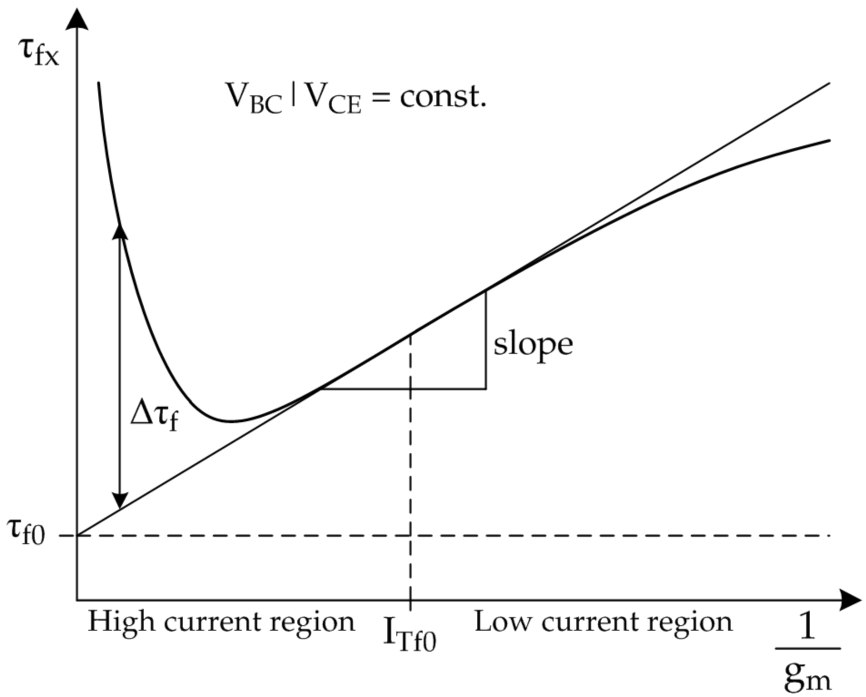
  - includes all parasitic time constants and impact of frequency dependence

- also: need to include bias dependence of (internal) depletion capacitances

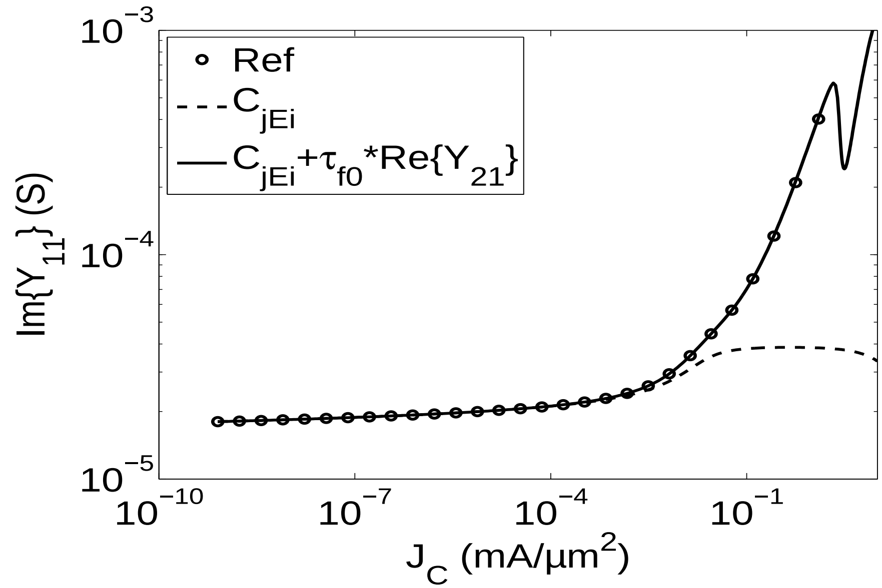
$\Rightarrow$ consistent deembedding method has been implemented and used

Extraction of τ_{f0}.

- Classical method uses linear extrapolation of $\tau_f(1/I_C)$ curve and a_{jEi} from slope at inflection point.



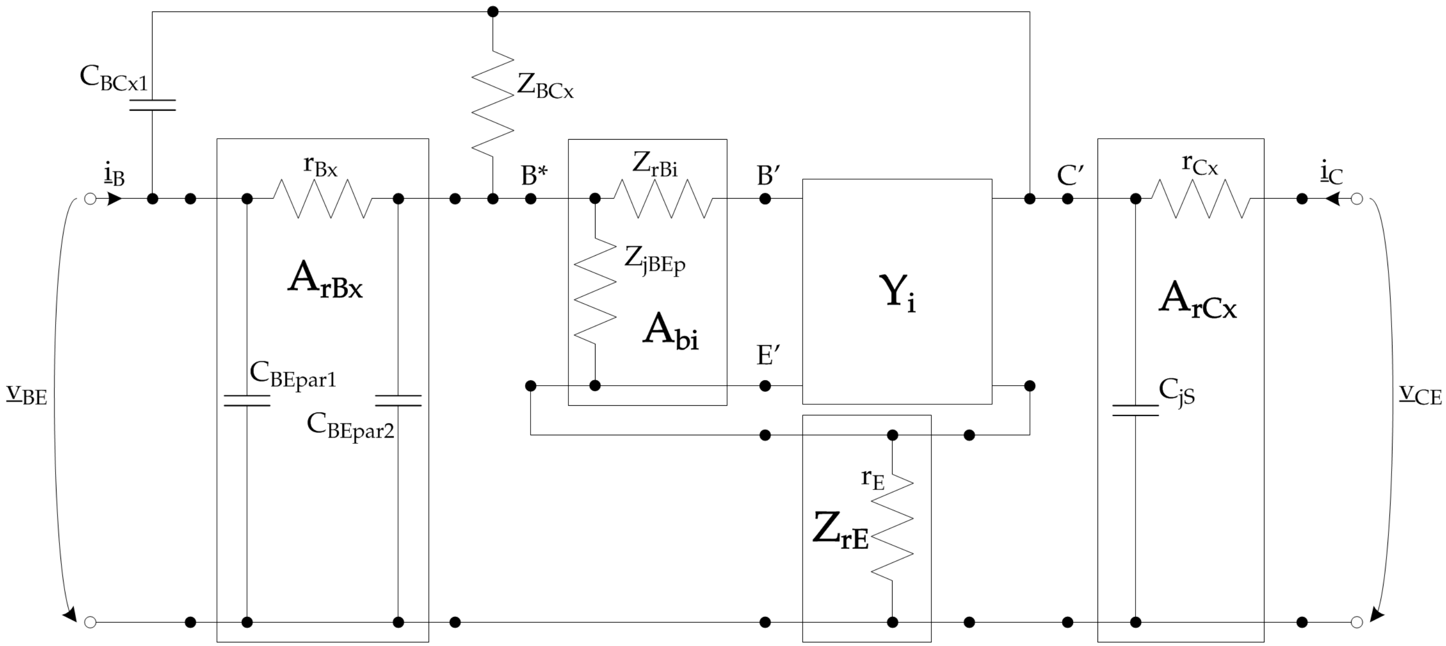
- Better: global fit of deembedded Y_{11} with C_{jEi} and τ_{f0} in the medium current range
  - $g_m = Re\{Y_{21}\}$ holds for each frequency for deembedded Y_{21}
  - No influence of time constants caused by series resistances
  - Fitting of a_{jEi} and τ_{f0} for the best agreement of resulting capacitance



=> consistency is key to accurate modeling

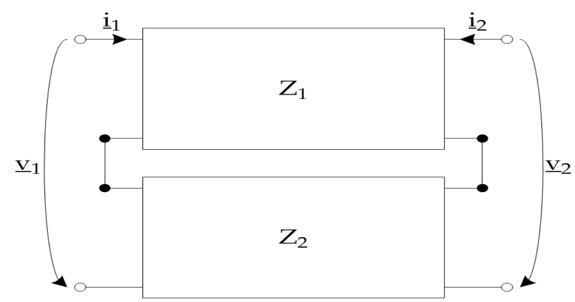
Deembedding of internal transistor

- use two-port parameters of each external element.



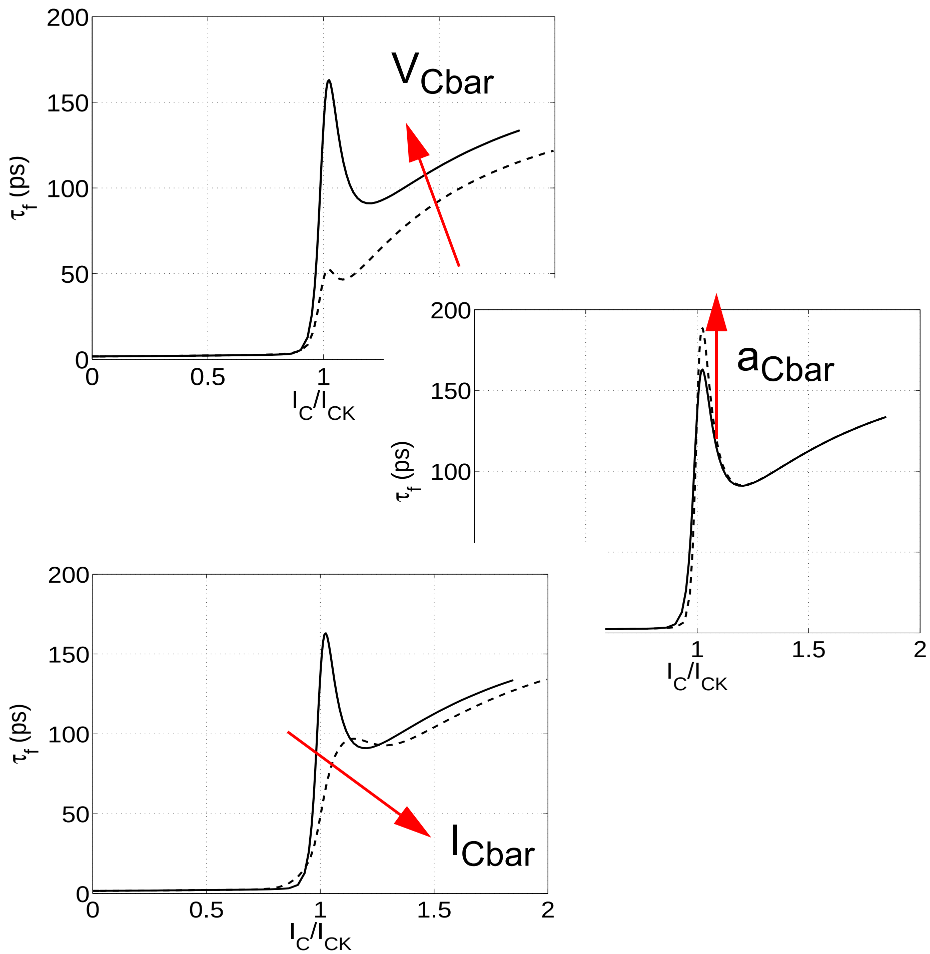
- example for deembedding of (impact of) R_E:

$$Z_{rE} = \begin{bmatrix} R_E & R_E \\ R_E & R_E \end{bmatrix} \quad \text{and} \quad Z_{int} = Z_{ext} - Z_{rE}$$

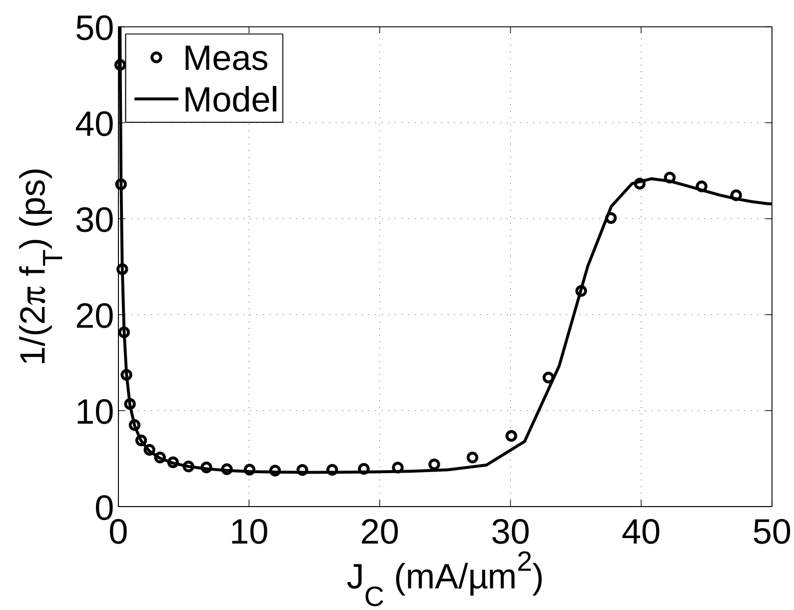


Extraction of barrier effect

- three model parameters
  - V_{Cbar}, a_{Cbar} and I_{Cbar} => shape of barrier effect.



Application to measurements



=> new formulation captures shape more accurately for HBTs

Transfer current

- forward bias ($V_{CE} \geq 0.2V$) HICUM/L2 transfer current equation reads:

$$I_T = \frac{c_{10}}{Q_{p0} + h_{jEi}Q_{jEi} + h_{jCi}Q_{jCi} + Q_{fT}} \exp\left(\frac{V_{BEi}}{V_T}\right)$$

- weighted minority charge is composed of different portions:

$$Q_{fT} = h_{f0}\tau_{f0}i_{Tf} + h_{fE}\Delta Q_{Ef} + \Delta Q_{Bf} + h_{fC}\Delta Q_{Cf}$$

$\Rightarrow$ Transfer current is related to the small signal parameters!

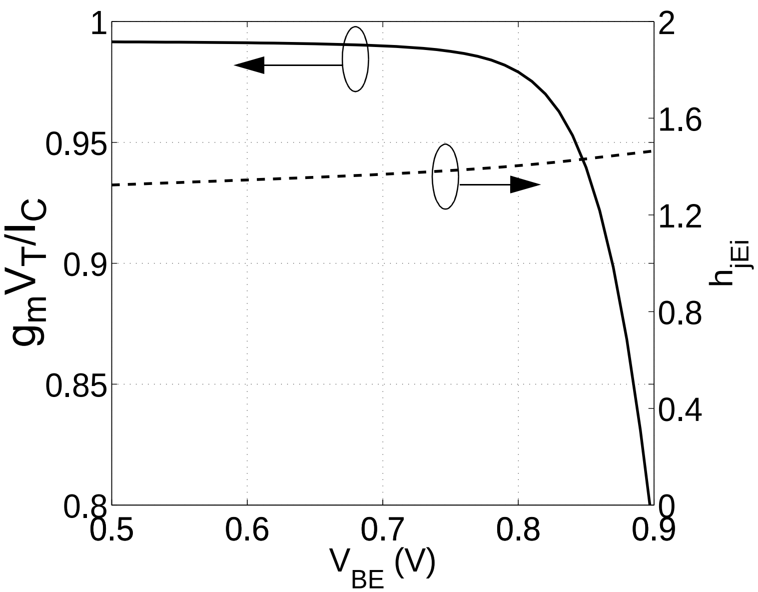
- extracted model parameters:
  - GICCR constant c_{10}
  - Weight factors h_{jEi}, h_{jCi}, h_{f0}, h_{fE}, h_{fC}
- required:
  - all parameters for C_{jEi} and C_{jCi} including temperature dependence
  - all parameters for the minority charges
  - Q_{p0} from tetrode measurements
    ($\Rightarrow$ may be set to arbitrary value, but this will lead to small deviations and also incorrect modeling of internal base resistance)

Low current region ⇒ parameters c_{10}, h_{jEi}

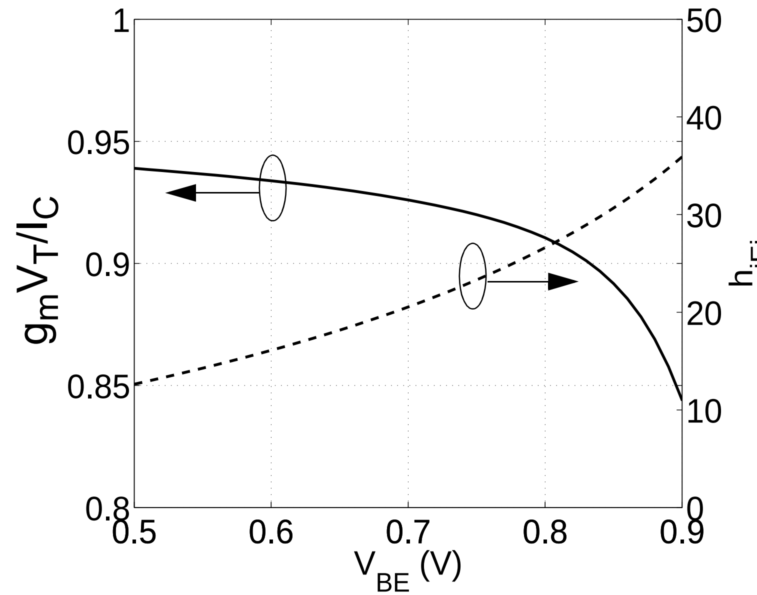
- extraction of bias *dependent* or bias *independent* h_{jEi} depending on normalized g_m

$$\frac{g_m}{I_T/V_T} = 1 - \frac{V_T}{Q_{pT}}\frac{dQ_{pT}}{dV_{BEi}}\bigg|_{V_{CE}} \quad \text{with} \quad \frac{dQ_{pT}}{dV_{BEi}}\bigg|_{V_{CE}} = h_{jEi}C_{jEi} + Q_{jEi}\frac{dh_{jEi}}{dV_{BEi}}$$

$$\text{and} \quad h_{jEi} = h_{jEi0}[\exp(u) - 1]/u \quad \text{with} \quad u = a_{hjEi}(1 - (V_{BEi}/V_{DEi})^{z_{Ei}})$$



normalized g_m is bias *independent*
=> h_{jEi} = const, $a_{hjEi} = 0$



normalized g_m is bias *dependent*
=> $h_{jEi}(V_{BEi})$, $a_{hjEi} > 0$

Reverse Early effect

- In case of $a_{hjEi} > 0 \Rightarrow$ separate extraction at low injection (V_{BEi}=V_{BE}) and V_{BC}=0

$$\text{set } I_S = \frac{c_{10}}{Q_{p0}}, \quad h(V_{BE}) = \frac{h_{jEi}(V_{BE})}{h_{jEi0}}, \quad V_{Er} = \frac{Q_{p0}}{h_{jEi0}C_{jEi0}}, \quad v_j = \frac{Q_{jEi}(V_{BE})}{C_{jEi0}}$$

$$\Rightarrow \quad I_T = I_C = \frac{I_S}{1 + h(V_{BE})v_j(V_{BE})/V_{Er}} \exp\left(\frac{V_{BE}}{V_T}\right) \quad \Rightarrow \quad \boxed{V_{Er} + hv_j = \frac{V_{Er}I_S\exp(V_{BE}/V_T)}{I_C}}$$

  - unknown constants V_{Er} and I_S can be removed by using four combinations of $I_C(V_{BE})$ values

$$\frac{h(v_1, a_{hjei})v_j(v_1) - h(v_2, a_{hjEi})v_j(v_2)}{h(v_3, a_{hjEi})v_j(v_3) - h(v_4, a_{hjEi})v_j(v_4)} = \frac{\left(\frac{\exp(V_{BE1}/V_T)}{I_{C1}} - \frac{\exp(V_{BE2}/V_T)}{I_{C2}}\right)}{\left(\frac{\exp(V_{BE3}/V_T)}{I_{C3}} - \frac{\exp(V_{BE4}/V_T)}{I_{C4}}\right)}$$

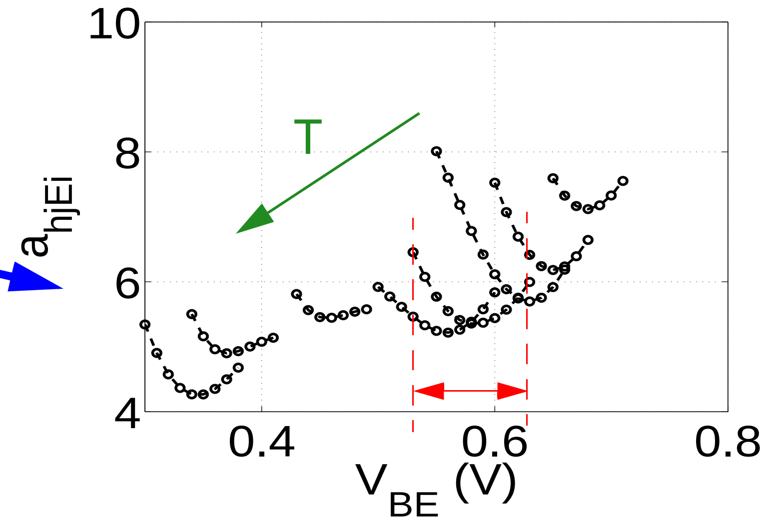
$\Rightarrow$ a_{hjEi} can now be calculated by solving non-linear equation above

- Note: differences between voltages may not be too small to avoid errors due to noise.In practical application, ΔV_{BE}=30 mV between each voltage was sufficient.

Extraction of a_{hjEi}

- extraction bias range criteria
  - V_{BE} large enough $\Rightarrow$ avoid too noisy results
  - V_{BE} low enough $\Rightarrow$ avoid errors due to high-current effects and self-heating
  - sweep center point within bias region

    $\Rightarrow$ a_{hjEi} = average within region

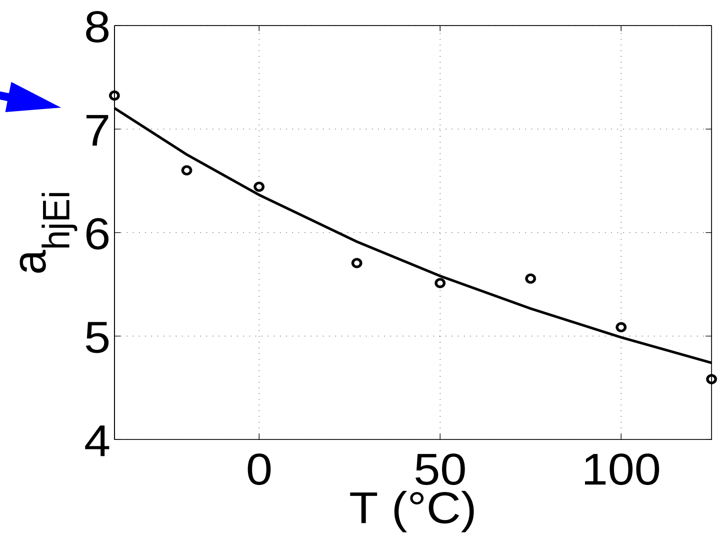


- perform extraction for each temperature
  - bias range to be adapted to temperature

    $\Rightarrow$ temperature dependence of a_{hjEi}
  - extraction of temperature coefficient ζ_{hjEi} according to model equation

$$a_{hjEi}(T) = a_{hjEi}(T_0)\left(\frac{T}{T_0}\right)^{\zeta_{hjEi}}$$

$\Rightarrow$ good agreement



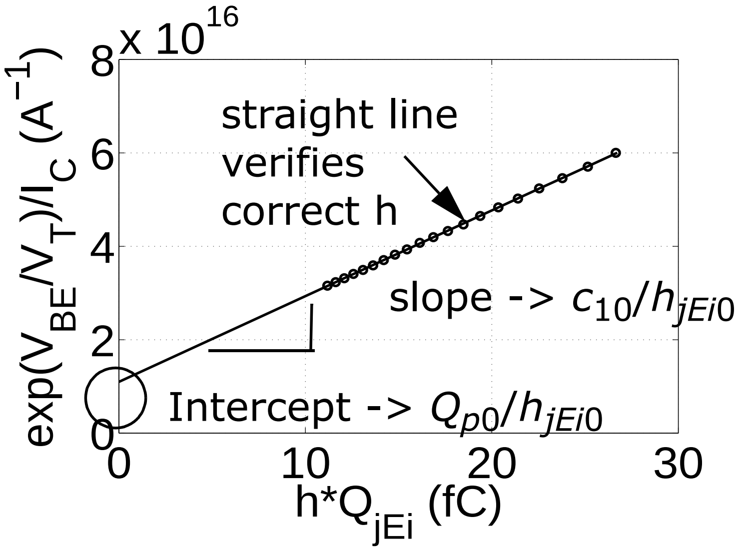
Extraction of h_{jEi0} and c_{10}

- Rewrite transfer current (at low injection):

$$hQ_{jEi} = -\frac{Q_{p0}}{h_{jEi0}} + \frac{c_{10}}{h_{jEi0}} \frac{\exp(V_{BE}/V_T)}{I_T}$$

  - $h = 1$ for $a_{hjEi} = 0$ or $h(a_{hjEi}, V_{BEi})$ for $a_{hjEi} > 0$

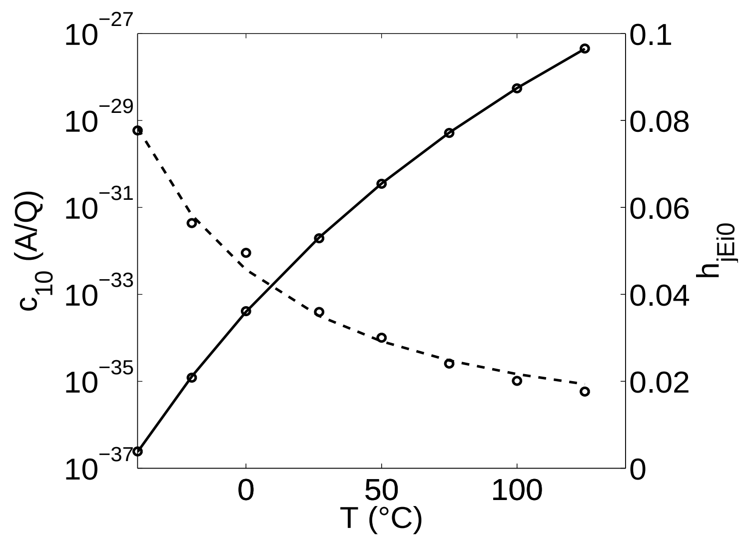
- plot $\dfrac{\exp(V_{BEi}/V_T)}{I_T}$ vs. hQ_{jEi}



- from linear fit:
  - h_{jEi0} from intercept with y-axis and known Q_{p0}
  - c_{10} from slope and known h_{jEi0}

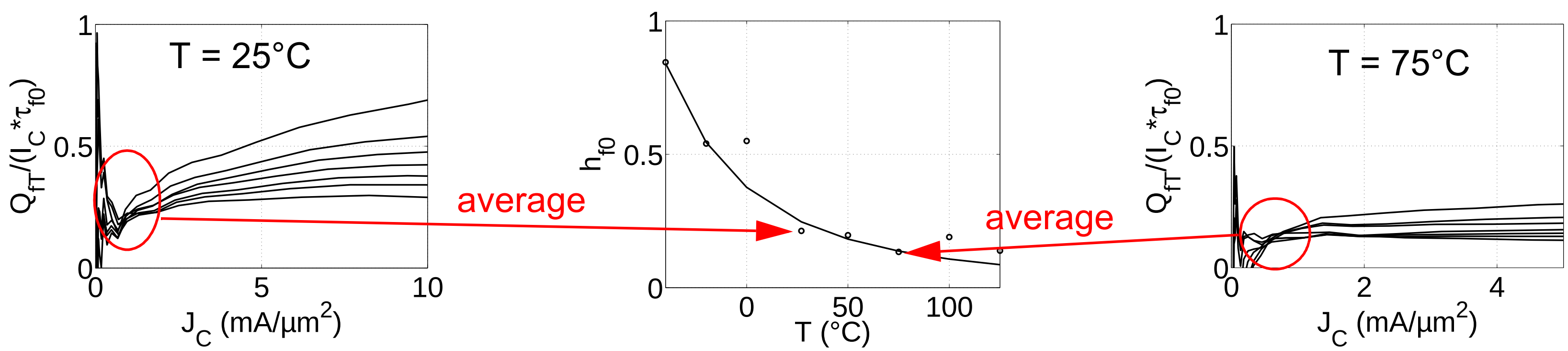
- performed for each temperature
$\Rightarrow$ extract ΔV_{gBE} and ζ_{vgBE} from

$$h_{jEi0}(T) = h_{jEi0}(T_0)\exp\left(\frac{\Delta V_{gBE}}{V_T}\left(\left(\frac{T}{T_0}\right)^{\zeta_{vgBE}} - 1\right)\right)$$



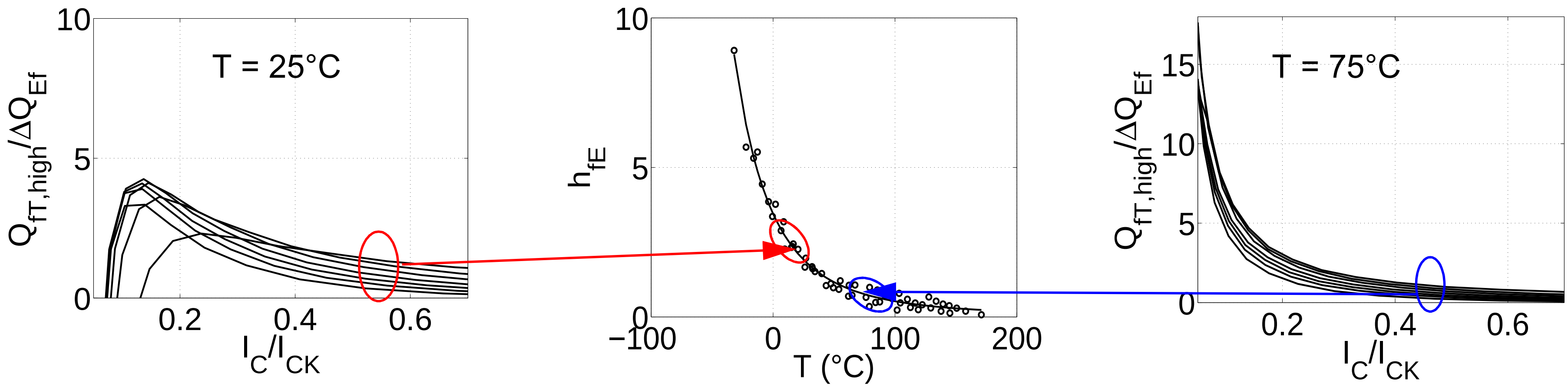
Extraction of mobile charge weight factor h_{f0}

- For reliable results, device temperature needs to be calculated:
  - $\Delta T = I_T V_{CEi} R_{th}$ with $I_T = I_C$ from measurements
    $\Rightarrow$ non-linear equation $\cdot\, I_C(V_{CE} - I_C r_{Cx}(\Delta T) - I_E r_E(\Delta T)) R_{th}$ solved at each bias point
- at $T = T_0 + \Delta T \Rightarrow$ $Q_{pT}(T) = \dfrac{c_{10}(T)\exp(V_{BEi}/V_T)}{I_C}$ with $V_{BEi} = V_{BEi} - I_B R_B(\mathrm{T}) - I_E R_E(T)$
- mobile charge for each operating point
  - $Q_{fT} = Q_{pT} - Q_{p0}(T) - h_{jEi}(V_{BEi}, T) Q_{jEi}(V_{BEi}, V_T) - h_{jCi} Q_{jCi}(V_{BCi}, T)$ @ $V_{BCx} = 0$
  - Extraction of $h_{f0} = \lim\limits_{I_T \to 0} \dfrac{Q_{fT}}{I_T \tau_{f0}} = h_{f0}(T_0)\exp\left(\dfrac{\Delta V_{gBE}}{V_T}\left(\dfrac{T}{T_0} - 1\right)\right)$ for each temperature



Extraction of h_{fE}

- calculate $h_{fE} = Q_{fT,high}/\Delta Q_{Ef}$ with $Q_{fT,high} = Q_{fT} - h_{f0}\tau_{f0}I_C$ at each temperature T
  - different device temperature due to impact of self-heating
    $\Rightarrow$ several h_{fE} for each ambient temperature
- plot h_{fE} vs. I_C/I_{CK}



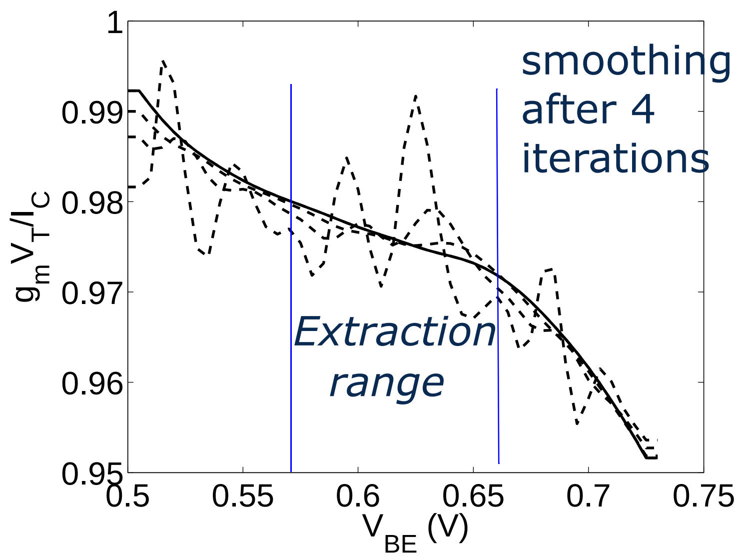
- select bias range where ΔQ_{Ef} has largest influence (e.g. I_T/I_{CK}=0.5)
  => h_{fE} = from each curve with corresponding self-heating corrected temperature
- apply similar procedure for extraction of h_{fC}

Summary

- increasing number of physical effects and demand for highly accurate large- and small-signal compact models increases model complexity
  $\Rightarrow$ increased number of model parameters
  $\Rightarrow$ increased parameter extraction effort
- automated extraction procedure and method depository $\Rightarrow$ reliable results
- inaccurate models can result from (in this priority)
  - inaccurate or inconsistent determination of device quantities (e.g. transit time)
  - inadequate parameter extraction methods (e.g. scalable vs. single device)
  - inaccurate measurements
  - inadequate model equations
- Note:
  - extraction methods and procedures depend on model equivalent circuit
    (e.g., the base impedance value is different in different models)
  - use (linear) independent measurements as much as possible for obtaining physically correct parameter values (e.g. cannot extract R_B, R_{CX}, parasitic caps from single device structure)
  - not all model parameters are needed for every process technology!
    $\Rightarrow$ you need to know which effects are relevant for your process and designers
    $\Rightarrow$ helps reducing parameter extraction effort

Smoothing of noisy measurement data

- smoothing method used: square interpolation



- Smoothing of noisy collector current based on normalized transconductance.