Cryo-CMOS Design and Validation for Quantum Computing Applications

Master's Thesis (PFE)

AQUA Laboratory — Rhonexum Project

Azel Bouraba

Thesis supervisor: Prof. Edoardo Charbon : edoardo.charbon@epfl.ch
Thesis co-supervisor: Hung-Chi Han : hung.han@epfl.ch

École Polytechnique Fédérale de Lausanne (EPFL)
Grenoble INP – Phelma, Université Grenoble Alpes
Politecnico di Torino

Lausanne, September 15, 2026

CONFIDENTIAL

Acknowledgements

I would like to express my sincere gratitude to all the members of Rhonexum for their warm welcome, support, and for making this internship such a valuable and enriching experience. I am particularly grateful to Hung-Chi Han and Vicente Charbon for their daily guidance, support, and advice throughout my internship, as well as for giving me the opportunity to begin my professional career at Rhonexum. I would also like to sincerely thank Vladimir Kopta and Christophe Enz for the weekly project meetings they held every Tuesday and for their valuable advice, which greatly contributed to the progress of my work. My deepest thanks go to Georgios Bantemits for the extensive technical guidance and expertise he shared with me throughout the different tasks I carried out, as well as for the many valuable insights and advice that have inspired me and will undoubtedly influence my future career. I am also very grateful to my fellow interns, Ramon Prats Consola and Gabriel Carre, with whom I shared the daily experience of this internship, collaborated closely, and worked together on numerous projects. Finally, I would like to thank my academic institutions, EPFL, Grenoble INP–Phelma, and Politecnico di Torino, for the education and opportunities that have prepared me for this experience, and finally my family for their constant support and encouragement throughout my studies.

Rhonexum

Rhonexum is a Swiss company developing cryogenic electronics for quantum computing. It is a spin-off of the Advanced Quantum Architecture Laboratory (AQUA) at EPFL, headed by Prof. Edoardo Charbon, and was founded by Vicente Carbon and Dr. Hung-Chi Han. The venture began as a project inside the laboratory a little over two years ago and was later incorporated in Lausanne, where it remains incubated at EPFL. In March 2026 the company closed a pre-seed round of USD 1 million led by QDNL Participations, with participation from Venture Kick and non-dilutive support from the EPFL Startup Launchpad, the Fondation pour l'Innovation Technologique and the Swiss National Science Foundation.

The positioning of the company answers a well identified bottleneck. Quantum processors operate at a few kelvin inside a dilution refrigerator, while their control electronics remain at room temperature. The number of coaxial cables required grows with the number of qubits and becomes an obstacle to scaling. Moving the control electronics into the cold environment, in the form of Cryo-CMOS integrated circuits, removes this obstacle. This approach faces a design difficulty. The compact transistor models supplied in commercial PDKs are validated only over the industrial temperature range, and the device behaviour below 77 K is not captured. A designer cannot verify circuit performance before tape-out, which makes the development cycle long and expensive.

Rhonexum addresses this gap with a modelling service called RX-Model. It combines automated low-temperature device characterisation, parameter extraction and a SPICE-compatible compact model built on the charge-based sEKV formulation and implemented in Verilog-A. The result is delivered as a cryogenic PDK that can be used in a commercial EDA environment and supports circuit simulation down to 4 K. The company targets both research institutions and industrial partners, and is in parallel preparing the internal design of selected cryogenic circuits with its own tools.

Given its age, the company has no division into departments or services. When I joined, the team consisted of five people: the two co-founders and three interns, including myself. This size implies a direct organisation without hierarchical intermediaries and daily communication between all members of the team. I was the first master's student hosted by the company for a thesis project, so the scope of the work was defined progressively as the project advanced.

Day-to-day supervision was provided by Hung-Chi Han, co-founder and chief technology officer of Rhonexum, who is also co-supervisor of this thesis. Academic supervision was provided by Prof. Edoardo Charbon. The work was carried out on the premises of the AQUA laboratory, where the company is incubated. My activity was placed at the interface between the development of the RX-Model platform and its application to circuit design, which corresponds to the validation step the company needs before the commercial release of its cryogenic PDKs.

Glossary

Acronyms and abbreviations

ABCD	Chain (transmission) matrix representation of a two-port.
AC	Alternating current; here the small-signal regime of the model.
ADE	Analog Design Environment, the simulation front-end of Cadence Virtuoso.
AHDL	Analog Hardware Description Language; the `ahdl_include` statement is what binds a Verilog-A module to a Spectre netlist.
AQUA	Advanced Quantum Architecture laboratory of EPFL, where this work was carried out.
BSIM-IMG	Berkeley Short-channel IGFET Model, Independent Multi-Gate. The reference compact model of the GF22FDX PDK.
CDF	Component Description Format. The Cadence description that declares the editable parameters of a cell, their defaults, the terminal order and the netlisting rules.
CEA-Leti	French research institute of the CEA, source of the OTA measurement used as the silicon reference of Section 3.1.2.
CIW	Command Interpreter Window, the SKILL console of Virtuoso.
CMOS	Complementary Metal-Oxide-Semiconductor.
CSV	Comma-Separated Values.
DC	Direct current; the static operating regime.
EDA	Electronic Design Automation.
EG	Extended-gate (thick-oxide) device option of the GF22FDX.
EKV	Enz–Krummenacher–Vittoz charge-based MOSFET model.
FDSOI	Fully Depleted Silicon On Insulator.
FFT	Fast Fourier Transform.
FinFET	Fin Field-Effect Transistor.
FOM	Figure of merit; here the gain–bandwidth–noise–power metric of Eq. (4.13).
GF22FDX	GlobalFoundries 22 nm FD-SOI technology platform, the target technology of this work.
HEMT	High Electron Mobility Transistor.
InP	Indium phosphide, the III-V technology of current cryogenic LNAs.
LNA	Low-Noise Amplifier.
LVT	Low threshold voltage device option.
MC	Monte-Carlo; the statistical simulation mode of the PDK.
MOSFET	Metal-Oxide-Semiconductor Field-Effect Transistor.
NF	Noise figure. Used only with this meaning in the text; the gate finger count is always written n_f, and `NF` in the code listings is the corresponding Verilog-A parameter name.
NMOS	n-channel MOSFET.
NQS	Non-Quasi-Static; the regime where the channel charge no longer follows the terminal voltages instantaneously.
OCEAN	Open Command Environment for Analysis, the batch simulation language of Cadence.

OTA	Operational Transconductance Amplifier.
PDK	Process Design Kit. The set of models, symbols and rules supplied by the foundry.
PEX	Parasitic Extraction; the post-layout extraction of interconnect resistance and capacitance.
ph	Photon. Appears only in the unit of the figure of merit, dB/mW/ph, which comes from the hf_0/kT_e factor of Eq. (4.13).
PMOS	p-channel MOSFET.
PSF	Parameter Storage Format, the binary result database written by Spectre and read back by OCEAN.
RF	Radio Frequency.
RMS	Root Mean Square.
RVT	Regular threshold voltage device option.
RX-Model	The cryogenic compact modelling and parameter extraction service developed by Rhonexum.
sEKV	Simplified EKV. The design-oriented reduced-parameter formulation of the EKV model, implemented here in Verilog-A and used throughout this report.
SG	Slim-gate (thin-oxide) device option of the GF22FDX, nominal supply 0.8 V.
SKILL	The proprietary Lisp-based scripting language of Cadence Virtuoso. Not an acronym.
SLVT	Super-low threshold voltage device option.
SOI	Silicon On Insulator.
SP	S-parameter analysis.
SPICE	Simulation Program with Integrated Circuit Emphasis; by extension, the class of circuit simulators and the netlist syntax they read.
TCAD	Technology Computer-Aided Design; device-level numerical simulation.
Verilog-A	Analog hardware description language used to write the sEKV model.
VM	Virtual machine.

Symbols

General notation and constants

Two conventions hold throughout. Upper-case italic symbols denote physical quantities (W, L, N_F, M), whereas the lower-case typewriter names (`w`, `l`, `nf`, `m`) are the corresponding Cadence CDF and Spectre parameter names. A leading Δ always denotes a variation or a geometrical offset, except for Δ_S which is the determinant of the scattering matrix; an overbar always denotes an average over the 4 to 6 GHz band or over a Monte-Carlo population.

Symbol	Meaning	Unit
x, y	Terminal indices of the four-terminal device, $x, y = D, G, S, B$ (or FG, BG for the two gates of the FDSOI device)	—
u	Position along the channel, from source to drain	m
f	Frequency	Hz
ω	Angular frequency, $\omega = 2\pi f$	rad/s
	Imaginary unit	—
q	Elementary charge	C

Symbol	Meaning	Unit
k	Boltzmann constant	J/K
h	Planck constant	J s
σ	Standard deviation of a Monte-Carlo distribution	—

Geometry and bias

Symbol	Meaning	Unit
W	Total gate width	m
W_f	Gate width per finger, $W/n_f + \Delta W$	m
W_{tot}	Total device width in a multi-finger sweep	m
$W_{M1..M3}$	Total width of the three LNA stages	m
W_1, W_2, W_3	Widths of the three devices T_1, T_2, T_3 of the transcapacitance testbench	m
L	Drawn channel length	m
L_1, L_2, L_3	Lengths of the three devices of the transcapacitance testbench	m
L_g	Effective gate length, $L + \Delta L$	m
ΔL, ΔW	Effective length and width offsets	m
n_f	Number of gate fingers (CDF parameter `nf`, Verilog-A parameter `NF`)	—
M	Instance multiplier (CDF parameter `m`)	—
V_{GS}	Gate–source voltage	V
V_{DS}	Drain–source voltage	V
V_{FG}	Front-gate voltage	V
V_{BG}	Back-gate voltage	V
V_{DB}	Drain–bulk voltage	V
V_{DD}	Supply voltage	V
$V_{G1..3}$	Gate bias of LNA stages 1 to 3	V
V_T	Threshold voltage	V
I_D, I_{DS}	Drain current	A
J_D	Drain current density, $\lvert I_D \rvert / (W \cdot M)$	mA/mm
J_{D1}	Drain current density of the LNA input stage	mA/mm

Small-signal parameters, charges and capacitances

Symbol	Meaning	Unit
g_m	Gate transconductance	S
G_{ms}	Source transconductance of the device, used in the degenerated small-signal expressions	S
g_{ds}	Output (drain–source) conductance	S
$g_{ds,n}$, $g_{ds,p}$	Output conductance of the NMOS and PMOS branches of the OTA	S
$g_{m,\mathrm{eff}}$, $g_{ds,\mathrm{eff}}$	Same quantities including inductive source degeneration	S
A_v	DC voltage gain of the OTA; $\overline{A_v}$ is its Monte-Carlo mean	V/V

Symbol	Meaning	Unit
Q_x	Charge on terminal x	C
q_x	Normalised charge, function of the bias point only	—
Q_{spec}	Specific charge of the sEKV normalisation	C/m^2
$\Delta\Phi_1$	Front-gate work-function difference	V
ψ_{s1}	Front surface potential along the channel	V
C_{xx}	Self-capacitance, $+\partial Q_x/\partial V_x$	F
C_{xy}	Transcapacitance, $-\partial Q_x/\partial V_y$, $x \neq y$	F
$C_{fg,fg}$	Front-gate self-capacitance	F
$C_{s,fg}$, $C_{d,fg}$	Source– and drain–to–front-gate transcapacitances	F
$C_{bg,fg}$	Back-gate–to–front-gate transcapacitance	F
C_{gs}, C_{gd}, C_{ds}	Gate–source, gate–drain and drain–source capacitances	F
C_{gg}	Total gate capacitance	F
C_{db}, C_m	Drain–bulk and transcapacitance term of the Y matrix	F
C_{ox1}	Front-gate oxide capacitance per unit area	F/m^2
C_{tot}	Total terminal capacitance returned by the simulator, intrinsic plus extrinsic	F
C_{int}	Intrinsic capacitance per unit channel area	F/m^2
C_{e1}, C_{e2}	Extrinsic coefficients scaling with W and with L	F/m
$C_{e1,s}$, $C_{e1,d}$	Source-side and drain-side split of C_{e1}	F/m
C_1, C_2, C_3	Total capacitance measured on the three testbench devices	F
T_1, T_2, T_3	The three devices of the transcapacitance testbench	—
c_{gsi}, c_{gdi}, c_{gbi}, c_{dbi}	Intrinsic elements of the EKV small-signal equivalent circuit	F
c_{gsov}, c_{gdov}, c_{gbov}	Gate–source, gate–drain and gate–bulk overlap capacitances	F
c_{sbj}, c_{dbj}	Source and drain junction capacitances	F
$R_{S\text{eff}}$, $R_{D\text{eff}}$	Effective source and drain access resistances	Ω
`CGSO`, `CGDO`	Gate–source and gate–drain overlap capacitance per unit width (model card parameters)	F/m
`CGBO`	Gate–substrate overlap capacitance per unit length	F/m
`cggi`, `cdgi`, `csgi`, `cbgi`	Intrinsic capacitances exposed by the PDK as operating-point variables	F

Gate resistance

Symbol	Meaning	Unit
R_{G}	Parasitic gate resistance	Ω
$R_{\square}$	Sheet resistance of the top gate conductor	Ω/sq
R_{ℓ}	Total resistance of one gate finger seen along its width, before lumping	Ω
α	Lumping factor of the distributed electrode, 1/3 for a single-side connection and 1/12 for a both-side connection	—
$I(u)$, I_{tot}	Current along the finger and total injected gate current, used in the derivation of α	A

Symbol	Meaning	Unit
ρ_v	Vertical (gate-stack interface) resistivity	$\Omega\mu m^2$
ρ_{int}, ρ_{vp}	Interfacial and top-layer contributions to ρ_v reported separately in the literature	$\Omega\mu m^2$
R_c	Resistance of one gate contact	Ω
n_c	Number of contacts per finger	—
w_c, p_c	Contact width and contact pitch	m
ρ_M, w_M, t_M	Resistivity, width and thickness of the strap metal	Ωm, m, m
s_{pp}	Gate-to-gate pitch	m
Δ_{ext}	Gate extension beyond the channel area	m
$R_{\min}$	Numerical floor on R_G	Ω
k_{gf}	Coefficient of the width-set access term	$\Omega\mu m$
R_{gf}	Width-set access term, $R_{gf} = k_{gf}/W_{tot}$	Ω
κ	Elmore constant of the non-quasi-static channel term	—
R_G^{S1}, R_G^{S2}	Two-term and single-term simplifications of R_G	Ω

Two-port and RF quantities

Symbol	Meaning	Unit
M1, M2, M3	The three common-source transistors of the LNA, from input to output	—
V_1, V_2, I_1, I_2	Port voltages and currents of a two-port, input and output	V, A
S_{11}, S_{22}	Input and output reflection coefficients	dB
S_{21}, S_{12}	Forward gain and reverse isolation	dB
Δ_S	Determinant of the scattering matrix, $S_{11}S_{22} - S_{12}S_{21}$	—
Y_{ij}	Admittance parameters of the two-port	S
Z_{in}	Input impedance seen at the amplifier access	Ω
Z_L	Load impedance closing the two-port	Ω
Z_0	Reference impedance, 50 Ω	Ω
Z_s	Impedance of the source degeneration inductor, ωL_s	Ω
A, B, C, D	Coefficients of the chain matrix	—
M_{series}, M_{shunt}	Chain matrices of an elementary series impedance and shunt admittance block	—
T_k, T_{LNA}	Chain matrix of block k and of the complete amplifier	—
μ, μ'	Edwards–Sinsky stability factors; unconditional stability for $\mu > 1$	—
K	Rollett stability factor, superseded here by μ	—
f_T	Transition frequency, unity current gain	Hz
$f_{\max}$	Maximum oscillation frequency	Hz
f_0	Band centre frequency, 5 GHz	Hz
f_h, f_l	Upper and lower band edges; their ratio $f_h/f_l = 1.5$ enters the FOM	Hz
Q	Quality factor of an inductor	—
L_{ind}	Value of a generic inductor of the passive network	H
R_s	Series loss resistance of an inductor, $2\pi f_0 L_{\text{ind}}/Q$	Ω

Symbol	Meaning	Unit
$L_{g,in}$	Series gate inductor of the input match	H
$L_{s,M1}$, $L_{s,M2}$	Source degeneration inductors	H
$L_{d1..3}$, L_{s23}	Drain load and inter-stage inductors	H
C_{out}	Output coupling capacitor	F
R_{d2}, R_{d3}	Drain load resistors of the second and third stages	Ω

Noise, power and statistics

Symbol	Meaning	Unit
T_e	Equivalent input noise temperature of the amplifier	K
$\overline{T_e}$	Noise temperature averaged over 4 to 6 GHz	K
F	Fano factor of the suppressed shot noise, $F \leq 1$	—
$\overline{\lvert i_{n,d}\rvert^2}$	Channel current noise power spectral density	A^2/Hz
P_{DC}	Static power consumption	mW
$P_{1dB,in}$	Input-referred 1 dB compression point	dB mW
P_{tone}	Power per readout tone at the LNA input	dB mW
G_{min}	Minimum gain over the band	dB
G_0	Small-signal gain in the linearity measurement	dB
N	Number of multiplexed readout tones	—
τ	Time constant of the transient ring-down	s

Terms

Back gate

The second gate of an FDSOI device, located under the buried oxide. Its voltage V_{BG} shifts the threshold voltage and is the only post-fabrication tuning knob of the reference LNA.

Bias tee

The network that superimposes a DC bias on an RF path while keeping the two separated. Ideal bias tees are used for the S-parameter analyses and replaced by realistic values for the transient ones, whose extreme component values otherwise prevent convergence.

Binning

The convention by which an industrial compact model splits the (W, L) plane into ranges and assigns a distinct set of extracted parameters to each. A device whose dimensions fall outside every bin selects no model card and carries exactly zero current, without any simulator warning.

Charge partitioning

The modelling choice that splits the inversion charge between the source and the drain terminals. It is a convention and not a physical observable, so BSIM-IMG and sEKV disagree on $C_{d,fg}$ and $C_{s,fg}$ while agreeing on their sum.

Compact model

An analytical description of a device, expressed as a set of closed-form equations with extracted parameters, fast enough to be evaluated inside a circuit simulator.

Cryo-CMOS

Circuits fabricated in a standard CMOS process and operated at cryogenic temperature, typically at the 4 K stage of a dilution refrigerator.

Device flavour
One device option of the technology, identified by its gate oxide and its threshold voltage adjustment (SG or EG, SLVT, LVT, RVT), in an n and a p version. The Cadence cell names follow this convention, for instance `slvtnfet` and `eguslvtpfet`. Each flavour carries its own extracted parameter set.

Dilution refrigerator
The cryostat that cools a superconducting quantum processor to a few tens of millikelvin, with intermediate stages of increasing temperature and cooling power.

Dispersive readout
The measurement of a qubit state through the frequency shift it imposes on a coupled resonator. It sets the power per tone at the LNA input, of the order of $-100\,\mathrm{dB\,mW}$.

Extrinsic capacitance
A capacitance originating outside the channel: gate–source and gate–drain overlaps, source and drain junctions, gate-to-substrate overlap at the finger ends. It behaves as a parasitic offset added on top of the intrinsic response and scales with W or with L, not with the area.

Fano factor
The suppression coefficient F of the channel shot noise. It is temperature-independent, which allows its extraction at room temperature, and bias-dependent, which is what makes a noise-aware bias optimisation possible.

Figure of merit
The scalar metric of Das et al. [1] combining minimum gain, fractional bandwidth, noise temperature and power, used here to compare configurations after selection and not to select them.

Gear-2 and `traponly`
The two transient integration methods of Spectre. Gear-2 is the default and introduces an artificial numerical damping that can hide a marginal oscillation, which is why the kick test is run with `traponly`.

Inductive source degeneration
An input stage topology in which an inductor in the source synthesises a real input resistance without adding a physical resistor, so the input can be matched to $50\,\Omega$ without the noise penalty of a resistive termination.

Intrinsic capacitance
A capacitance describing the response of the inversion charge inside the channel. It is strongly bias-dependent, proportional to the channel area WL, and is the quantity a compact model such as sEKV predicts.

Inversion coefficient
The normalised current of the EKV formulation, which places the operating point on the continuum from weak to strong inversion and exposes the analog design trade-offs directly.

Kick test
A transient verification in which a current impulse is injected on an internal node and the response is checked to be a damped ring-down. It detects inter-stage oscillations that the μ factor, computed at the external ports only, cannot see.

Model card
The block of a SPICE-like file that lists the extracted parameter values of one device for one bin.

Noise match
The source impedance that minimises the noise figure of an amplifier, in general different from the impedance that maximises power transfer.

Operating point
The DC solution of a circuit, from which the simulator derives the small-signal parameters. Read

here through the `dcOpInfo` structure of Cadence.

Quasi-static
The regime in which the channel charge follows the terminal voltages instantaneously. The AC extraction frequency is kept low enough (1 kHz) for this assumption to hold.

Qubit
The elementary information unit of a quantum processor. The superconducting qubits targeted here are read out in the 4 to 8 GHz band and require the amplification chain of which the LNA is the first stage.

Skillbridge
The Python-to-SKILL bridge used to drive Cadence Virtuoso from an external Python session, which is what makes the geometry sweeps of the R_G extraction automatic.

Strap
The metal rail that connects the gate fingers of a multi-finger device together. Being shared by all the fingers, it is a series element and therefore grows with the finger count instead of benefiting from the $1/n_f$ parallel factor.

Suppressed shot noise
The description of the channel noise of a cryogenic MOSFET as $2qFI_{DS}$ per unit bandwidth, that is, full shot noise reduced by the Fano factor.

Tape-out
The submission of a finished design to fabrication. At cryogenic temperature each iteration costs a fabrication run followed by a cool-down, which is the cost that a validated compact model removes.

Transcapacitance
An off-diagonal element $C_{xy} = -\partial Q_x/\partial V_y$ of the capacitance matrix. The matrix is non-reciprocal: $C_{xy} \neq C_{yx}$ in general.

Unconditional stability
The property that no combination of passive source and load impedances can make the amplifier oscillate, tested here with the Edwards–Sinsky factor $\mu > 1$.

`vdc` **and** `acm`
The DC value and the AC magnitude of a Cadence source. Every terminal of the extraction testbenches carries a `vdc`, and only the excited terminal carries `acm` = 1.

Contents

Acknowledgements

Rhonexum **i**

Glossary **ii**

List of Figures **xiii**

List of Tables **xvi**

1 Introduction **1**

2 Cryogenic PDK generation and integration **3**
2.1 Automated Generation of Transistor Cells for Cadence Virtuoso 3
2.1.1 Motivation 3
2.1.2 Structure of the generator 3
2.1.3 Scope and conclusion 6

3 Cryogenic PDK validation **7**
3.1 Room-temperature DC validation of the short-channel model 7
3.1.1 Device-level comparison 7
3.1.2 Circuit-level comparison: OTA DC gain 8
3.1.3 Summary 10
3.2 AC operating mode: extraction of the intrinsic transcapacitances 10
3.2.1 Context: intrinsic versus extrinsic 10
3.2.2 Extraction method 11
3.2.3 Validation of the extraction 13
3.2.4 Comparison between the GF22FDX PDK and the sEKV model 16
3.2.5 Conclusion 17
3.3 Gate resistance: extraction, calibration and modelling 17
3.3.1 Why the sEKV needs a gate resistance at all 17
3.3.2 The physical model of the gate electrode in the literature 18
3.3.3 Extraction method 19
3.3.4 Calibration 20
3.3.5 Simplification for the sEKV 22
3.3.6 Implementation in the sEKV 24

4 Cryogenic CMOS design and optimization **26**
4.1 Design verification and optimization of a 4–6 GHz cryogenic CMOS LNA using the sEKV compact model 26
4.1.1 Part I: verification of the model 27
4.1.2 Part II: optimization 30
4.1.3 Part III: verification of the optimized circuit 36
4.1.4 Conclusion 38

5 Conclusion and perspectives **39**
5.1 Conclusion 39

A Monte-Carlo analysis of the OTA DC gain **41**

B Supporting figures **42**
B.1 Intrinsic transcapacitances reported in the literature 42
B.2 Stability and time-domain verification of the optimized LNA 43

C Derivation of the ABCD model **44**

D Project schedule **48**

Resumé Français **49**

English Summary **50**

Bibliography **51**

Competency self-assessment report **53**
D.1 Background and self-assessment approach 53
D.2 C1: Designing or implementing engineering solutions to meet a specification 54
D.2.1 Level 1: follow an established approach or protocol 54
D.2.1.1 Context 54
D.2.1.2 Argument 54
D.2.1.3 Resources drawn on 55
D.2.1.4 Evidence 55
D.2.2 Level 2: propose an approach or a protocol 56
D.2.2.1 Context 56
D.2.2.2 Argument 56
D.2.2.3 Resources drawn on 57
D.2.2.4 Evidence 58
D.2.3 Progression radars 58
D.2.4 Final assessment 58
D.3 C2: Implementing a fundamental research approach or an applied research approach for innovation purposes 59
D.3.1 Level 1: analyse and use an innovation or research approach 59
D.3.1.1 Context 59
D.3.1.2 Argument 59
D.3.1.3 Resources drawn on 60
D.3.1.4 Evidence 60
D.3.2 Level 2: carry out a research or innovation process 60
D.3.2.1 Context 60
D.3.2.2 Argument 60
D.3.2.3 Resources drawn on 62
D.3.2.4 Evidence 62
D.3.3 Progression radars 62
D.3.4 Final assessment 62
D.4 C3: Collaborate within a team and in project mode 62
D.4.1 Level 1: participate in a team 62
D.4.1.1 Context 62
D.4.1.2 Argument 62
D.4.1.3 Resources drawn on 64

D.4.2 Level 2: lead or facilitate a project . . . 64
D.4.2.1 Context . . . 64
D.4.2.2 Argument . . . 64
D.4.2.3 Resources drawn on . . . 66
D.4.2.4 Evidence . . . 66
D.4.3 Progression radars . . . 66
D.4.4 Final assessment . . . 66
D.5 C4: Work and communicate in an international and intercultural environment . . . 67
D.5.1 Part A: justify international mobility . . . 67
D.5.1.1 Context . . . 67
D.5.1.2 Comparative analysis of work and training habits . . . 67
D.5.1.3 Reflective argumentation on the competency criteria . . . 68
D.5.1.4 A reservation I state . . . 68
D.5.2 Part B: master one or more foreign languages . . . 68
D.5.2.1 Argument . . . 68
D.5.2.2 Evidence . . . 69
D.5.3 Final assessment . . . 69
D.6 C5: Integrate technological, environmental and societal transitions into the work . . . 69
D.6.1 Level 1: analyse an existing situation . . . 69
D.6.1.1 Context . . . 69
D.6.1.2 Argument . . . 69
D.6.1.3 Evidence . . . 71
D.6.2 Level 2: propose solutions . . . 71
D.6.2.1 Argument . . . 71
D.6.2.2 Resources drawn on . . . 72
D.6.2.3 Evidence . . . 73
D.6.3 Progression radars . . . 73
D.6.4 Final assessment . . . 73
D.7 Summary and overall assessment . . . 73
D.8 Evidence and appendices . . . 74
D.8.1 Master's thesis . . . 74
D.8.2 Impact analysis document from the second-year internship . . . 75
D.8.3 Planned and actual schedule of the final-year project . . . 75
D.8.4 Nanogenerator group project . . . 75
D.8.5 Transcripts and language certification . . . 75
D.8.6 Competency radars . . . 75

List of Figures

2.1 Generation flow. The `.lib` file produced by the parameter extraction is parsed by the Python script, which writes one SKILL file per transistor cell. Executing these files in the Virtuoso CIW populates the Cadence library `rx_model_cryo_PDK_22FDX` with cells having a symbol, a `spectre` view and a complete CDF. 4

3.1 Long-channel SG FDSOI NMOS ($L = 500\,\text{nm}$, $W = 1\,\mu\text{m}$) at room temperature. Left: transfer characteristic $I_D(V_{GS})$ at $V_{DS} = 0.5$ and $0.8\,\text{V}$, logarithmic (left axis) and linear (right axis) scales. Right: output characteristic $I_D(V_{DS})$ for $V_{GS} = 0.1$ to $1.0\,\text{V}$. Blue: sEKV (ngspice). Red: GF22FDX PDK (Cadence). 7

3.2 Short-channel SG FDSOI NMOS at room temperature, transfer characteristic $I_D(V_{GS})$ at $V_{DS} = 0.5$ and $0.8\,\text{V}$. Left: $L = 20\,\text{nm}$, $W = 1\,\mu\text{m}$. Right: $L = 40\,\text{nm}$, $W = 1\,\mu\text{m}$. Blue: sEKV (ngspice). Red: GF22FDX PDK (Cadence). 8

3.3 Schematic of the five-transistor OTA used for the circuit-level DC comparison, as entered in Cadence Virtuoso. 9

3.4 Small-signal equivalent circuit of the EKV model. The dashed box delimits the intrinsic elements; all the other capacitances are extrinsic contributions located outside the channel. 11

3.5 Cadence implementation of the extraction testbench: three `slvtnfet` devices T_1, T_2, T_3 sharing the same bias conditions and the same AC excitation on the front gate, and differing only by their W and L. 13

3.6 Normalised intrinsic transcapacitances (C/WLC_{ox}) extracted by the three-transistor testbench in Cadence (red) compared with the intrinsic ngspice reference (blue), for the sEKV model with extrinsic capacitances enabled in the verilog-A. 14

3.7 Intrinsic transcapacitances of the GF22FDX PDK at $V_{DS} = 0\,\text{V}$ ($L = 500\,\text{nm}$, $W = 1\,\mu\text{m}$): testbench extraction (dashed) versus the operating-point variables of the PDK (solid). 14

3.8 Same comparison as Fig. 3.7 but at $V_{DS} = 0.5\,\text{V}$. $C_{fg,fg}$ and $C_{bg,fg}$ still match, whereas $C_{d,fg}$ and $C_{s,fg}$ deviate increasingly with V_{FG}. 15

3.9 Intrinsic transcapacitances of the GF22FDX PDK (solid, from the operating-point variables) and of the sEKV model (dashed, extracted with the three-transistor testbench), for an SG NMOS $L = 500\,\text{nm}$, $W = 1\,\mu\text{m}$. 16

3.10 Front-gate transcapacitances normalised to $C_{ox1}WL$, GF22FDX PDK (Cadence) versus sEKV (ngspice), for $V_{DS} = 0\,\text{V}$ (top) and $V_{DS} = 0.5\,\text{V}$ (bottom), on SG $L = 500\,\text{nm}$, $W = 1\,\mu\text{m}$ devices. 17

3.11 R_G extraction testbench: front gate on port 1, drain on port 2, source and back-gate grounded, device biased in the linear region. 20

3.12 Calibrated model (solid) against the extracted points over the three geometry sweeps. (a) The minimum near $1\,\mu\text{m}$ is the crossover between the vertical access, dominant at small W, and the horizontal electrode, dominant at large W. (b) The $1/L$ decay of the electrode term over almost two decades. (c) The drop is the $1/N_F$ parallel factor; the upturn beyond $n_f \approx 10$ is the strap, which is a series element and therefore grows with the finger count. 22

3.13 Contribution of each term of (3.10) at $L = 70\,\text{nm}$, $n_f = 1$. The horizontal electrode dominates, the vertical access carries the small-W rise, the contacts are a nearly constant offset and the strap stays below $0.1\,\Omega$. 22

3.14 AC testbench: one `slvtnfet` with the added gate resistor `R0` (`r=res`) in series with the front gate. 23
3.15 Relative deviation of the extracted transcapacitances versus added gate resistance, referenced to the lowest-R_{gate} point, $L = 250\,\text{nm}$. 24

4.1 Schematic of the three-stage cryogenic LNA, reproduced from [1], with the component values published in the article. All transistors are minimum length. 27
4.2 Fano factor versus drain current density: digitized measurement points from [27] and the linear regression of equation (4.3) implemented in the Verilog-A model. 28
4.3 Simulated drain current density versus gate voltage for the three device geometries. The markers indicate the gate voltages that reproduce the current densities of the article. 28
4.4 S-parameters of the circuit simulated in Cadence with the sEKV model, compared with the curves of [1] digitized by hand. 29
4.5 Analytical ABCD model evaluated with the small-signal parameters of Table I of [1], compared with the measured and simulated curves of the article. 31
4.6 S-parameters and noise of the simulated configurations compared with the measurement of [1]. 33
4.7 Bias grids (V_{G2}, V_{G3}) explored at the two supply voltages, at fixed V_{G1}. Discs satisfy the three specifications and are unconditionally stable, colored by the power consumption; open triangles satisfy the specifications but not the μ condition; crosses violate at least one specification. The circle marks the retained point. Note the different axis ranges of the two panels. 34
4.8 S-parameters and noise temperature of the simulated configurations compared with the measurement of [1]. The legend applies to the four panels. 35
4.9 Gain compression at 5 GHz extracted from transient simulation with coherent FFT, at the two optimized bias points. The dashed lines mark $G_0 - 1$ dB. 37

A.1 Left: Monte-Carlo distribution of the OTA DC gain simulated with the GF22FDX PDK at room temperature (1000 samples, $\mu = 14.26$ V/V, $\sigma = 3.48$ V/V). Right: the three DC gain values of Table 3.1 placed on the $\pm1\sigma$, $\pm2\sigma$ and $\pm3\sigma$ bands of that distribution. . 41

B.1 Intrinsic transcapacitances $C_{fg,fg}$, $C_{d,fg}$, $C_{s,fg}$ and $C_{bg,fg}$ normalised to $C_{ox1}WL$ versus front-gate voltage, from Lu et al. [16] ($T_{ox1} = 1.2\,\text{nm}$, $T_{si} = 15\,\text{nm}$, $T_{ox2} = 20\,\text{nm}$, $V_{ds} = 0.5\,\text{V}$, $V_{bg} = 0$). The curves are obtained from TCAD on an idealised device and are used in Section 3.2.1 only to establish the expected shape and sign of each transcapacitance. 42
B.2 Edwards–Sinsky stability factors μ and μ' of the optimized circuit at both supply voltages, computed from the S-parameters of Section 4.1.2 with the inductor series resistances included. The dashed line marks the unconditional stability limit $\mu = 1$. Both factors stay above unity over the whole sweep, the minima being $\mu'_{\text{min}} = 1.0042$ at $V_{DD} = 0.6$ V and 1.0012 at $V_{DD} = 0.3$ V. 43
B.3 Kick test: response at the output node to a current impulse injected on the drain of M1, for two impulse amplitudes. Left: raw response at $V_{DD} = 0.6$ V. Right: the two responses normalised by the impulse amplitude, which superpose and confirm that the circuit answers linearly to the perturbation. 43

D.1 Project schedule 48

D.2 Competency C1: radar filled in during the second-year internship, on the left, and radar updated after the final-year project, on the right. The dashed grey outline on the right recalls the level reached in the second year. 58
D.3 Competency C2: radar of the second year and radar updated after the final-year project. 63
D.4 Competency C3: radar of the second year and radar updated after the final-year project. 66
D.5 Competency C5: radar of the second year and radar updated after the final-year project. 73

List of Tables

3.1 DC gain of the OTA at room temperature. 9
3.2 Validity of the three-transistor extraction on the GF22FDX PDK, assessed against the operating-point variables of the PDK. 15
3.3 Symbols of the general expression (3.10). Another device flavour (other V_t, the PFET) reuses the same structure with its own coefficients. 19
3.4 Published reference values for the coefficients of (3.10). The spread on the vertical term reflects the gate stack: the older $TiSi_2$-on-polysilicon stacks are an order of magnitude worse than modern NiSi and metal-gate stacks. 19
3.5 L sweep at $W = 1\,\mu m$, $n_f = 1$. The residual is flat versus L ($\approx 35\,\Omega$), not the $1/L$ growth a vertical term would give nor the growth with L an NQS term would give. Read on the n_f sweep at $W_{tot} = 8\,\mu m$ the same residual is only about $4.5\,\Omega$, again flat, hence the $\approx 36\,\Omega\mu m/W_{tot}$ law. 21
3.6 Calibrated coefficient set (RMS 2.2 % over the three sweeps). The purely geometric quantities (ΔL, ΔW, w_c, p_c, s_{pp}, the strap cross-section) are held at their nominal values and not fitted. 21
3.7 Deviation of S1 and S2 against the calibrated expression. The large relative errors at high n_f sit on a near-zero base of a few ohms and reflect the missing strap, not a meaningful absolute error. 23
3.8 Resistance error crossed with the measured AC sensitivity, $L = 250\,nm$. $\Delta R_g = \text{full} - S_i$; the AC deviation is read from Fig. 3.15 at that added resistance on the matching width curve. 24

4.1 Small-signal parameters: measurement of [1] (Table I) versus sEKV model, at $J_D =$ 37.5 mA/mm. 28
4.2 Based on the paper's schematic: comparison with Das *et al.* [1] over 4–6 GHz. 29
4.3 Sensitivity of the main metrics to a 15 % variation of each passive element. Only the elements with a significant effect are listed; the gate bias resistors and the supply filtering elements were found to be inert on all six metrics. 31
4.4 Effect of the passive recalibration over 4–6 GHz, with the measurement of [1] as reference. 33
4.5 Recalibrated post-layout + bias optimized for FOM, compared with Das *et al.* [1] at both supply voltages. The reference is the nominal bias of the article at $V_{DD} = 0.6$ V and its published low-power bias at $V_{DD} = 0.3$ V. All band metrics are taken over 4–6 GHz. 35
4.6 Figure of merit of all the configurations, computed with equation (4.13) and the $G_{\min}$ convention used throughout this work. With the gain convention of [1], the published values for the two bias points of the article are 0.192 and 0.517. 36

C.1 Mapping between the `dcOpInfo` record and the parameters of Eq. (C.4). Off-diagonal capacitances are taken as positive magnitudes, $C_{gd} = -\partial Q_G/\partial V_D$ and $C_{dg} = -\partial Q_D/\partial V_G$. 45
C.2 Decomposition of the LNA into elementary two-ports. Y_{Gk} is the gate bias conductance, $Y_{Lk} = 1/Z_{Lk}$ the drain load admittance. Numerical values are those of the netlist; L_{s1} is the degeneration inductor of M_1. 46

D.1 Summary of the self-assessment against the third-year targets. 74

1 Introduction

Superconducting quantum processors operate at a few tens of millikelvin, while the electronics that generates their control pulses and reads their state sits at room temperature. Each qubit therefore requires its own line running down the cryostat. This works for the few tens of qubits of current processors but does not extend to the thousands of physical qubits required by fault-tolerant architectures: the cabling occupies the available volume, carries a thermal load at every stage, and limits the number of channels a dilution refrigerator can host [2]. The approach studied to remove this limit is to bring the control and readout electronics inside the cryostat, typically at the 4 K stage where the available cooling power is of the order of one watt. Circuits fabricated in a standard CMOS process and operated at that temperature are known as cryo-CMOS [3], and their interest is integration: a die placed close to the quantum processor replaces a large number of cables, and the analog front end can be multiplexed instead of duplicated. The first block of the readout chain is a low-noise amplifier [4], which must deliver more than 35 dB of gain with a noise temperature of a few kelvins over the 4 to 8 GHz band, within the cooling budget of the stage it sits on. This function is currently implemented in InP HEMT technology, and the trade-off between noise temperature and dissipated power that governs this class of amplifier is reviewed in Bardin [5]; a CMOS implementation would be compatible with the rest of an integrated front end, and reaching the required noise and power figures in CMOS remains an open problem [6].

At room temperature, an analog designer works inside a closed loop: a schematic is entered in an EDA environment, simulated with the compact models of the process design kit (PDK), and the predicted performance is trusted well enough to commit to a fabrication run. At cryogenic temperature this loop is broken. The models shipped with an industrial PDK are calibrated over a temperature range that stops far above 4 K, and nothing guarantees their extrapolation below it, while the device physics itself changes: the threshold voltage increases, the subthreshold slope stops following kT/q and saturates, and the mobility is modified [7, 8, 9]. The performance of a cryogenic circuit therefore cannot be verified before tape-out, and each design iteration costs a fabrication run followed by a cool-down. Closing this gap requires a compact model that is valid at cryogenic temperature, usable in a SPICE-class simulator, and that a designer can place on a schematic in the same way as a PDK device.

This project was carried out in the AQUA laboratory at EPFL, within Rhonexum, an incubated startup project of the laboratory dedicated to cryogenic electronics. Rhonexum develops a modelling service called RX-Model, which covers transistor compact modelling and parameter extraction at cryogenic temperature. Its compact model, referred to as sEKV in this report, is based on the simplified EKV charge-based MOSFET model [10], a design-oriented formulation that exposes the trade-offs of the analog design octagon directly to the designer through the inversion coefficient [11, 12]. The target technology is GlobalFoundries 22FDX (22 nm FDSOI), and an important point for the rest of this report is that sEKV is not a PDK built from scratch: the original GF22FDX kit is kept for the passive devices and for the whole design infrastructure, and only the default transistor compact model is replaced by the cryogenic sEKV implementation written in Verilog-A. The parameter extraction is performed separately in the laboratory, flavour by flavour and geometry by geometry, and produces a SPICE-like `.lib` file that contains the model cards and the binning rules. This report concerns the use of that file and not its production: the AC description of the model and

the cryogenic extraction flow that feeds it were developed in parallel by colleagues of the laboratory. What was missing is what this work addresses. The `.lib` file was not yet a design environment, its predictions had not been checked against an independent reference, and no circuit had been designed with it. These three gaps define the three parts of this work and follow the three steps of the project description.

Part 2 builds the design environment. A Python tool, `gen_skillfinit.py`, takes the `.lib` file and derives from it everything Cadence needs: a SKILL script that builds the library, the cell, its symbol and its CDF, and a Spectre netlist that declares the Verilog-A module and reproduces the binning logic. Both files come from the same parse of the same input, so the consistency between what the schematic knows about a device and what the simulator knows about it is obtained by construction rather than by review. The purpose is to make the creation of a cryogenic library from an extraction output an automated step rather than a manual one.

Part 3 checks the model against an independent reference at device level, in DC and in AC, using the GF22FDX PDK at room temperature. Room temperature is the only operating point where such a reference exists on both sides of the comparison, which is what makes the exercise conclusive. The DC characteristics are compared on long- and short-channel devices and, at circuit level, on a published OTA of the same technology. A method is then developed to extract the intrinsic transcapacitances from standard AC simulations without any access to the internals of the compact model, relying on the different geometrical scaling of the intrinsic and extrinsic charges. Finally, a physical model of the gate resistance is taken from the literature, calibrated on the device with an automated extraction flow, and simplified for implementation in the Verilog-A description, the simplification being decided from a measured sensitivity study rather than from a tolerance argument.

Part 4 puts the model to work on a circuit. A published 4 to 6 GHz cryogenic CMOS LNA [1] is reproduced and simulated at 4.4 K, after a channel noise model with a bias-dependent Fano factor is added to the Verilog-A description. The model is first verified at transistor level and on the complete published circuit, and each deviation is traced to an identified mechanism. An analytical model of the amplifier is then developed to obtain the sensitivity of every metric to every passive element, which guides a recalibration of the passive network, and the bias of the three stages, which the original methodology could not explore because it relied on a small-signal model extracted at a single operating point, is optimized on a constrained grid. The stability and the linearity of the resulting circuit are verified. A general conclusion summarises what the model can and cannot be used for at this stage, and lists the extraction work it calls for.

2 Cryogenic PDK generation and integration

2.1 Automated Generation of Transistor Cells for Cadence Virtuoso

2.1.1 Motivation

As stated in the introduction, the sEKV replaces the transistor compact model of the GF22FDX kit and nothing else. The Verilog-A implementation is attached to the simulator through a Spectre model library, so cryogenic circuit simulation runs inside the standard Cadence design flow.

The sEKV parameters are extracted separately in the laboratory, flavour by flavour (`sg`, `slvt`, `lvt`, `rvt` and the other device options of the 22FDX) and geometry by geometry within each flavour. The result is a SPICE-like `.lib` file carrying the model cards and the binning rules that select the right card for a given W and L, following the geometrical binning convention that industrial compact models have used since Sheu et al. [13] and that the reference model of this technology, BSIM-IMG [14], still applies. It holds all the physics, but it describes model cards and not devices that can be placed on a schematic. To be usable, the model has to be presented to Cadence exactly like a PDK device, through three objects that must agree with each other: a **symbol** view, so that the device can be instantiated and wired; a **CDF** (Component Description Format) description, which declares the parameters the designer can edit (w, l, n_f, m), their default values, the terminal order and the netlisting rules; and a **Spectre netlist**, which binds the instance to the Verilog-A module and to the correct parameter set.

Building these by hand is feasible for one device and impractical for a complete kit. The operation is repeated for every device option of the 22FDX, in n and p versions, each with its own symbol, CDF, binning table and model cards, and each card carries enough numerical parameters for at least one transcription error to be likely. The extraction is also an ongoing activity: every run produces an updated `.lib`, with new parameter values and sometimes modified binning ranges, so the objects bound to Cadence have to be rebuilt as easily as the `.lib` is reproduced. And an inconsistency between the three objects produces no error message: a terminal order declared in the CDF that does not match the port order of the `.subckt`, or a parameter name differing by one character between the schematic and the netlist, still converges and gives a wrong result. The `.lib` file is therefore taken as the single source of truth, and everything Cadence needs is derived from it automatically by the Python tool `gen_skillfinit.py`.

2.1.2 Structure of the generator

For each subcircuit found in the `.lib`, the tool writes two files (Fig. 2.1): a SKILL script (`.il`), SKILL being the native scripting language of Virtuoso, which builds the library, the cell, its symbol and its CDF when it is executed; and a Spectre netlist (`<subckt>_cryo.scs`), which declares the Verilog-A module, reproduces the binning logic and carries the model cards. The split follows the division of responsibility inside Cadence: the SKILL script describes what the schematic knows about the device, the Spectre netlist what the simulator knows about it. Both are written from the same parse of the same input, so the parameter names, the terminal order and the model names are written once and used twice, which removes the silent inconsistency described above.

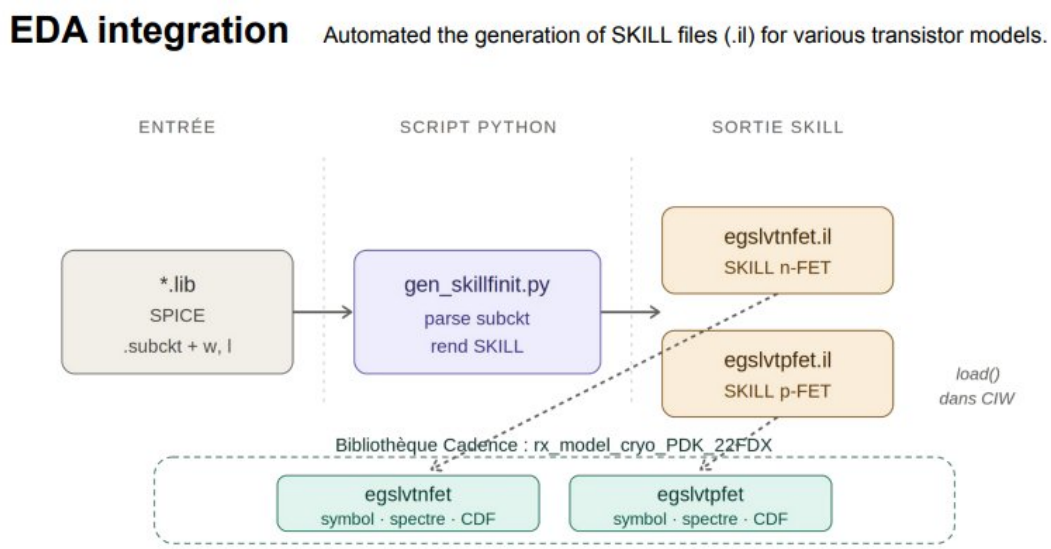


Figure 2.1: Generation flow. The `.lib` file produced by the parameter extraction is parsed by the Python script, which writes one SKILL file per transistor cell. Executing these files in the Virtuoso CIW populates the Cadence library `rx_model_cryo_PDK_22FDX` with cells having a symbol, a `spectre` view and a complete CDF.

The two files also change at different rates. The schematic side is stable, because the symbol, the terminal order and the list of editable parameters do not depend on the numerical content of the model; a cell already loaded in Cadence is rebuilt only when its identity changes, for instance for a new flavour or when the library is renamed or moved. The simulator side follows the extraction work, the `.scs` being regenerated after each new extraction so that the model cards match the new parameter set. Updating the environment after an extraction therefore costs one file and not the whole library.

Parsing of the `.lib` file. The parser reads only what it needs: the library name on the `.lib` line; the cell names, port lists and default dimensions on the `.subckt` lines; and, when a netlist has to be produced, the binning blocks and the `.model` cards with all their continuation lines. Comments, Verilog-A includes, temperature settings and corner sections are skipped, since they describe how a simulation is configured and not how a cell is described.

Three choices make the parser tolerant to the way the input file is written. Tokens on a `.subckt` line are matched by pattern and not by position, so the `w=` and `l=` assignments can appear anywhere relative to the port list and a file written by another extraction script is accepted without editing. Default dimensions are stored as strings and never converted to floating-point numbers, which keeps the SPICE suffixes (`1u`, `30n`) as written and avoids the rounding of a numerical conversion. Incomplete inputs are also accepted: a subcircuit with no binning and no model card still produces its `.il` file and no `.scs`, which allows a symbol to be created before the corresponding extraction is finished.

Two further choices are approximations. The defaults written in the CDF are not the literal `w` and `l` of the `.subckt` line but the midpoints of the W and L ranges of a chosen bin; a device whose dimensions fall outside every bin selects no model and carries no current (see below), so taking the defaults inside a bin guarantees that a newly placed instance can be simulated immediately. And since the `.lib` file does not say whether a subcircuit is an NMOS or a PMOS, while the symbol needs its arrow in the right direction, the type is deduced from the cell name: a name containing `pfet` or `pmos`, or ending with `p`, is treated as PMOS. This works with the naming convention of the models

used here, but it remains a weak point of the tool.

Generation of the cell. Executing the `.il` file creates the library if needed and deletes any cell of the same name, so that running it again after a new extraction overwrites the old cell instead of leaving it partly updated. The symbol is not drawn directly: a `schematic` view with four `inputOutput` pins (D, G, S, B) is created first and a `schematicSymbol` view is generated from it, which guarantees that the pins of the symbol are the pins Cadence associates with the cell. The graphics produced by that conversion are then removed and the symbol is redrawn, with the pins repositioned, the transistor body drawn segment by segment, the source arrow oriented according to the channel type and the three label types placed (`cdsTerm` for the terminals, `cdsParam` for the parameters, `cdsName` for the instance name). Redrawing rather than keeping the automatic symbol gives the same appearance and the same pin placement for every flavour, whatever the port order of the input file. The symbol is finally copied into a `spectre` view and a `symbol` view and the intermediate schematic view is deleted; both are needed, since the netlister looks for the `spectre` view when it resolves the instance towards the simulator and the schematic editor looks for the `symbol` view when the cell is placed.

The CDF gives the cell its electrical meaning. Four parameters are editable (w, l, n_f, m) and a fifth, `model`, is read-only and holds the original subcircuit name; it is the link between the schematic instance and the Spectre netlist, and making it read-only prevents a designer from breaking that link by accident. The `simInfo` table states how the instance is netlisted:

```
otherParameters = (model)     ; subcircuit to instantiate
instParameters  = (w l nf m)  ; parameters passed to that instance
termOrder       = (D G S B)   ; symbol pins <-> subckt ports
```

The `termOrder` entry and the associated `termMapping` fix the correspondence between the pins of the symbol and the ports of the `.subckt`. An error here gives no warning: it produces a transistor with drain and source exchanged, which simulates normally and gives a wrong result. Since the symbol pins and the netlist ports come from the same parsed port list, this error cannot occur.

Generation of the netlist. The second file is a Spectre subcircuit built from the parsed `.lib`: an `ahdl_include` of `sekv.va`, an `inline subckt` whose ports are those read in the input, a `parameters` line with the default dimensions and the instance parameters n_f and m, the binning logic, and the model cards with all their parameters.

The binning is translated from the SPICE `.if` / `.elseif` form to the Spectre `if` / `else if` form, each branch instantiating the same device with the model card valid over its geometry range. The selection uses the width per finger w/n_f and not the total width, which is the correct criterion since the cards are extracted per finger: a device declared with $w = 2\,\mu\text{m}$ and $n_f = 4$ must select the bin of 500 nm. As n_f is available in the CDF, changing the number of fingers on the schematic selects the right card automatically.

```
if ( (w/nf >= WMIN_1) && (w/nf < WMAX_1) && (l >= LMIN_1) && (l < LMAX_1) ) {
    n1 (d g s b) model_bin1 w=w l=l nf=nf m=m
} else if ( ... ) { ... }
```

The branches keep the order of the `.lib`, so the first-match behaviour of the original file is preserved and the generator does not introduce an order of its own. The multiplier m is passed to the instantiated device but is not used in the selection, since it replicates a device without changing the geometry for which a card was extracted.

One consequence deserves attention, because the symptom is misleading: an instance given dimensions outside every range takes no branch, instantiates no device inside the subcircuit and carries no current, while the simulation converges without any warning and reports a drain current

of exactly zero. This is why the CDF defaults are taken inside a bin (see above), and why a device that unexpectedly carries no current should have its geometry compared with the binning table before the bias conditions or the model itself are questioned.

Two syntax rules of Spectre that SPICE model files do not always follow are applied while writing: an identifier starting with a digit is prefixed, and a dot inside a model name is replaced by an underscore. Both substitutions are applied to the definition and to every reference in the file, so the netlist stays consistent whatever naming convention the extraction used.

2.1.3 Scope and conclusion

The tool deals with integration and not with modelling: it does not inspect or validate `sekv.va`, which is only referenced by its path, and it does not check that the model cards are physically meaningful, so an unrealistic parameter is copied faithfully and the cell simulates accordingly. Checking that the model gives the correct electrical behaviour is the subject of the rest of this report.

Three limitations remain. The channel type is deduced from the cell name instead of being read in the file. The namespace used for net and terminal names is not available as a settable CDF attribute in the version of Virtuoso used here, which was checked by listing the available attributes with `baseCDF~>?`, and it is the only step of the flow that stays manual, once per cell. Finally, a subcircuit without binning or model cards gives a symbol with no model behind it, so the presence of a cell in the library does not prove that it can be simulated.

Within these limits, the generator turns the file produced by the parameter extraction into a Cadence library whose cells have a correct symbol, a consistent CDF and a Spectre netlist reproducing the original binning, the consistency between the schematic side and the simulator side being obtained by construction and not by review. For the designer, `sEKV` then behaves like the PDK it replaces: the device is placed on a schematic, its property form shows w, l, n_f and m as editable fields and `model` as read-only, and the simulation runs. All the characterisation work of the following sections and the circuit simulations of the verification part were done with cells produced by this flow.

3 Cryogenic PDK validation

3.1 Room-temperature DC validation of the short-channel model

The sEKV parameters are extracted by machine learning from measurements taken at cryogenic and at room temperature. My task was to verify the model on an SG transistor in DC, which requires a reference. Room temperature is the only operating point where such a reference exists on both sides of the comparison: the foundry PDK is guaranteed valid at this temperature, and Rhonexum holds silicon measurement data on the same technology. This section therefore compares the sEKV implementation with the GF22FDX PDK at room temperature, first on a single transistor and then on a complete circuit.

3.1.1 Device-level comparison

The device under test is an SG FDSOI NMOS. Three geometries are compared: a long-channel device ($L = 500\,\text{nm}$, $W = 1\,\mu\text{m}$), used as a reference case where short-channel effects are essentially absent, and two short-channel devices ($L = 20\,\text{nm}$ and $L = 40\,\text{nm}$, both $W = 1\,\mu\text{m}$), which correspond to the regime actually used in design.

Two sweeps are performed. The transfer characteristic $I_D(V_{GS})$ is swept from 0 V to 1 V at $V_{DS} = 0.5\,\text{V}$ and 0.8 V, on a logarithmic and on a linear axis simultaneously so that weak and strong inversion are read on the same figure. The output characteristic $I_D(V_{DS})$ is swept up to 0.9 V for $V_{GS} = 0.1$, 0.4, 0.6, 0.8 and 1.0 V. The GF22FDX curves come from Cadence Spectre and the sEKV curves from ngspice, since a Cadence licence covering the Verilog-A model was not yet available at the time of this study. Both simulations run the same model.

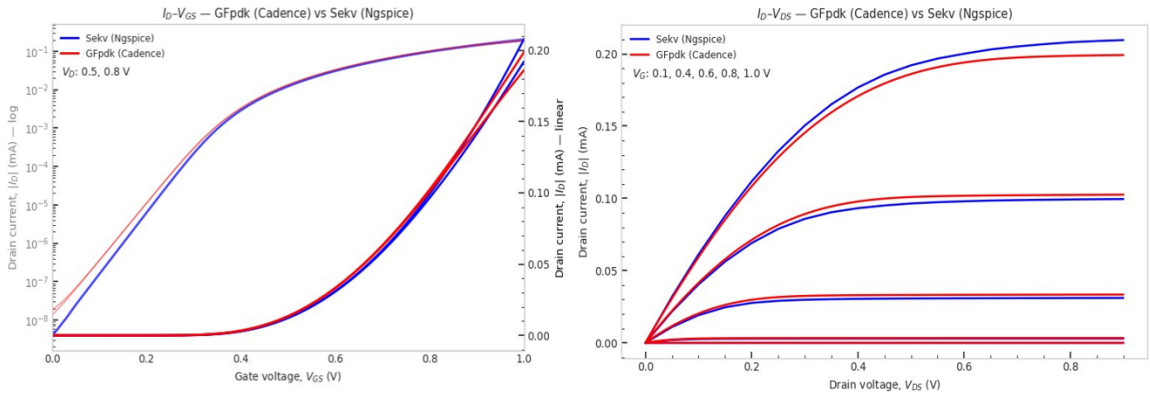


Figure 3.1: Long-channel SG FDSOI NMOS ($L = 500\,\text{nm}$, $W = 1\,\mu\text{m}$) at room temperature. Left: transfer characteristic $I_D(V_{GS})$ at $V_{DS} = 0.5$ and 0.8 V, logarithmic (left axis) and linear (right axis) scales. Right: output characteristic $I_D(V_{DS})$ for $V_{GS} = 0.1$ to 1.0 V. Blue: sEKV (ngspice). Red: GF22FDX PDK (Cadence).

For the long-channel device the two models overlap over the whole sweep. The subthreshold branches follow the same slope over several decades of current, the strong-inversion branches are almost indistinguishable up to $V_{GS} = 1\,\text{V}$, the saturation current is reproduced for every gate bias within a few percent, and the linear-to-saturation transition occurs at the same drain voltage. The charge and transport equations are therefore correctly implemented and correctly parameterised in the regime where the geometry imposes no additional physics.

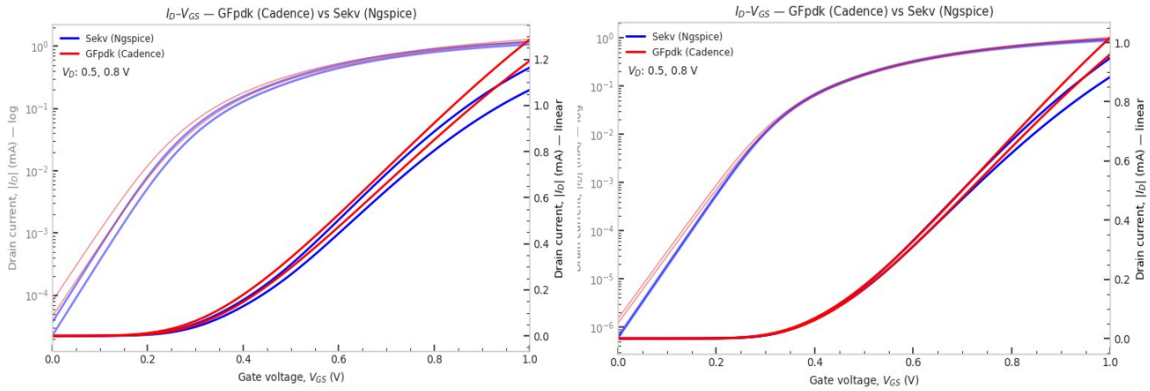


Figure 3.2: Short-channel SG FDSOI NMOS at room temperature, transfer characteristic $I_D(V_{GS})$ at $V_{DS} = 0.5$ and 0.8 V. Left: $L = 20\,\text{nm}$, $W = 1\,\mu\text{m}$. Right: $L = 40\,\text{nm}$, $W = 1\,\mu\text{m}$. Blue: sEKV (ngspice). Red: GF22FDX PDK (Cadence).

For the short-channel devices the agreement remains good but is no longer perfect. In strong inversion, which is the regime that sets the bias point of an analogue stage, the linear branches track each other closely for both lengths. Two deviations remain, and they do not carry the same weight.

The first is the on-current at $V_{GS} = 1\,\text{V}$, where the model sits slightly below the PDK. This point lies outside the operating range of the device. The SG flavour of the GF22FDX has a nominal supply of 0.8 V, set by the gate oxide, and is not operated above it; the circuits designed in this work run at 0.6 V and 0.3 V. A deviation appearing only beyond 0.8 V therefore has no consequence for the designs targeted here.

The second is the moderate-inversion region of the $L = 20\,\text{nm}$ device, where the two logarithmic branches separate by a visible but bounded amount before merging again in strong inversion. This one sits inside the operating range and cannot be dismissed on the same grounds. It is accepted at this stage for a different reason. The sEKV is a design-oriented compact model built on a deliberately reduced parameter set: it aims at capturing the physical trends that drive analogue design trade-offs, not at reproducing an industrial compact model bit for bit. The GF22FDX model is a full production model in which a large number of second-order effects are described by dedicated equations and by a much larger set of fitted parameters. Several of these refinements are not yet present in the current sEKV implementation, and the extraction flow itself is still being consolidated. Measured against that context, an agreement of a few percent in strong inversion and a bounded discrepancy in moderate inversion is a satisfactory result.

3.1.2 Circuit-level comparison: OTA DC gain

Agreement at device level does not guarantee agreement at circuit level, where individual small-signal errors combine and can be amplified. The second test therefore moves to a complete circuit. The reference chosen is the OTA published by Martinez *et al.* [15], designed and measured in the same GF22FDX 22 nm FDSOI technology used throughout this work. Using a published and measured circuit adds a third, independent term to the comparison: silicon.

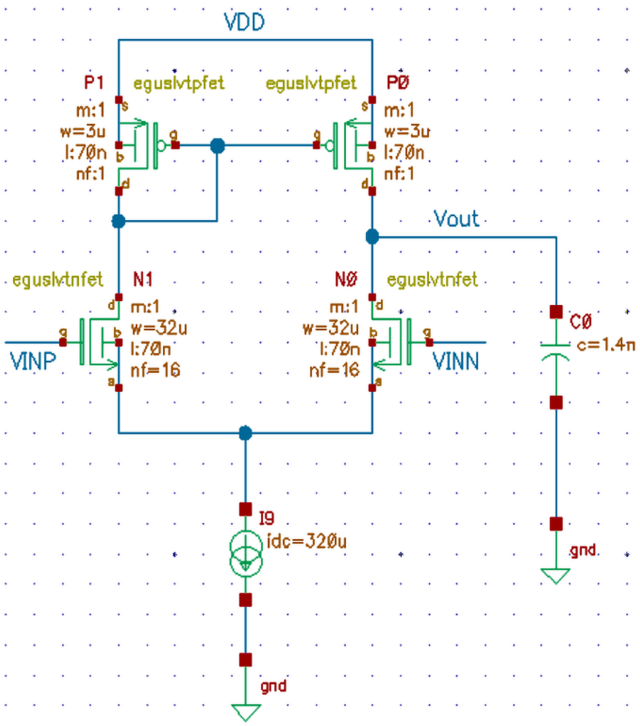


Figure 3.3: Schematic of the five-transistor OTA used for the circuit-level DC comparison, as entered in Cadence Virtuoso.

The circuit, shown in Fig. 3.3, is a five-transistor OTA: an NMOS differential input pair (`eguslvtnfet`, $W = 32\,\mu\text{m}$, $L = 70\,\text{nm}$, $n_f = 16$), a PMOS current-mirror load (`eguslvtpfet`, $W = 3\,\mu\text{m}$, $L = 70\,\text{nm}$), an ideal tail current source of 320 µA and a 1.4 nF capacitive load at the output. The schematic, the bias conditions and the testbench are strictly identical in the two simulations; only the transistor model is swapped between the GF22FDX PDK and the sEKV. The DC gain is read on the low-frequency plateau of the small-signal response, where the gain is flat and equal to its DC value. All simulations are performed at room temperature.

Table 3.1: DC gain of the OTA at room temperature.

Source	DC gain (V/V)	Deviation vs. PDK nominal
GF22FDX PDK (nominal)	14.56	reference
sEKV (sEKV)	9.84	−32%
CEA-Leti measurement [15]	10.54	−28%

Table 3.1 gives the three values. Two observations follow. First, the sEKV underestimates the gain with respect to the PDK nominal simulation by about 32 %. Second, and more significantly, the measured silicon deviates from that same nominal value in the same direction and by a comparable amount, 28 %. The sEKV prediction is within 7 % of the measured gain, that is, closer to the measured device than the foundry nominal simulation is.

This raises the question of whether the gap to the PDK nominal value reflects a modelling error or simply the normal spread of the technology. A Monte-Carlo analysis of the same OTA was run with the GF22FDX PDK alone, at room temperature, over 1000 samples. The resulting distribution of the DC gain is Gaussian, with mean $\mu = 14.26\,\text{V/V}$ and standard deviation $\sigma = 3.48\,\text{V/V}$, i.e. a relative spread of about 24 %. Placed on that distribution, the sEKV prediction lies 1.27σ below the mean and the CEA-Leti measurement 1.07σ below it. Both fall inside the $\pm 2\sigma$ interval and on the same side of the mean at nearly the same distance.

The spread produced by process variability alone is therefore of the same order as the gap between the sEKV and the PDK nominal value, so a single deterministic simulation at the typical corner is not a sharp enough reference to declare the model wrong. What the Monte-Carlo result establishes is that the residual error is contained within the envelope of the technology's own variability, which is the practically relevant criterion for a design-oriented model. It does not prove that the whole gap originates from variability: the DC gain of this topology is set by $g_m/(g_{ds,n}+g_{ds,p})$, a moderate error on g_{ds} alone is enough to shift it by tens of percent, and refining the output conductance description remains one of the open items of the model. The value of Martinez *et al.* is also measured on a fabricated die and therefore contains the layout parasitics, whereas the two simulations are pre-layout, so the 7 % agreement with silicon should be read as an order of magnitude rather than as a precise figure. The distribution itself, the position of the three values on it and the detailed discussion of this pre-layout bias are given in Appendix A.

3.1.3 Summary

The DC core of the model is consistent enough to be used as a design tool. It follows the GF22FDX closely at long channel, deviates in a bounded way at 20 and 40 nm, and places the OTA gain 1.27σ below the mean of the PDK process distribution, inside the spread that the foundry itself accepts. This is the prerequisite for extending the comparison to the AC regime and to cryogenic temperature.

3.2 AC operating mode: extraction of the intrinsic transcapacitances

3.2.1 Context: intrinsic versus extrinsic

The DC validation of the sEKV model only probes the static behaviour of the transistor. The AC operating mode has to be validated as well, because the dynamic performance of any analog block (gain–bandwidth product, dominant pole, settling) is set by the terminal capacitances and not by the currents alone.

In a MOSFET these capacitances belong to two families of very different nature:

- the **intrinsic** capacitances, which describe the response of the inversion charge inside the channel. They are strongly bias-dependent and are the quantities predicted by a compact model such as sEKV;
- the **extrinsic** capacitances, which originate outside the channel (gate–source and gate–drain overlaps, source/drain junctions, gate-to-substrate overlap at the ends of the gate finger). They behave as a parasitic offset added on top of the intrinsic response.

Figure 3.4 shows the small-signal equivalent circuit of the EKV model. The dashed box delimits the intrinsic elements (c_{gsi}, c_{gdi}, c_{gbi}, c_{dbi}); the elements drawn outside it are the extrinsic ones (c_{gsov}, c_{gdov}, c_{gbov}, c_{sbj}, c_{dbj}) together with the access resistances R_{Seff} and R_{Deff}.

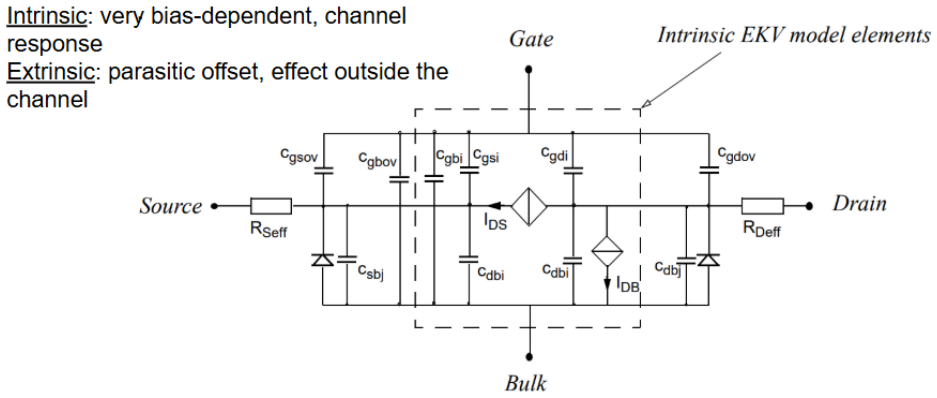


Figure 3.4: Small-signal equivalent circuit of the EKV model. The dashed box delimits the intrinsic elements; all the other capacitances are extrinsic contributions located outside the channel.

A circuit simulator only gives access to the *total* terminal capacitance, that is the sum of the two contributions. Comparing the sEKV model with the GF22FDX PDK at AC level therefore requires a method able to isolate the intrinsic part from a standard Cadence simulation.

Which reference? Two candidate references exist, and neither of them can be used directly.

The first one is the literature. Lu et al. [16] report, for an FD-SOI device, the intrinsic transcapacitances normalised to $C_{ox1}WL$ as a function of the front-gate voltage, obtained from TCAD and reproduced in Fig. B.1 of Appendix B. These curves describe an *idealised* transistor: no routing parasitics, direct access to the intrinsic region and a simplified 2D/3D interconnection scheme. They give the expected shape of the curves, but they do not correspond to the technology target of this work.

The second element is the GF22FDX software development kit (PDK) itself. It targets the exact same technology as the model developed here and includes all the parasitic effects of a realistic device but its BSIM-IMG model map is encrypted, so the internal partition between intrinsic and extrinsic contributions is not accessible from the netlist. The only visible trace is the list of DcOp operating points printed by Cadence (`cggi`,`cgg` `cgsi`, `cgdi`, `cbgi`, ...) are the intrinsic capacitances.

The strategy adopted in this work follows from this observation: build an extraction testbench that does not require any knowledge of the model internals, and then use the operating-point variables of the PDK as an independent check of what this testbench returns.

3.2.2 Extraction method

Formalism and Cadence implementation. The charge-based description of a MOSFET, in which the terminal charges rather than the terminal capacitances are the state variables, was introduced by Ward and Dutton [17] to guarantee charge conservation, and it is what makes the capacitance matrix non-reciprocal: $C_{KL} \neq C_{LK}$ in general [18]. Following the convention of the EKV model [10], the transcapacitances of a four-terminal device are defined at the operating point by

$$C_{KK} = +\left.\frac{\partial Q_K}{\partial V_K}\right|_{op}, \qquad C_{KL} = -\left.\frac{\partial Q_K}{\partial V_L}\right|_{op}, \qquad K,L = D,G,S,B. \tag{3.1}$$

The self-capacitances (diagonal terms) are positive and the transcapacitances (off-diagonal terms) carry a minus sign, so that all the elements of the capacitance matrix are positive for a well-behaved device.

In a small-signal AC analysis, a unit AC excitation is applied on a single terminal y while all the other terminals are held at their DC value. The imaginary part of the current flowing into terminal x

gives directly

$$C_{xy} = -\frac{\mathrm{Im}(I_x)}{\omega\, V_y}, \qquad C_{xx} = +\frac{\mathrm{Im}(I_x)}{\omega\, V_x}. \tag{3.2}$$

In practice each terminal is biased by an independent DC source and only one of them carries the AC excitation (acm = 1). A DC sweep is performed on V_{FG} with $V_{BG} = 0\,\mathrm{V}$, and the analysis is repeated for $V_{DS} = 0\,\mathrm{V}$ and $V_{DS} = 0.5\,\mathrm{V}$. The AC frequency is kept low enough (1 kHz) for $\mathrm{Im}(I_x)/\omega$ to represent a quasi-static capacitance, that is, well below the frequency at which the channel charge stops following the terminal voltages instantaneously [18]. The excitation is placed on the front gate, which gives the four front-gate transcapacitances $C_{fg,fg}$, $C_{s,fg}$, $C_{d,fg}$ and $C_{bg,fg}$.

How can the intrinsic part be isolated? Equation (3.2) returns the total capacitance. The separation between the two families relies on a purely geometrical argument: the intrinsic and extrinsic charges do not scale in the same way with the device dimensions.

In the sEKV model the terminal charges are written as

$$Q_K = Q_{\mathrm{spec}} \cdot W \cdot L \cdot q_K, \qquad K = S, D, I, FG, BG, \tag{3.3}$$

where q_K is a normalised charge that depends only on the bias point. The same structure appears in the analytical derivation of Lu et al. [16], where the front-gate charge is obtained as a surface integral of the charge density along the channel:

$$Q_{fg} = W \int_0^L C_{ox1} \Big[V_{fg} - \Delta\Phi_1 - \psi_{s1}(y) \Big]\, \mathrm{d}y = C_{ox1} W L \Big(\cdots \Big). \tag{3.4}$$

In both formulations the intrinsic charge, and therefore the intrinsic capacitance, is proportional to the channel *area* $W \cdot L$. The extrinsic contributions are localised along the edges of the device: the gate–source and gate–drain overlaps and the junction capacitances scale with the gate width W (parameters `CGSO` and `CGDO`, in F/m), while the gate-to-substrate overlap at the two ends of the gate finger scales with the gate length L (parameter `CGBO`, in F/m). This leads to the first-order model used throughout this chapter:

$$\boxed{C_{\mathrm{tot}}(W, L) = C_{\mathrm{int}} \cdot WL \;+\; C_{e1} \cdot W \;+\; C_{e2} \cdot L} \tag{3.5}$$

where C_{int} is expressed in $\mathrm{F/m^2}$ and C_{e1}, C_{e2} in F/m. The unknown C_{int} is the quantity to be compared with the sEKV model, while C_{e1} and C_{e2} act as nuisance parameters that the extraction eliminates.

Implementation of the testbench. Model (3.5) has three unknowns, so three transistors with distinct geometries are required. Writing (3.5) for each of them gives the linear system

$$\begin{pmatrix} W_1 L_1 & W_1 & L_1 \\ W_2 L_2 & W_2 & L_2 \\ W_3 L_3 & W_3 & L_3 \end{pmatrix} \begin{pmatrix} C_{\mathrm{int}} \\ C_{e1} \\ C_{e2} \end{pmatrix} = \begin{pmatrix} C_1 \\ C_2 \\ C_3 \end{pmatrix}. \tag{3.6}$$

The geometries were chosen as T_1 ($W_1 = 1\,\mu\mathrm{m}$, $L_1 = 0.5\,\mu\mathrm{m}$), T_2 ($W_2 = 4\,\mu\mathrm{m}$, $L_2 = L_1$) and T_3 ($W_3 = W_1$, $L_3 = 4\,\mu\mathrm{m}$), so that T_2 differs from T_1 only by its width and T_3 only by its length. The system then decouples: the pair (T_1, T_2) isolates the W-scaling term and the pair (T_1, T_3) the L-scaling term. Solving (3.6) yields

$$C_{\mathrm{int}} = \frac{1}{W_1 L_1} \left[W_1 \frac{C_2 - C_1}{W_2 - W_1} + L_1 \frac{C_3 - C_1}{L_3 - L_1} - C_1 \right], \tag{3.7}$$

which, with the numerical values above (all dimensions in µm), reduces to the compact expression

implemented in the ADE calculator:

$$C_{\text{int}} = 2\left[\frac{C_2 - C_1}{3} + \frac{C_3 - C_1}{7} - C_1\right], \qquad C_k = \pm\frac{\text{Im}(I_k)}{\omega}, \tag{3.8}$$

the sign of C_k being given by (3.2) depending on whether a self-capacitance or a transcapacitance is measured. Since the three devices share the same bias conditions and the same AC excitation terminal, the three schematics are simply replicated in a single testbench (Fig. 3.5).

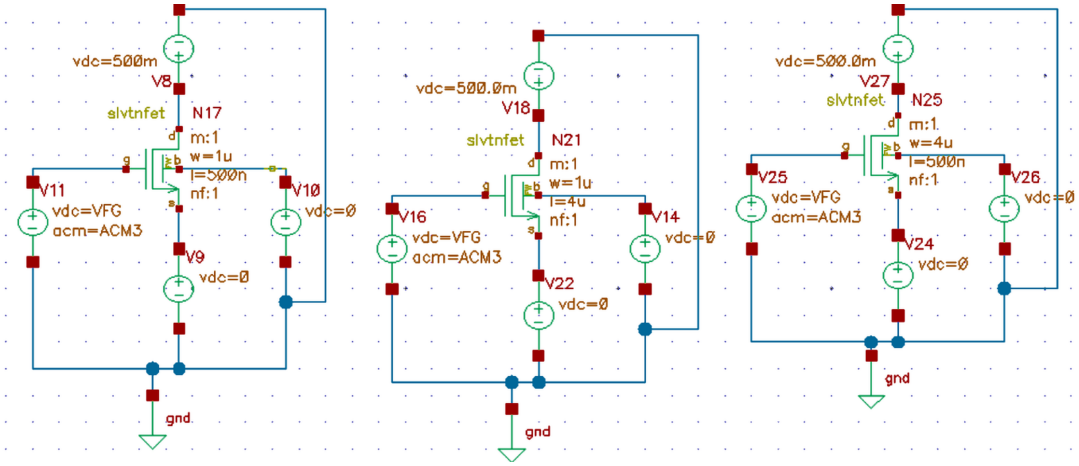


Figure 3.5: Cadence implementation of the extraction testbench: three `slvtnfet` devices T_1, T_2, T_3 sharing the same bias conditions and the same AC excitation on the front gate, and differing only by their W and L.

3.2.3 Validation of the extraction

Validation on the sEKV model. Before being applied to the PDK, the testbench must be validated on a model whose intrinsic capacitances are known independently. The sEKV model, being implemented in Verilog-A, offers exactly this possibility: its extrinsic capacitances can be switched off in ngspice, which provides a ground truth.

Figure 3.7 compares the transcapacitances extracted by the three-transistor testbench in Cadence, on the sEKV model *with* its extrinsic capacitances enabled, with the purely intrinsic ngspice reference. The two sets of curves superimpose over the whole sweep, for the four front-gate transcapacitances. This validates the method itself: when the extrinsic capacitances follow the W and L scaling assumed in (3.5), the extraction returns the intrinsic transcapacitances exactly.

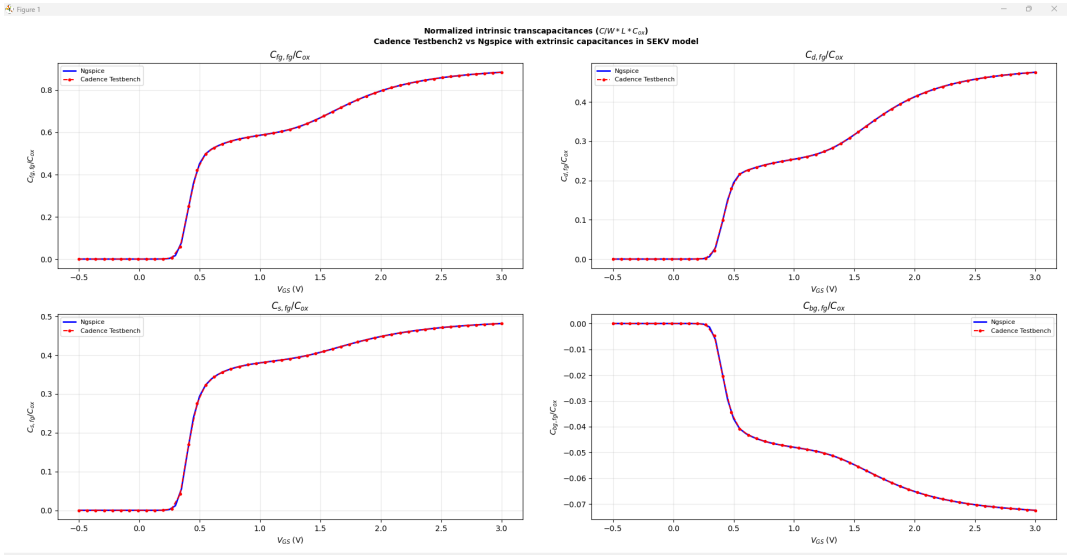


Figure 3.6: Normalised intrinsic transcapacitances (C/WLC_{ox}) extracted by the three-transistor testbench in Cadence (red) compared with the intrinsic ngspice reference (blue), for the sEKV model with extrinsic capacitances enabled in the verilog-A.

Application to the GF22FDX PDK at $V_{DS} = 0\,\mathrm{V}$. The same testbench is then applied to an `slvtnfet` device of the GF22FDX PDK. Because the PDK exposes its own intrinsic capacitances through the operating-point variables `cggi`, `cdgi`, `csgi` and `cbgi`, the extraction can be directly verified against them, and as we can observe in the Figure 3.6 the match is perfect between the intrinsic capacity extracted with the testbench I developed and the PDK reference data.

Figure 3.7 shows this comparison at $V_{DS} = 0\,\mathrm{V}$ for a $L = 500\,\mathrm{nm}$, $W = 1\,\mu\mathrm{m}$ device. The agreement is excellent on the four transcapacitances over the whole V_{FG} range: a black-box extraction, based only on terminal measurements and on the geometrical scaling of (3.5), reproduces the internal partitioning of an encrypted industrial model.

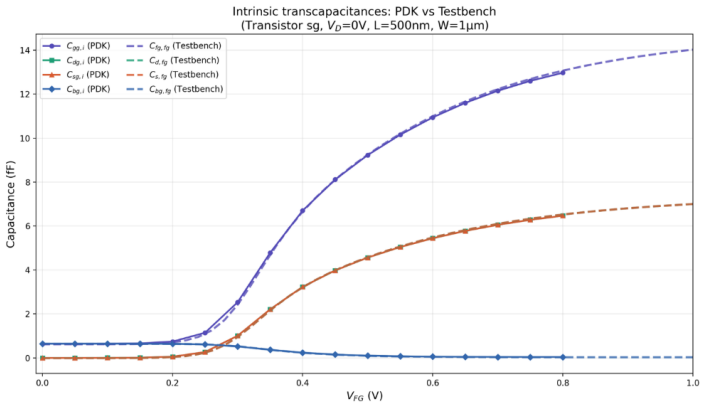


Figure 3.7: Intrinsic transcapacitances of the GF22FDX PDK at $V_{DS} = 0\,\mathrm{V}$ ($L = 500\,\mathrm{nm}$, $W = 1\,\mu\mathrm{m}$): testbench extraction (dashed) versus the operating-point variables of the PDK (solid).

Application at $V_{DS} = 0.5\,\mathrm{V}$ and origin of the $C_{d,fg}$ / $C_{s,fg}$ mismatch. The picture changes when the drain is biased. Figure 3.8 repeats the comparison at $V_{DS} = 0.5\,\mathrm{V}$. The self-capacitance $C_{fg,fg}$ and the back-gate transcapacitance $C_{bg,fg}$ still match the PDK almost perfectly, but a visible discrepancy appears on $C_{d,fg}$ and $C_{s,fg}$, and it grows with V_{FG}.

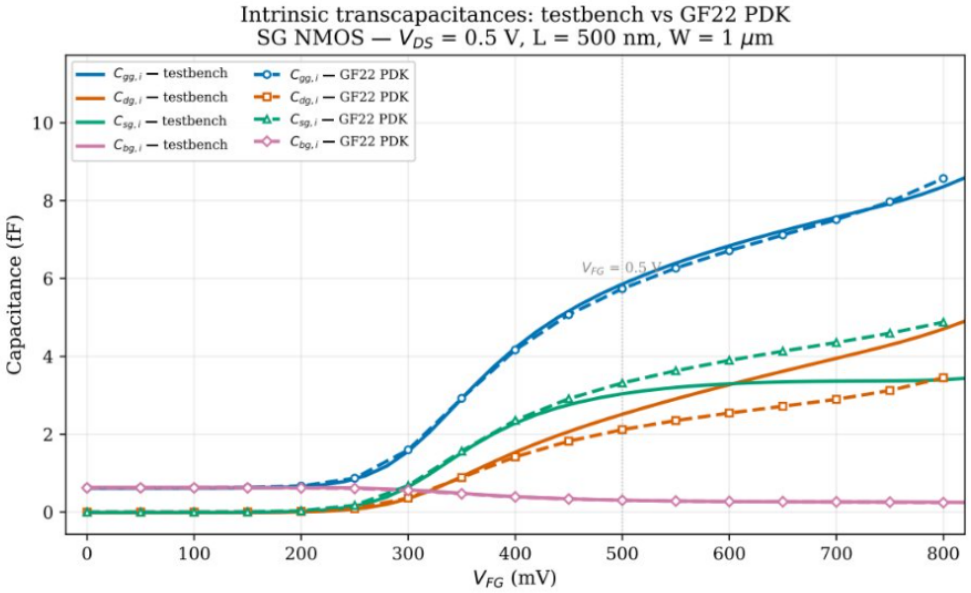


Figure 3.8: Same comparison as Fig. 3.7 but at $V_{DS} = 0.5$V. $C_{fg,fg}$ and $C_{bg,fg}$ still match, whereas $C_{d,fg}$ and $C_{s,fg}$ deviate increasingly with V_{FG}.

The origin of this deviation lies in the way model (3.5) treats the W-scaling extrinsics. A *single* coefficient C_{e1} lumps together the source-side overlap, the drain-side overlap and both junction capacitances, which is legitimate only if the two sides are identical. At $V_{DS} = 0$V the device is electrically symmetric, $C_{e1,s} = C_{e1,d}$, the assumption holds and the extraction is exact, which is what Fig. 3.7 shows. As soon as $V_{DS} \neq 0$, the drain-side contributions become bias-dependent through V_{DB} (the overlap capacitance $C_{gdov}(V_{DB})$ and the junction capacitance $C_{jdb}(V_{DB})$) while the source-side contributions remain static. The single coefficient C_{e1} then behaves as the average of two unequal quantities, which produces a systematic and *antisymmetric* error: what is subtracted in excess on the drain side is subtracted in default on the source side. The self-capacitance $C_{fg,fg}$, which sums both sides, is left almost unaffected, and so is $C_{bg,fg}$, which does not involve the source/drain overlaps.

Removing this limitation requires splitting the coefficient into a source and a drain contribution:

$$C_{\text{tot}}(W,L) = C_{\text{int}} \cdot WL + C_{e1,s} \cdot W + C_{e1,d} \cdot W + C_{e2} \cdot L, \tag{3.9}$$

which introduces a fourth unknown and therefore requires a fourth geometry. This refinement was not implemented here and is left as a perspective. Table 3.2 summarises the resulting domain of validity.

Table 3.2: Validity of the three-transistor extraction on the GF22FDX PDK, assessed against the operating-point variables of the PDK.

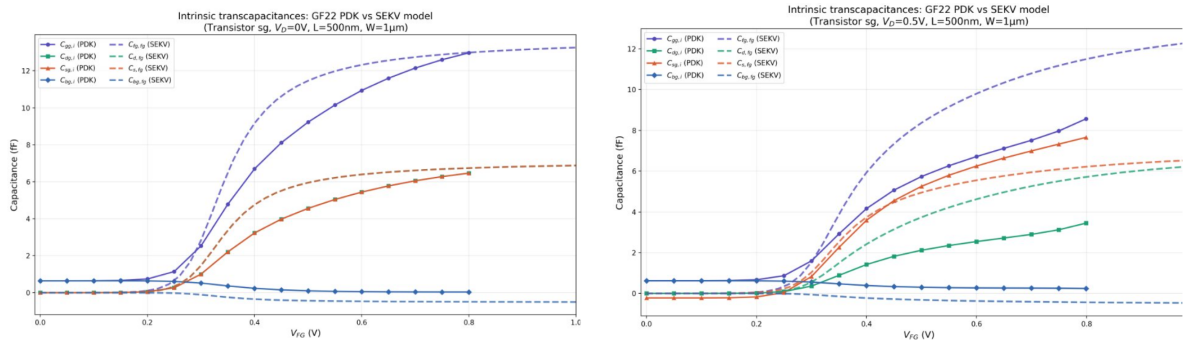
Transcapacitance	$V_{DS} = 0$V	$V_{DS} = 0.5$V
$C_{fg,fg}$	✓	✓
$C_{bg,fg}$	✓	✓
$C_{d,fg}$	✓	×
$C_{s,fg}$	✓	×

The testbench remains useful for the purpose of this work despite this restriction. $C_{fg,fg}$ and $C_{bg,fg}$ are valid over the whole bias range and they are the dominant circuit-level observables, since they set the input capacitance, the gain–bandwidth product and the noise contribution of the device. The method also provides an independent black-box cross-check of the internal partitioning of

the PDK, confirming that `cggi` and `cbgi` are consistent with what is seen at the terminals. Most importantly, it allows the sEKV model to be validated at AC level without any modification of the Verilog-A source, that is exactly under the conditions in which the model will be used for circuit design.

3.2.4 Comparison between the GF22FDX PDK and the sEKV model

The extraction can now be used for its actual purpose: comparing the intrinsic transcapacitances predicted by the sEKV model, using the parameter set extracted from the DC characterisation, with those of the GF22FDX PDK. The comparison is performed on an SG NMOS device with $L = 500\,\text{nm}$ and $W = 1\,\mu\text{m}$, at $V_{DS} = 0\,\text{V}$ and $V_{DS} = 0.5\,\text{V}$ (Fig. 3.9).



(a) $V_{DS} = 0\,\text{V}$ **(b)** $V_{DS} = 0.5\,\text{V}$

Figure 3.9: Intrinsic transcapacitances of the GF22FDX PDK (solid, from the operating-point variables) and of the sEKV model (dashed, extracted with the three-transistor testbench), for an SG NMOS $L = 500\,\text{nm}$, $W = 1\,\mu\text{m}$.

The two models agree on the essential features: the threshold voltage at which the capacitances start to rise, the sign and the order of magnitude of each transcapacitance, and the monotonic behaviour of $C_{bg,fg}$. The residual differences are of two kinds. The first one is a difference of *charge-partitioning convention*. The way the inversion charge is split between the source and the drain terminals is a modelling choice, not a physical observable: BSIM-IMG and sEKV do not use the same partitioning, which directly affects $C_{d,fg}$ and $C_{s,fg}$ while leaving their sum, and therefore $C_{fg,fg}$, unchanged. This matches Fig. 3.9, where the disagreement is concentrated on the source/drain terms. The second one is an amplitude offset on $C_{fg,fg}$ in strong inversion. The sEKV model, calibrated on DC measurements only, has no AC-specific parameter fitted at this stage. Extracting the sEKV parameters dedicated to the AC operating mode is identified as the next step of this work. The AC parameters here were determined manually. Finally, Fig. 3.10 puts the intrinsic and the total (intrinsic + extrinsic) transcapacitances side by side, normalised to $C_{ox1}WL$, for both bias conditions. It shows the magnitude of the extrinsic offset that the extraction removes: below threshold, where no inversion charge is present, the intrinsic capacitances tend to zero whereas the total ones keep a finite plateau of the order of a few percent of $C_{ox1}WL$.

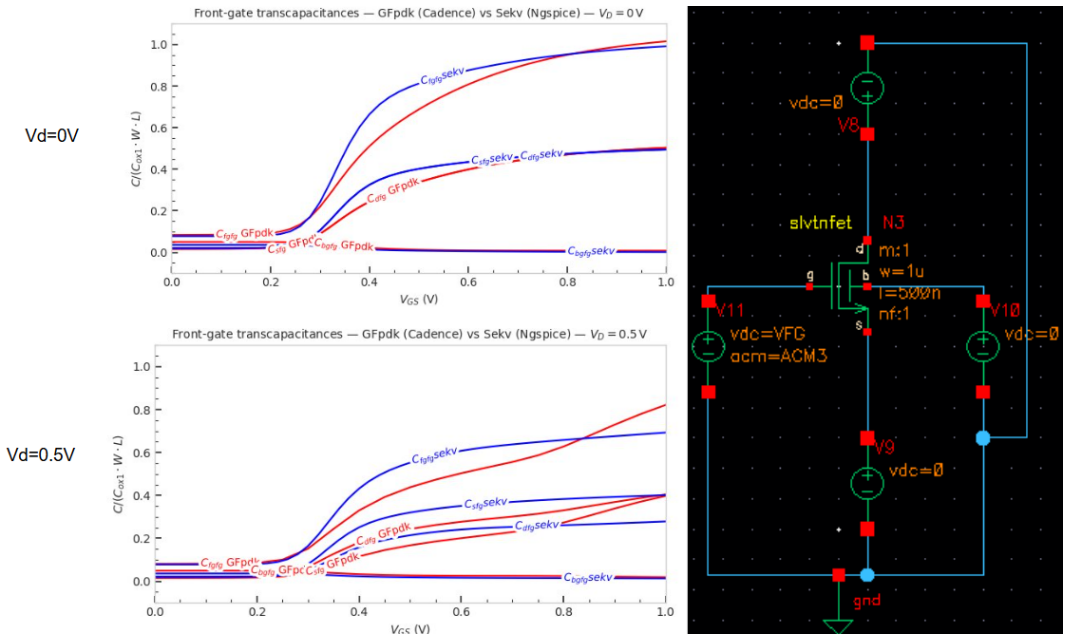


Figure 3.10: Front-gate transcapacitances normalised to $C_{ox1}WL$, GF22FDX PDK (Cadence) versus sEKV (ngspice), for $V_{DS} = 0$ V (top) and $V_{DS} = 0.5$ V (bottom), on SG $L = 500$ nm, $W = 1$ µm devices.

3.2.5 Conclusion

The extraction reproduces the operating-point variables of the encrypted GF22FDX PDK on the four front-gate transcapacitances at $V_{DS} = 0$ V, and remains exact on $C_{fg,fg}$ and $C_{bg,fg}$ at $V_{DS} = 0.5$ V. The deviation on $C_{d,fg}$ and $C_{s,fg}$ comes from the single W-scaling coefficient, which cannot represent the bias-dependent asymmetry between the two sides; the refined model (3.9), which separates $C_{e1,s}$ from $C_{e1,d}$ and requires a fourth geometry, is the direct perspective. The residual differences between the sEKV and the PDK are attributable to the charge-partitioning convention and to the absence of an extraction dedicated to the AC regime.

3.3 Gate resistance: extraction, calibration and modelling

The gate resistance R_G does not affect the DC characteristics of a MOSFET, but it sets the input impedance, f_{max} and the thermal noise of any RF or small-signal use of the device, so a compact model intended for LNA-type design has to carry it. This chapter starts from the physical model of the gate electrode established in the literature, which decomposes R_G into a distributed horizontal electrode term, a vertical access term through the gate stack, a contact term and a strap term. The published reference values of the coefficients that best match the device studied here (a high-κ/metal-gate FDSOI transistor) are taken as a starting point, and an extraction flow is then built to calibrate those coefficients on the device itself, using an automated operating-point read and an independent small-signal Y-parameter extraction over W, L and n_f. The calibrated model reproduces the extraction to 2.2 % RMS. A component decomposition and a measured AC sensitivity study then justify dropping every term but one, which is implemented in a single Verilog-A line.

3.3.1 Why the sEKV needs a gate resistance at all

At DC the gate of a MOSFET carries no current, so R_G has no influence whatsoever on I_d–V_g or I_d–V_d. It is invisible to the DC part of the model and tempting to ignore. It is not ignorable in the small-signal regime: the gate resistance sets the real part of the input impedance, it enters f_{max}, and it is the dominant extrinsic contributor to the input-referred thermal noise [19], [20]. Since the sEKV built here is meant to be used for the LNA of the Rhonexum front-end, R_G has to be present, scalable

in W, L and n_f, and physically meaningful, but also as light as possible, since every parameter added is a parameter that has to be extracted and maintained. This tension is what the chapter resolves.

3.3.2 The physical model of the gate electrode in the literature

Horizontal electrode and the factor $1/3$**.** The gate is not an equipotential node: the signal is injected at the contact and propagates along the finger width through a resistive electrode which is capacitively loaded by the channel all along its length. Razavi, Yan and Lee showed that, for the purposes of f_{max}, thermal noise and time response, this distributed line can be lumped into a single series resistor equal to the total electrode resistance divided by three for a single-side gate connection [19]. The derivation is reproduced here because the factor is what identifies this term in the extracted data. Modelling the finger as a line of total resistance R_ℓ with the gate current injected uniformly along the width W, the current grows linearly from the open end, $I(x) = I_{tot}\,x/W$, so the dissipated power is

$$\int_0^W I(x)^2\,\frac{R_\ell}{W}\,dx \;=\; I_{tot}^2\,\frac{R_\ell}{3},$$

i.e. an effective lumped resistance $R_\ell/3$. Contacting both sides halves the length and puts the two halves in parallel, which gives $R_\ell/12$.

Vertical access through the gate stack. Litwin pointed out that the horizontal term alone is not enough: the signal also has to travel *vertically*, from the low-resistance top of the gate stack down through the interface into the layer that actually controls the channel, and that this interfacial contact resistance is of the same order as the sheet resistance of the top layer itself [21]. Because it is a current crossing an area rather than running along a line, it scales as $1/(W_f L)$, that is inversely with the gate area: negligible on large devices, dominant on narrow and short ones. Jo and Matsuoka combined the two mechanisms into a single scalable expression and validated it on a 130 nm process [22]; Dormieu *et al.* revisited the same decomposition specifically for high-κ/metal-gate stacks, where the top conductor is a metal rather than a silicide but the vertical access through the stack remains [23]. The approach has since been carried to 16 nm FinFETs [24], which is reassuring for a 22 nm FDSOI device.

Contacts, gate extension and strap. Two extrinsic terms complete the picture: the contacts between the metal and the top of the gate stack, which appear as N_{cg} identical resistances in parallel per finger, and the portion of the gate that extends beyond the channel area before reaching the contact, which simply adds a length to the horizontal term [22], [23]. A multi-finger device needs one further element that none of the above papers has to carry, because their fingers are contacted individually: the metal rail that straps the fingers together is a *series* element shared by all of them, so it does not benefit from the $1/N_F$ parallel factor and it grows with the finger count.

Resulting general expression. Putting the four mechanisms together gives the expression taken as the model here, written in a flavour-independent form:

$$R_{\mathrm{G}} = \frac{1}{N_F}\left[\underbrace{\frac{R_\square}{L_g}\left(\frac{W_f}{3}+\Delta_{ext}\right)}_{\text{horizontal electrode}} + \underbrace{\frac{\rho_v}{W_f L_g}}_{\text{vertical access}} + \underbrace{\frac{R_c}{n_c}}_{\text{contacts}}\right] + \underbrace{\frac{\rho_M\,[N_F L_g + (N_F-1)\,s_{pp}]}{12\,w_M\,t_M}}_{\text{strap}} + R_{\min} \tag{3.10}$$

with the effective geometry

$$L_g = L + \Delta L, \qquad W_f = \frac{W}{N_F} + \Delta W, \qquad n_c = 1 + \left\lfloor \frac{L_g - w_c}{p_c} \right\rfloor .$$

The first two terms are exactly Eq. (4) of [22] with $k = 1$ (single-side connection), the third is the contact term of [22], [23], and the fourth is the strap. The symbols are listed in Table 3.3.

Symbol	Meaning	Unit
$R_\square$	sheet resistance of the top gate conductor	Ω/sq
ρ_v	vertical (gate-stack interface) resistivity	$\Omega \cdot \mu m^2$
R_c	resistance of one gate contact	Ω
ρ_M, w_M, t_M	strap metal resistivity, width, thickness	Ω·m, m, m
w_c, p_c	contact width and pitch (they set n_c)	m
s_{pp}	gate-to-gate pitch $s_{pp}(L)$	m
Δ_{ext}	gate extension beyond the channel area	m
ΔL, ΔW	effective length / width offsets	m
$R_{\min}$	numerical floor	Ω

Table 3.3: Symbols of the general expression (3.10). Another device flavour (other V_t, the PFET) reuses the same structure with its own coefficients.

Reference values. The structure of (3.10) is generic; the coefficients are not. A starting point is needed before calibrating anything, and the choice that matters most is ρ_v, because it is the one coefficient for which the published values span an order of magnitude and it is also the one that decides the behaviour at small W.

Coefficient	Reported value	Context	Source
ρ_{int}	$\approx 25\,\Omega\mu m^2$	$TiSi_2$ on poly	[21]
ρ_{int}	≈ 2 to $3\,\Omega\mu m^2$	NiSi	[25]
$\rho_{int} + \rho_{vp}$	$12.5\,\Omega\mu m^2$	fitted, 130 nm poly gate	[22]
k (lumping factor)	1/3 / 1/12	single- / both-side connection	[19], [22]
NQS Elmore constant κ	≈ 3–5	$L < 1\,\mu m$ / $L > 1\,\mu m$	[22]

Table 3.4: Published reference values for the coefficients of (3.10). The spread on the vertical term reflects the gate stack: the older $TiSi_2$-on-polysilicon stacks are an order of magnitude worse than modern NiSi and metal-gate stacks.

The device studied here is a high-κ/metal-gate FDSOI transistor, so the $TiSi_2$ figure of [21] is not representative: the relevant branch is the low one, and the starting point is $\rho_v = 5\,\Omega\mu m^2$, which sits just above the NiSi range of [25] and well below the poly-gate fit of [22], consistently with a metal-gate stack. The remaining coefficients are set from the nominal technology figures for this device: $R_\square = 27\,\Omega$/sq, $R_c = 27\,\Omega$, $\rho_M/(w_M t_M)$ giving $0.238\,\Omega/\mu m$ for the strap, $\Delta L = 6\,nm$, $\Delta W = 14\,nm$, $\Delta_{ext} = 0.07\,\mu m$, $w_c = 40\,nm$, $p_c = 100\,nm$, and $s_{pp}(L)$ from the gate-pitch rule. The gate is contacted on one side ($k = 1/3$) and the strap carries the 1/12 factor of a line contacted at both ends. These are starting values; the next two sections check them against the device.

3.3.3 Extraction method

Calibrating (3.10) requires a measurement of R_G that can be repeated over a large number of geometries. Two methods were used, deliberately different in nature so that they cross-check each other.

Method A: operating-point read (reference). The first value is obtained directly from the simulator, and the whole flow is automated:

1. from **VSCode**, a **Python** session runs on the VM and drives **Cadence Virtuoso / ADE Maestro** through **Skillbridge** (the Python/SKILL bridge);

2. an **OCEAN** script configures and launches a **DC operating point** on the transistor (instance `N0`);
3. once it finishes, the gate resistor that the device model instantiates between the external gate and the model core is read from the operating-point information (`dcOpInfo`) and written to a CSV;
4. Python sweeps W, L and n_f parametrically (`maeSetVar` + `maeRunSimulation`) and accumulates the points.

```
openResults(psf_path)
selectResult('dcOpInfo)
rg = getData("N0.rgres" ?result 'dcOpInfo :res)    ; the :res suffix is required
fprintf(port "%g,%g,%g,%g\n" W L NF rg)            ; one point written to the CSV
```

This gives a dense, cheap reference over the whole geometry space. Its limitation is that it is a *static* read: it gives the resistance the model carries, not the resistance a circuit actually sees at the frequencies where R_G matters.

Method B: small-signal Y-parameters (calibration). The second method is the standard high-frequency characterisation of R_G [26], and it is the one used for the calibration, because it probes the device in exactly the small-signal regime the sEKV is meant to reproduce. The testbench (Fig. 3.11) is a single `slvtnfet` with the front gate on port 1 and the drain on port 2, source and back-gate grounded. The device is biased in the linear region ($V_{DS} = 0\,\mathrm{V}$, V_{GS} 0.4 V above threshold) so that a distributed channel is present under the gate; an unbiased device has no channel to load the electrode and returns a meaningless MΩ-range result. A single `sp` analysis from 100 MHz to 10 GHz gives the full two-port Y matrix, from which

$$R_G = \frac{\mathrm{Re}(Y_{12})}{\mathrm{Im}(Y_{12})\,\mathrm{Im}(Y_{11})}, \tag{3.11}$$

read on the low-frequency plateau (here 1 GHz). One run yields Y_{11} and Y_{12} together: the two terminal conditions of the method are the definitions of the two-port parameters, applied inside the S-to-Y conversion, not two separate benches.

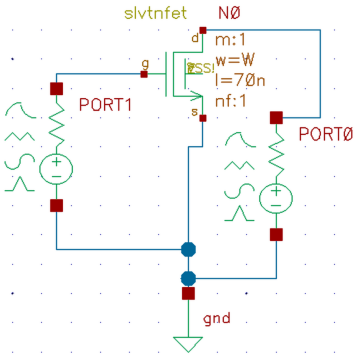


Figure 3.11: R_G extraction testbench: front gate on port 1, drain on port 2, source and back-gate grounded, device biased in the linear region.

3.3.4 Calibration

Method B was run over three sweeps, namely W at $L = 70\,\mathrm{nm}$, $n_f = 1$; L at $W = 1\,\mu\mathrm{m}$, $n_f = 1$; and n_f at $W_{tot} = 8\,\mu\mathrm{m}$, $L = 1\,\mu\mathrm{m}$, and the free coefficients of (3.10) were fitted to the three sweeps simultaneously by least squares. Fitting all three at once is what makes the result meaningful: each sweep constrains a different term (W separates horizontal from vertical, L constrains the contacts, n_f constrains the strap), so a coefficient set that fits all three cannot be absorbing an error into the wrong term.

A missing term. The first fit refused to close: the extracted points sat systematically *above* (3.10) by an offset that was flat versus L (Table 3.5) and versus n_f, but that scaled with the total device width. That behaviour rules out the two natural candidates. An underestimated vertical access would grow as $1/L$ and diverge at short L, which it does not. A non-quasi-static channel term κ/g_m [22] would grow with L, which it also does not. What remains is a width-set access resistance, added as one extra term

$$R_{gf} = \frac{k_{gf}}{W_{tot}},$$

after which the fit closes to 2.2 % RMS over all three sweeps.

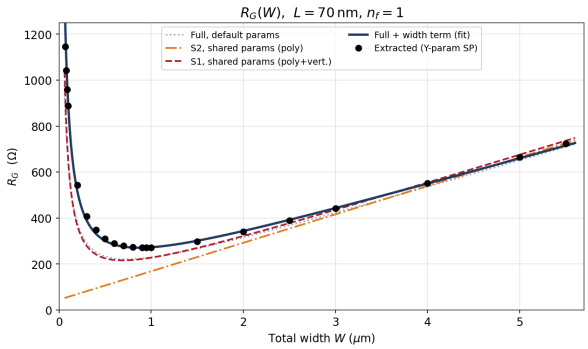
L (nm)	R_G extracted (Ω)	(3.10), $\rho_v = 5$ (Ω)	offset (Ω)
22	623	597	27
70	273	237	36
200	127	91	36
500	72	37	35
950	54	20	34

Table 3.5: L sweep at $W = 1\,\mu$m, $n_f = 1$. The residual is flat versus L ($\approx 35\,\Omega$), not the $1/L$ growth a vertical term would give nor the growth with L an NQS term would give. Read on the n_f sweep at $W_{tot} = 8\,\mu$m the same residual is only about 4.5 Ω, again flat, hence the $\approx 36\,\Omega\mu\text{m}/W_{tot}$ law.

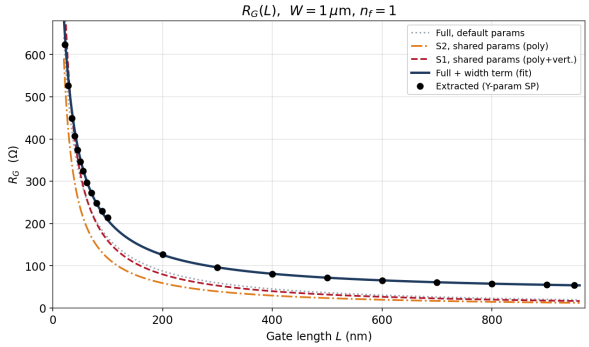
Coefficient	Starting value	Calibrated	Comment
$R_\square$	27 Ω/sq	26.0 Ω/sq	−3.7 %, well within tolerance
ρ_v	5 Ωμm^2	4.14 Ωμm^2	consistent with a metal-gate stack [25]
R_c	27 Ω	40 Ω	saturated at the upper bound allowed
Δ_{ext}	0.07 μm	0.121 μm	absorbs part of the layout extension
k_{gf}	—	33.0 Ωμm	added width-access term, $R_{gf} = k_{gf}/W_{tot}$

Table 3.6: Calibrated coefficient set (RMS 2.2 % over the three sweeps). The purely geometric quantities (ΔL, ΔW, w_c, p_c, s_{pp}, the strap cross-section) are held at their nominal values and not fitted.

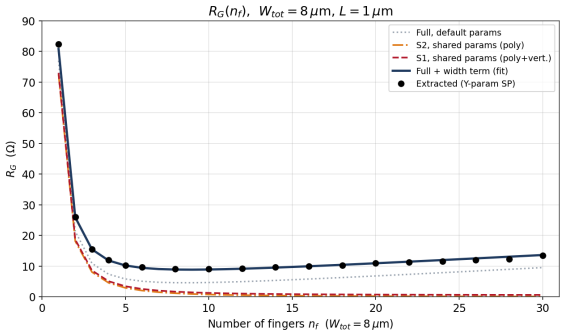
Two of the calibrated values carry information. $\rho_v = 4.14\,\Omega\mu\text{m}^2$ lands almost exactly where Table 3.4 predicts a metal-gate stack should be: an order of magnitude below the $TiSi_2$ figure of [21] and a factor three below the poly-gate fit of [22]. The fit reaches this value on its own, from an entirely independent small-signal measurement, which supports the term of (3.10) being physical and not a curve-fitting artefact. R_c, on the other hand, saturated against the upper bound of the range allowed, which indicates that the contact term is being asked to absorb something else, most likely part of the same width-set access resistance, and that no physical contact resistance should be read out of it.



(a) $R_G(W)$ at $L = 70$ nm, $n_f = 1$



(b) $R_G(L)$ at $W = 1\,\mu$m, $n_f = 1$



(c) $R_G(n_f)$ at $W_{tot} = 8\,\mu$m, $L = 1\,\mu$m

Figure 3.12: Calibrated model (solid) against the extracted points over the three geometry sweeps. (a) The minimum near 1 µm is the crossover between the vertical access, dominant at small W, and the horizontal electrode, dominant at large W. (b) The $1/L$ decay of the electrode term over almost two decades. (c) The drop is the $1/N_F$ parallel factor; the upturn beyond $n_f \approx 10$ is the strap, which is a series element and therefore grows with the finger count.

3.3.5 Simplification for the sEKV

A five-coefficient expression is acceptable for a characterisation report but not for a compact model, where every coefficient has to be maintained and where some of these terms will be extracted a second time by the layout parasitic extraction. Two criteria decide what to keep: the magnitude of each term, and whether a resistive PEX would re-extract it.

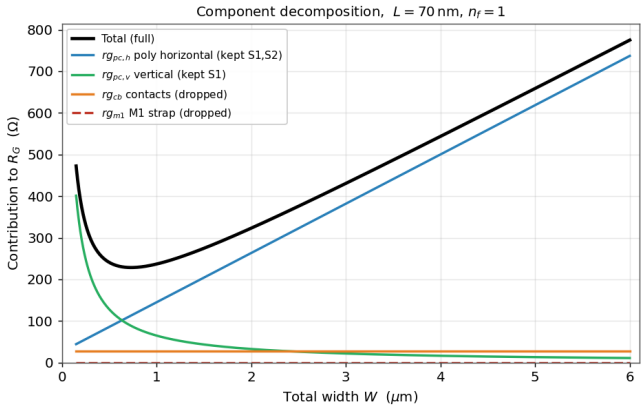


Figure 3.13: Contribution of each term of (3.10) at $L = 70$ nm, $n_f = 1$. The horizontal electrode dominates, the vertical access carries the small-W rise, the contacts are a nearly constant offset and the strap stays below 0.1 Ω.

The decomposition settles the two extrinsic terms immediately. The strap stays below 0.1 Ω over

the whole width range, and the contacts are a moderate, nearly constant offset of about 9 Ω; both are local interconnect that a PEX re-extracts, so keeping them in the compact model would be double-counting on top of being negligible. Dropping them leaves

$$\boxed{R_{\mathrm{G}}^{S1} = \frac{R_{\square}}{n_f^2}\,\frac{W/3 + \Delta_{ext}}{L} + \frac{\rho_v}{W\,L}} \tag{3.12}$$

and dropping the vertical access as well leaves a single term:

$$\boxed{R_{\mathrm{G}}^{S2} = \frac{R_{\square}}{n_f^2}\,\frac{W/3 + \Delta_{ext}}{L} \approx \frac{9\,W}{n_f^2\,L}} \tag{3.13}$$

That second step is the only one that removes an *intrinsic* term, so it is the only one that needs a real justification. Table 3.7 shows the price in resistance, and it is not small in relative terms.

Sweep	point	full (Ω)	S1 (Ω)	dev. S1	S2 (Ω)	ΔS2 (Ω) / dev.
W ($L = 250\,$nm)	$W = 0.3\,\mu$m	89.7	80.6	−10.1 %	18.4	71.3 / −79.5 %
	$W = 1\,\mu$m	71.4	62.3	−12.7 %	43.0	28.4 / −39.7 %
	$W = 6\,\mu$m	231.1	222.1	−3.9 %	218.8	12.3 / −5.3 %
L ($W = 1\,\mu$m)	$L = 30\,$nm	470.0	443.0	−5.7 %	306.0	164.0 / −34.9 %
	$L = 250\,$nm	71.4	62.3	−12.7 %	43.0	28.4 / −39.7 %
	$L = 1\,\mu$m	18.8	15.9	−15.6 %	11.0	7.8 / −41.7 %
n_f ($W_{tot} = 8\,\mu$m)	$n_f = 1$	77.1	74.2	−3.8 %	73.6	3.5 / −4.6 %
	$n_f = 8$	4.6	2.0	−56.9 %	1.4	3.2 / −70.3 %
	$n_f = 30$	9.5	0.7	−92.3 %	0.1	9.4 / −98.5 %

Table 3.7: Deviation of S1 and S2 against the calibrated expression. The large relative errors at high n_f sit on a near-zero base of a few ohms and reflect the missing strap, not a meaningful absolute error.

Does a resistance error of that size actually matter? A relative error on R_{G} is only a problem if it moves something that matters, so rather than argue about tolerances it was measured. A controlled resistor `R0` was inserted in series with the front gate of a single device (Fig. 3.14), its value was swept, and the resulting displacement of the extracted transcapacitances was recorded. Every terminal carries a DC bias (`vdc`) and an AC magnitude (`acm`), and the transcapacitances are read from the imaginary part of the terminal currents, $C = \mathrm{imag}(I)/2\pi f$. At each run the added resistance and the width were changed; each curve is referenced to its own lowest-resistance point.

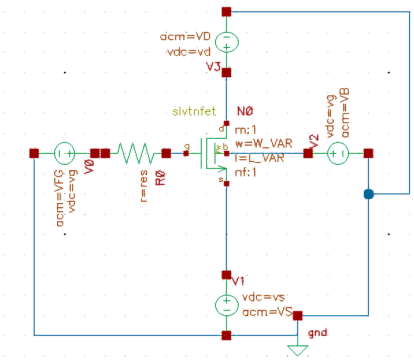


Figure 3.14: AC testbench: one `slvtnfet` with the added gate resistor `R0` (`r=res`) in series with the front gate.

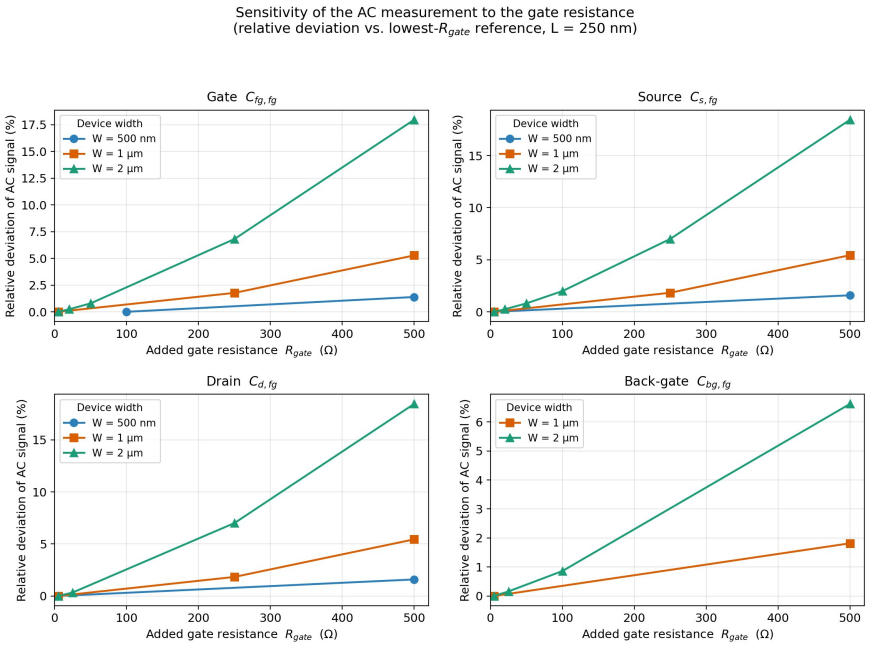


Figure 3.15: Relative deviation of the extracted transcapacitances versus added gate resistance, referenced to the lowest-R_{gate} point, $L = 250\,\text{nm}$.

Result. The sensitivity is strongly width-dependent. At $W = 500\,\text{nm}$ every capacitance stays within about 1.5 % even for 500 Ω of added gate resistance; at $W = 1\,\mu\text{m}$ the deviation reaches roughly 5 %; only at $W = 2\,\mu\text{m}$ does it climb to about 18 %. The width beyond which an error on R_G starts to matter for the AC signal is therefore around 1 µm.

Crossing the two studies quantitatively closes the argument (Table 3.8): the resistance error of S2, read on the sensitivity curve of the matching width, costs at most about 0.3 % on the extracted capacitances.

W (μm)	ΔR_g(S1) (Ω)	AC dev. S1	ΔR_g(S2) (Ω)	AC dev. S2
0.5	9.1	$\lesssim 0.05\,\%$	47.1	$\lesssim 0.1\,\%$
1	9.1	$\lesssim 0.05\,\%$	28.3	$\lesssim 0.2\,\%$
2	9.1	$\lesssim 0.1\,\%$	18.8	$\lesssim 0.3\,\%$

Table 3.8: Resistance error crossed with the measured AC sensitivity, $L = 250\,\text{nm}$. $\Delta R_g = \text{full} - \text{S}_i$; the AC deviation is read from Fig. 3.15 at that added resistance on the matching width curve.

Decision. S2 is the form implemented. The largest relative errors of S2 occur exactly where the AC is least sensitive: at small W, where the missing vertical term is largest but where the capacitances barely move, and at high n_f, where the absolute resistance is only a few ohms. In the region where the AC does respond ($W \gtrsim 2\,\mu\text{m}$) S2 is already accurate to a few percent. The residual k_{gf}/W_{tot} term found during calibration follows the identical logic: it is largest on narrow devices and falls to a few ohms on the wide ones. Combined with the fact that R_G is invisible at DC, the single-term form costs nothing measurable and gives a model with one parameter instead of five and no overlap with PEX.

3.3.6 Implementation in the sEKV

```
parameter real RGH_SHEET = 26.0;    // ohm/sq  gate sheet resistance (calibrated)
parameter real EXT_DELL  = 6e-9;    // m       length offset
parameter real FRINGE    = 121e-9;  // m       gate extension (calibrated)
parameter real RGMIN     = 1e-3;    // ohm     numerical floor
real Lg, Wf, RGeff;
```

```
@(initial_step) begin
    Lg    = L + EXT_DELL;
    Wf    = W/NF;
    RGeff = max(RGH_SHEET*(Wf/3.0 + FRINGE)/(NF*Lg), RGMIN);
end
// V(g,gi) <+ RGeff*I(g,gi);
// (+ white_noise(4*`P_K*$temperature*RGeff) for RF noise)
```

One computed line, two calibrated coefficients. The contacts and the strap, being post-layout parasitics, are left to extraction. Two limits are worth stating explicitly. First, this conclusion holds under the tolerance, bias point and frequency of the extraction reported here: if any of them tightens, for instance for a noise-critical LNA design or for operation well above 1 GHz, the vertical term of (3.12) should be reconsidered, since it is the one intrinsic mechanism that was dropped. Second, if the PEX deck were configured to extract the gate electrode itself and not only the local interconnect, even the horizontal term would become redundant and the sEKV should carry no gate resistance at all; that is worth confirming on the extraction flow before tape-out.

4 Cryogenic CMOS design and optimization

4.1 Design verification and optimization of a 4–6 GHz cryogenic CMOS LNA using the sEKV compact model

The readout of superconducting qubits requires an amplification chain operating at cryogenic temperature, in which a semiconductor low-noise amplifier (LNA) provides a gain above 35 dB with a noise temperature of a few kelvins in the 4–8 GHz range. These amplifiers are today implemented in InP HEMT technology. A CMOS implementation would be attractive for integration, but reaching the required noise and power performance in CMOS is still an open problem.

The reference design of this work is the cryogenic LNA published by Das, Raman and Bardin [1]: a three-stage common-source amplifier with inductive source degeneration, fabricated in GlobalFoundries 22 nm FDSOI (GF22FDX). At its nominal bias (V_{DD} = 0.6 V, 15.8 mW) it reaches a peak gain of 38 dB and an average noise temperature T_e of 4.4 K over 4–6 GHz. A low-power bias point is also reported (34 dB peak gain, T_e = 5.1 K, 4.5 mW).

The point that motivates this work is a limitation of their design methodology. The authors did not have a compact model of the transistor valid at cryogenic temperature and usable in a circuit simulator such as Cadence Spectre. Their transistor description is a small-signal model extracted at a *single* bias point (J_D = 37.5 mA/mm, Table I of [1]), and the simulated curves shown in the article are computed from this fixed model. This is the same type of analytical small-signal study as the ABCD model presented in Part II of this report: it predicts the response at the extraction bias, but it cannot be evaluated at another operating point. As a consequence, the bias of the three stages is a trade-off fixed by hand at design time, and the only knob left after fabrication is the back-gate voltage, which is common to the three stages. No systematic optimization was possible.

The tool that the authors did not have is exactly what this work brings: the sEKV compact model, presented earlier in this report, implemented in Verilog-A and running in Cadence. Since this model is continuous in bias, any operating point can be simulated, the channel noise can be made bias-dependent, and the three stages can be biased independently. The objective of this part is therefore double. First, to verify to what extent the sEKV model reproduces the measured transistor and the published circuit (Part I). Second, to use the verified model to perform what the authors could not: a systematic optimization of the passive network and of the per-stage biasing, followed by a verification of the stability and linearity of the optimized circuit (Part II). The specifications used in this work are $S_{21} > 35$ dB, $S_{11} < -10$ dB and $S_{22} < -10$ dB over the 4–6 GHz band.

All the simulations of this study are run at 4.4 K, the same temperature as the reference work, and they all include the series resistance of the inductors. Each inductor L is modeled with a series resistance

$$R_s = \frac{2\pi f_0 L}{Q}, \qquad Q = 40, \quad f_0 = 5\ \text{GHz}, \tag{4.1}$$

which corresponds to the quality factor reported at cryogenic temperature by the electromagnetic simulations of the article. The losses are therefore taken into account everywhere, with the approximation that R_s is computed once at the band centre and kept constant in frequency.

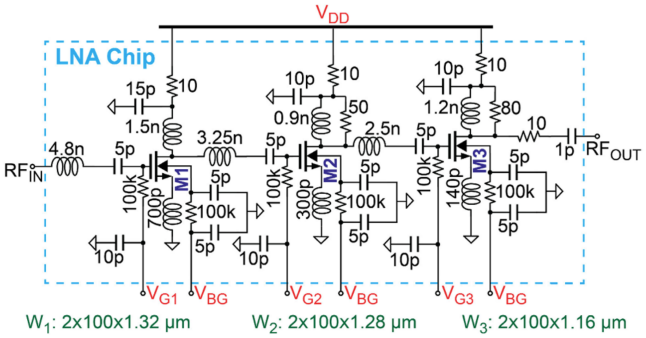


Figure 4.1: Schematic of the three-stage cryogenic LNA, reproduced from [1], with the component values published in the article. All transistors are minimum length.

4.1.1 Part I: verification of the model

Before optimizing anything, we must establish how much the sEKV model can be trusted. The verification is done in two steps: first at the transistor level, against the small-signal parameters measured in [1], then at the circuit level, by simulating the complete published LNA.

Implementation of the channel noise model. The sEKV model did not include a description of the channel noise valid at cryogenic temperature. Following [1], the channel current noise of a cryogenic MOSFET is described as suppressed shot noise,

$$\overline{|i_{n,d}|^2} = 2\,q\,F\,I_{DS}\,\Delta f, \tag{4.2}$$

where q is the elementary charge, I_{DS} the DC drain current and $F \leq 1$ the Fano factor, which accounts for the shot-noise suppression. The authors showed that F is temperature-independent, which allows its extraction from room-temperature measurements.

This noise source was added to the Verilog-A description of the sEKV model. The important point of the implementation is that F is not a constant. The dedicated paper of the same authors [27] provides measured values of F at several current densities, and a linear regression on these points (figure 4.2) gives

$$F(J_D) = 0.8723 - 0.004337\,J_D \quad [J_D \text{ in mA/mm}], \tag{4.3}$$

where $J_D = |I_D|/(W \cdot M)$ is evaluated inside the model at the actual operating point of each transistor. At the bias point of the article ($J_D = 37.5$ mA/mm), equation (4.3) gives $F \approx 0.71$, consistent with the value $F = 0.7$ of Table I in [1]. This bias dependence is what makes the noise follow the polarization, which is a necessary condition for the bias optimization of Part II.

The noise is extracted in Cadence through the noise figure of the amplifier. The noise temperature of the ports is set to $-268.6\,^{\circ}$C, so that the noise figure is referenced to the operating temperature of the circuit rather than to the 290 K room temperature convention, and the noise temperature is then obtained with

$$T_e = 4.4\left(10^{\mathrm{NF}/10} - 1\right), \tag{4.4}$$

which is the standard relation written with the reference temperature used here.

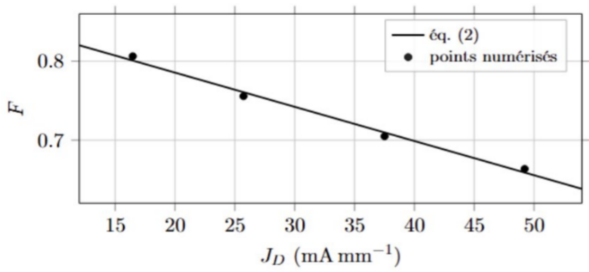


Figure 4.2: Fano factor versus drain current density: digitized measurement points from [27] and the linear regression of equation (4.3) implemented in the Verilog-A model.

Verification at the transistor level. The only quantitative reference available at the device level is Table I of [1], which gives the small-signal parameters extracted at $J_D = 37.5$ mA/mm. The verification is therefore done at a single point. This limits its scope, but it is the best comparison the published data allows.

The simulated devices use the geometry of the article ($W_1 = 264$ µm, $W_2 = 256$ µm, $W_3 = 232$ µm, minimum length, $NF = 100$, $M = 2$). The gate–drain overlap capacitance C_{gdo} is not provided by the cryogenic extraction. It was set manually so that its value matches the one given by the GF22 PDK at room temperature, which is justified by the fact that the overlap capacitance is essentially geometric and is therefore not expected to change significantly with temperature. The exact measurement network of the transistor was reproduced in Cadence, and the bias point of the article was recovered by sweeping V_G.

Table 4.1: Small-signal parameters: measurement of [1] (Table I) versus sEKV model, at $J_D = 37.5$ mA/mm.

	Ref.	M1		M2	
Quantity	Table I	Value	Dev.	Value	Dev.
J_D [mA/mm]	37.5	37.9	+1%	37.6	+0.3%
g_m [mS/µm]	0.92	0.950	+3%	0.950	+3%
g_{ds} [mS/µm]	0.08	0.053	−34%	0.053	−34%
C_{gd} [fF/µm]	0.23	0.169	−27%	0.169	−27%
C_{gs} [fF/µm]	0.45	0.846	+88%	0.846	+88%
C_{ds} [fF/µm]	0.24	0.271	+13%	0.271	+13%
f_T [GHz]	221	166.4	−25%	166.3	−25%
$f_{\max}$ [GHz]	277	238.3	−14%	246.7	−11%
R_g [Ω·µm]	480	483.6	+1.2%	474	−1.2%

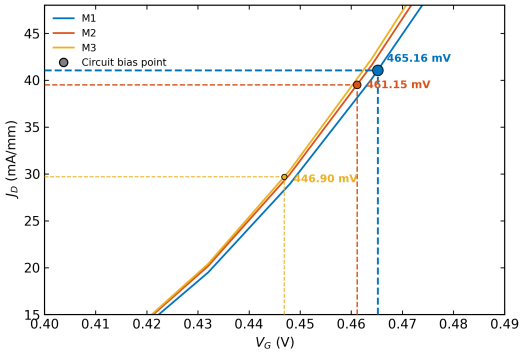


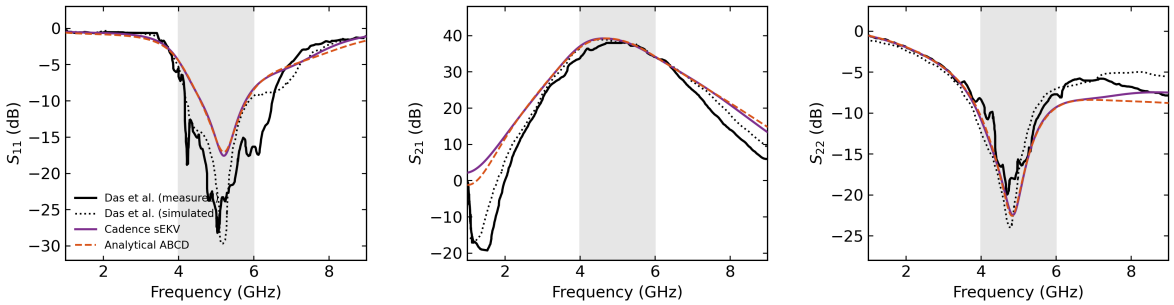
Figure 4.3: Simulated drain current density versus gate voltage for the three device geometries. The markers indicate the gate voltages that reproduce the current densities of the article.

Two families of parameters must be distinguished. The quantities related to transport and biasing are well reproduced: the current density is within 1 %, the transconductance within 3 % and the gate resistance within 1.2 %. Since the channel noise of equation (4.2) is driven by I_{DS} and $F(J_D)$, and since the noise match of the input stage is driven by g_m, the quantities on which the noise predictions rely can be trusted.

The AC parameters, on the other hand, deviate significantly: C_{gs} is overestimated by 88 % and g_{ds} is underestimated by 34 %, which propagates to f_T (−25 %). Improving these parameters would require a new extraction campaign dedicated to the AC behaviour of the model, to be carried out by the person in charge of the model extraction in the group. This is future work on the model itself and is outside the scope of this project. The practical consequence is that conclusions based on currents and noise are solid, while the exact resonance frequencies and gain levels must be read

with a tolerance.

Verification on the published circuit. The complete circuit of figure 4.1 was entered in Cadence with the passive values published in the article. This configuration is called "based on the paper's schematic" in the rest of the report. The gate voltages were adjusted to reproduce the drain currents of the article (9.9, 9.6 and 6.8 mA), which gives $V_{G1} = 464.62$ mV, $V_{G2} = 460.6$ mV and $V_{G3} = 446.2$ mV, with $V_{DD} = 0.6$ V and $V_{BG} = -1$ V. The simulated power is 15.74 mW, within 0.4 % of the 15.80 mW of the article. The reference curves of the article were digitized by hand from its figures, which introduces a reading error of its own, on top of the ± 1.1 K (2σ) uncertainty quoted by the authors on the noise measurement. The comparison must therefore be read as an agreement within a margin, not as an exact match. The results are shown in figure 4.4 and summarized in table 4.2. The behaviour is the expected one: a matched low-noise amplifier with a peak gain of 39.19 dB against 38.06 dB measured, and a minimum noise temperature of 2.41 K. The main deviation is a shift of the whole response towards lower frequencies, of about 10 %: the gain peak falls near 4.6 GHz instead of 5 GHz and the T_e minimum near 4.2 GHz instead of 5.2 GHz as we can see in the figure 4.6d. Since the band of interest stays fixed at 4–6 GHz, this detuning penalizes the band-averaged metrics, and the average T_e over 4–6 GHz reaches 6.30 K.



(a) S_{11} **(b)** S_{21} **(c)** S_{22}

Figure 4.4: S-parameters of the circuit simulated in Cadence with the sEKV model, compared with the curves of [1] digitized by hand.

Table 4.2: Based on the paper's schematic: comparison with Das *et al.* [1] over 4–6 GHz.

Metric	Das *et al.*	This work	Difference
Peak gain S_{21} (dB)	38.06	39.19	+3.0 %
Minimum gain, 4–6 GHz (dB)	33.73	34.13	+1.2 %
Max. S_{11}, 4–6 GHz (dB)	−5.52	−4.47	+19.0 %
Max. S_{22}, 4–6 GHz (dB)	−7.73	−9.31	−20.4 %
Minimum T_e (K)	3.65	2.41	−34.0 %
Average T_e, 4–6 GHz (K)	4.35	6.30	+44.8 %
P_{DC} (mW)	15.80	15.74	−0.4 %
FOM (dB/mW/ph)	0.177	0.124	−29.9 %

This frequency shift has two causes that add up. The first is the model: the input match of an inductively degenerated stage is set by the series resonance between the gate and degeneration inductors and C_{gs}, and the resonance frequency varies as $1/\sqrt{C_{gs}}$. The 88 % overestimation of C_{gs} found at the transistor level must therefore push the response downwards; taken alone it would predict a much larger shift than the 10 % observed, but C_{gs} is only one contribution to the total capacitance of the input resonator, so its influence is diluted. The second cause is the reference itself: the article measures a fabricated chip while the simulation uses the published schematic values without layout extraction, so the two responses cannot fall at exactly the same frequency.. In

the same way, the gain being 1.1 dB above the measurement is consistent with the underestimated g_{ds}, which raises the output impedance of each stage.

The circuit is therefore a credible but detuned starting point. Before any bias study, the passive network has to be re-tuned so that the amplifier is centred on its band, which also absorbs the schematic-to-layout difference. This is the object of Part II.

4.1.2 Part II: optimization

The optimization uses two levers, in two steps. First, the passive elements are re-tuned to recenter the response on the 4–6 GHz band and to compensate the gap with the extracted circuit; this configuration is called "recalibrated considering the post-layout". Second, with the passives fixed, the bias of the three stages is optimized to maximize a figure of merit combining gain, noise and power. Each retained circuit was verified to be unconditionally stable with the μ condition, explained in section 4.1.3. The optimized circuit is finally verified in the time domain and in linearity.

Analytical tool: an ABCD model of the amplifier. Cadence gives the performance of the circuit, but it does not say which element is responsible for which metric, and a full Spectre run for every trial is slow. To get this information, I developed an analytical model of the complete amplifier in Python, based on the ABCD (chain) matrix formalism.

A linear two-port is described by its chain matrix, which relates the input voltage and current to the output ones,

$$\begin{pmatrix} V_1 \\ I_1 \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} V_2 \\ -I_2 \end{pmatrix}, \tag{4.5}$$

and the matrix of a cascade of two-ports is the product of the individual matrices, taken in the order of the signal path. The amplifier is therefore cut into elementary blocks: a series impedance Z (coupling capacitor, gate inductor, inductor series resistance) and a shunt admittance Y (bias network, drain load seen at a node) have the matrices

$$M_{\text{series}} = \begin{pmatrix} 1 & Z \\ 0 & 1 \end{pmatrix}, \qquad M_{\text{shunt}} = \begin{pmatrix} 1 & 0 \\ Y & 1 \end{pmatrix}. \tag{4.6}$$

The transistors are described by their admittance parameters, taken from the charge-based formulation of Enz [10],

$$Y_{11} = \frac{\omega C_{gg}}{1+\omega R_g C_{gg}}, \qquad Y_{12} = \frac{-\omega C_{gd}}{1+\omega R_g C_{gg}}, \tag{4.7}$$

$$Y_{21} = \frac{g_{m,\text{eff}} - \omega(C_{gd}+C_m)}{1+\omega R_g C_{gg}}, \qquad Y_{22} = \frac{g_{ds,\text{eff}} + \omega^2 R_g C_{gd} C_m + \omega(C_{gd}+C_{db})}{1+\omega R_g C_{gg}}, \tag{4.8}$$

where the source degeneration is included directly in the effective transconductance and output conductance,

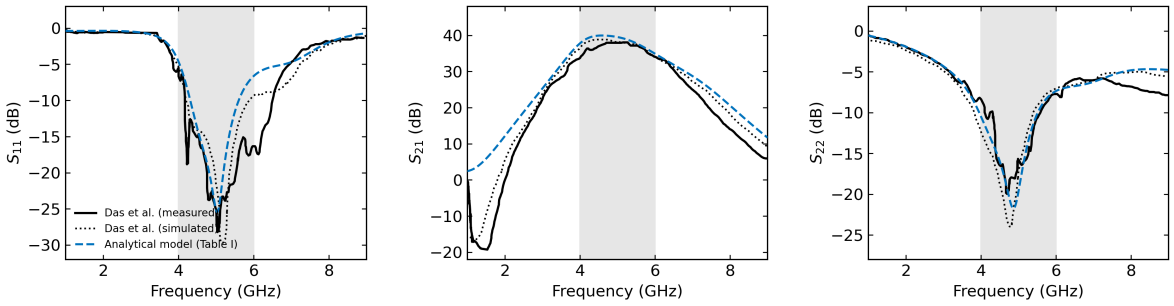
$$g_{m,\text{eff}} = \frac{g_m}{1+G_{ms}R_s}, \qquad g_{ds,\text{eff}} = \frac{g_{ds}}{1+G_{ms}R_s}. \tag{4.9}$$

Each transistor block is then converted from admittance to chain form with the usual relations

$$A = -\frac{Y_{22}}{Y_{21}}, \quad B = -\frac{1}{Y_{21}}, \quad C = -\frac{\det Y}{Y_{21}}, \quad D = -\frac{Y_{11}}{Y_{21}}, \tag{4.10}$$

$$\underbrace{\texttt{dcOpInfo}}_{g_m,\, g_{ds},\, C_{ij}} \longrightarrow Y_{\text{transistor}} \xrightarrow{Y\to Z\to Y} Y_{\text{transistor}+L_s} \longrightarrow T_{ABCD} \xrightarrow{\prod_k T_k} T_{\text{LNA}} \longrightarrow S(f). \tag{4.11}$$

The chain matrix of the amplifier is the product of all the block matrices, and the S-parameters follow from the standard conversion in a $Z_0 = 50\,\Omega$ environment. Each of the three representations is used where it is the most natural: admittance (Y) for the transistor, whose four Y_{ij} are read directly from the operating-point quantities (g_m, g_{ds}, C_{ij}) because its small-signal equivalent circuit is a set of elements connected between its terminals; transmission (T) for the assembly, because cascaded blocks multiply in the order of the signal path; scattering (S) for the result, because this is what is measured at RF and what the article publishes, so the comparison is direct. The full algebra is given in Appendix C. This simulation corresponds well to the one done by the lab to validate their measurements, as shown in figure 4.5. The small-signal parameters feeding these expressions are not fitted: they are read from the operating point of the simulation. As a first validation, the model was run with the parameters of Table I of [1], so that it can be compared directly with the curves published by the authors. The result is shown in figure 4.5: the agreement is very good over the whole sweep and excellent in the 4–6 GHz band, on the three S-parameters. The residual differences outside the band come from the approximations kept in the admittance expressions above, and could be reduced by using a less approximated form.



(a) S_{11} **(b)** S_{21} **(c)** S_{22}

Figure 4.5: Analytical ABCD model evaluated with the small-signal parameters of Table I of [1], compared with the measured and simulated curves of the article.

Once validated, the model evaluates in milliseconds instead of a full Spectre run, so the effect of a modification is tested immediately.Its main use is a sensitivity analysis, in which a sensitivity analysis, in which each passive element is varied by 15 % and the resulting displacement of the characteristic frequencies and change of the S-parameter levels are recorded, the dominant results being collected in table 4.3.

	Gain peak		Input match		Output match	
Element	Δf (MHz)	ΔS_{21} (dB)	Δf (MHz)	ΔS_{11} (dB)	Δf (MHz)	ΔS_{22} (dB)
$L_{s,M1}$	+32	−0.69	+152	+1.23	+12	+0.29
$L_{s,M2}$	+60	−0.56	−32	−1.43	−28	+0.20
$L_{g,in}$	−36	+0.25	−252	+0.53	−12	−0.53
L_{d1}	−72	+0.60	−156	+0.17	+0	−0.91
L_{d2}	−64	+0.20	+28	+0.06	+52	+0.88
L_{d3}	−64	+0.36	−24	−0.06	−224	+0.78
L_{s23}	−44	+0.45	−64	−0.40	−192	+1.36
R_{d2}	−32	+0.39	+0	+0.35	−20	+0.68
R_{d3}	+8	+0.36	+0	−0.16	−12	−2.97
C_{out}	−8	+0.03	+4	+0.06	−8	−3.19

Table 4.3: Sensitivity of the main metrics to a 15 % variation of each passive element. Only the elements with a significant effect are listed; the gate bias resistors and the supply filtering elements were found to be inert on all six metrics.

The table shows a clear division of labour between the elements, and this is what makes a guided manual tuning possible. The gate inductor $L_{g,in}$ is by far the strongest control of the input-match frequency, −252 MHz for a 15 % change, while it barely moves the level. The degeneration inductances $L_{s,M1}$ and $L_{s,M2}$ act mainly on the depth of the S_{11} dip, +1.23 and −1.43 dB, and much less on the gain. The three drain inductances each move the gain peak by roughly the same amount, 60 to 70 MHz, which is exactly what is needed to detune the three stage resonances independently. The output match is controlled by a separate group: C_{out} and R_{d3} set its level, −3.19 and −2.97 dB, while L_{d3} and L_{s23} set its frequency, −224 and −192 MHz. The bias resistors and the supply filtering elements have no measurable effect on any metric and were left at their published values.

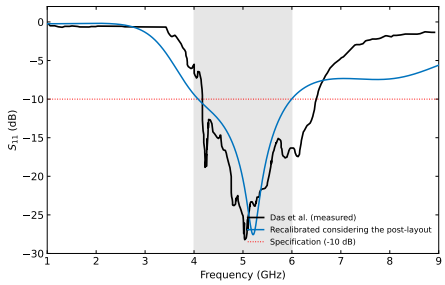
First optimization: recalibration of the passive network. The goal of this step is to bring the simulated response back onto the band, so that it matches the article better and so that the gap coming from the layout extraction is compensated. Table 4.2 shows that S_{11} is the metric that diverges the most, and it is also the one that drives the noise, since the input match of the first stage sets both the reflection and the noise match. It was therefore treated first.

The chain matrix gives the input impedance directly. Writing its definition explicitly, $V_1 = AV_2 + B(-I_2)$ and $I_1 = CV_2 + D(-I_2)$; when the amplifier is closed on a load, the output quantities are related by $V_2 = Z_L(-I_2)$. Substituting the load relation into the two equations and taking their ratio, the output load is transformed to the input through the four chain coefficients,

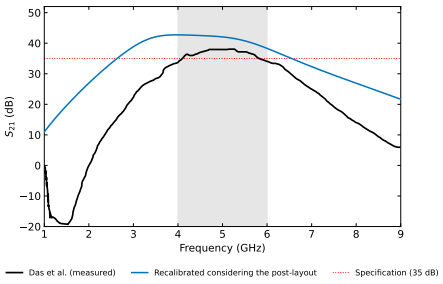
$$Z_{in} = \frac{V_1}{I_1} = \frac{AZ_L + B}{CZ_L + D}, \tag{4.12}$$

so with $Z_L = 50\ \Omega$ the input impedance is obtained from the four coefficients of the total chain matrix at each frequency. The tuning of the input network was then done in Python by minimizing the distance between $Z_{in}(f)$ and 50 Ω over the 4–6 GHz band, using $L_{g,in}$ to move the resonance and $L_{s,M1}$ to set its depth, as indicated by table 4.3. The inter-stage and output networks were adjusted next, the drain inductances being slightly staggered around 5 GHz so that the three stage responses overlap and give a flat gain over the band instead of a single sharp peak, and C_{out} and R_{d3} being used to recover the output match. Each candidate set of values was then re-simulated in Cadence.

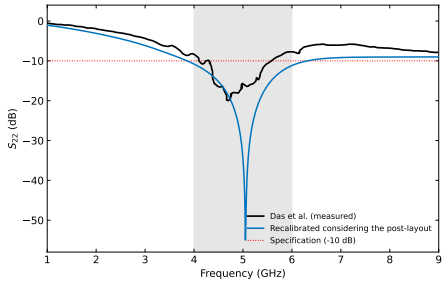
Throughout the design process, the stability factor μ, defined in Section 4.1.3, was continuously monitored as it is strongly affected by passive elements, particularly the degeneration inductance of the first stage. The final component values therefore represent a compromise between achieving a good in-band input match and maintaining sufficient stability margin outside the operating band. The result is shown in Figure 4.6d and summarized in Table 4.4. Recalibration shifts the noise minimum from 4.2 to 5.2 GHz, reducing the band-averaged noise temperature by 1.6×, from 6.30 to 3.84 K, while flattening the response from 3.4 to 4.5 K across 4–6 GHz instead of the 2.4 K minimum followed by a steep rise above 5 GHz. This behavior is consistent with staggered stage resonances. The input match improves from −4.47 to −9.38 dB, just above the −10 dB specification at the band edge, which is corrected by the bias optimization in the next section. These improvements require essentially no additional power (15.83 vs. 15.74 mW), increasing the figure of merit from 0.124 to 0.227 through the response shape rather than power consumption. Compared with the measurement, the gain is 4.6 dB higher, as expected from the underestimated g_{ds} identified in Part I, which also contributes to the higher figure of merit through $G_{\min}$. The average noise temperature is 11.8 %, below the measured value, remaining within the ±1.1 K measurement uncertainty and consistent with the absence of layout parasitics in the simulation.



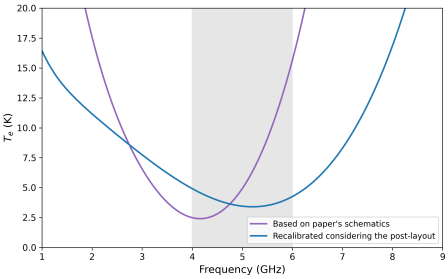
(a) S_{11}



(b) S_{21}



(c) S_{22}



(d) Noise temperature before and after the recalibration of the passive network.

Figure 4.6: S-parameters and noise of the simulated configurations compared with the measurement of [1].

Table 4.4: Effect of the passive recalibration over 4–6 GHz, with the measurement of [1] as reference.

Metric	Schematic-based	Recalibrated	Change	Measured (ref.)
Peak gain S_{21} (dB)	39.19	42.78	+9.2 %	38.06
Minimum gain, 4–6 GHz (dB)	34.13	38.32	+12.3 %	33.73
Max. S_{11}, 4–6 GHz (dB)	−4.47	−9.38	−109.8 %	−5.52
Max. S_{22}, 4–6 GHz (dB)	−9.31	−10.81	−16.1 %	−7.73
Minimum T_e (K)	2.41	3.41	+41.5 %	3.65
Average T_e, 4–6 GHz (K)	6.30	3.84	−39.0 %	4.35
P_{DC} (mW)	15.74	15.83	+0.6 %	15.80
FOM (dB/mW/ph)	0.124	0.227	+83.1 %	0.177

Second optimization: per-stage bias. With the passives fixed, the bias is the last lever: the operating point sets the transconductance and the capacitances of each transistor, hence the S-parameters, and it sets the noise through g_m and through $F(J_D)$. A point is retained if it satisfies the three specifications of Section 4.1 over 4–6 GHz and is unconditionally stable, and among those, the one of lowest power is kept. The figure of merit of Das et al. [1],

$$\mathrm{FOM} = G_{\min} \times \frac{f_h}{f_l} \times \frac{h f_0}{k T_e} \times \frac{1}{P_{DC}}, \tag{4.13}$$

evaluated with $f_0 = 5$ GHz and $f_h/f_l = 1.5$, is used to compare the configurations afterwards, not to select the operating point.

The noise of the recalibrated circuit was decomposed with the Spectre noise summary at 5 GHz. One precaution is essential: the dominant line, PORT0 with 55.9 % of the total output noise, is the reference noise of the source port and not a noise of the amplifier, and must be excluded. Of the amplifier noise, the channel noise of M1 contributes about 86 %, the series resistances of the input inductors about 6 %, and M2 and M3 together 6.5 %. The strategy follows directly. M1 cannot be starved: lowering its current lowers g_m, degrading the noise match, and raises $F(J_{D1})$ through (4.3), so T_e is penalised twice. M2 and M3 carry almost no noise and their current can be cut until the

gain stops it, the only cost being S_{21} and, through $G_{\min}$, the figure of merit. Their minimum current density constraint was accordingly relaxed from 13 to 9 mA/mm.

One-dimensional sweeps of V_{G2} and V_{G3} showed that the in-band S-parameters are not monotonic with the bias and that the constraints become active in different regions of the bias plane, which makes a local optimizer unreliable. A full two-dimensional grid was therefore run at both supply voltages at fixed V_{G1} (Figure 4.7), and every point was screened against the three specifications and against the μ condition of Section 4.1.3 over the whole frequency sweep.

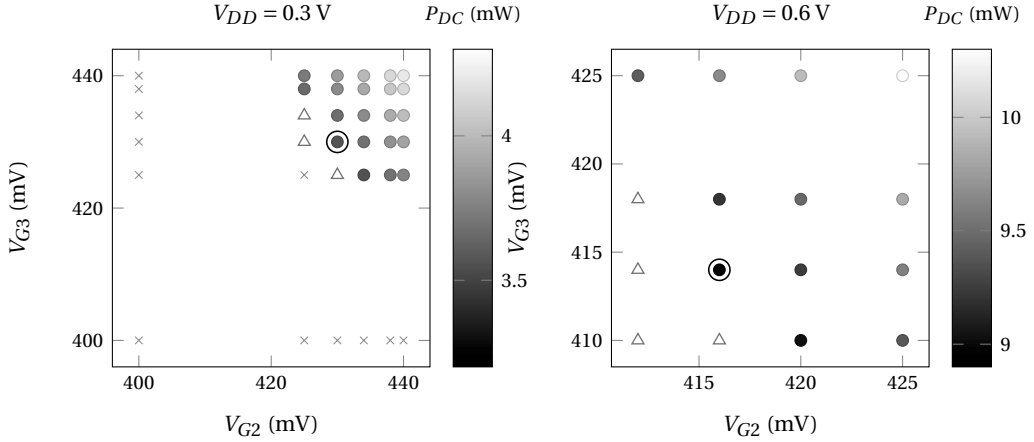


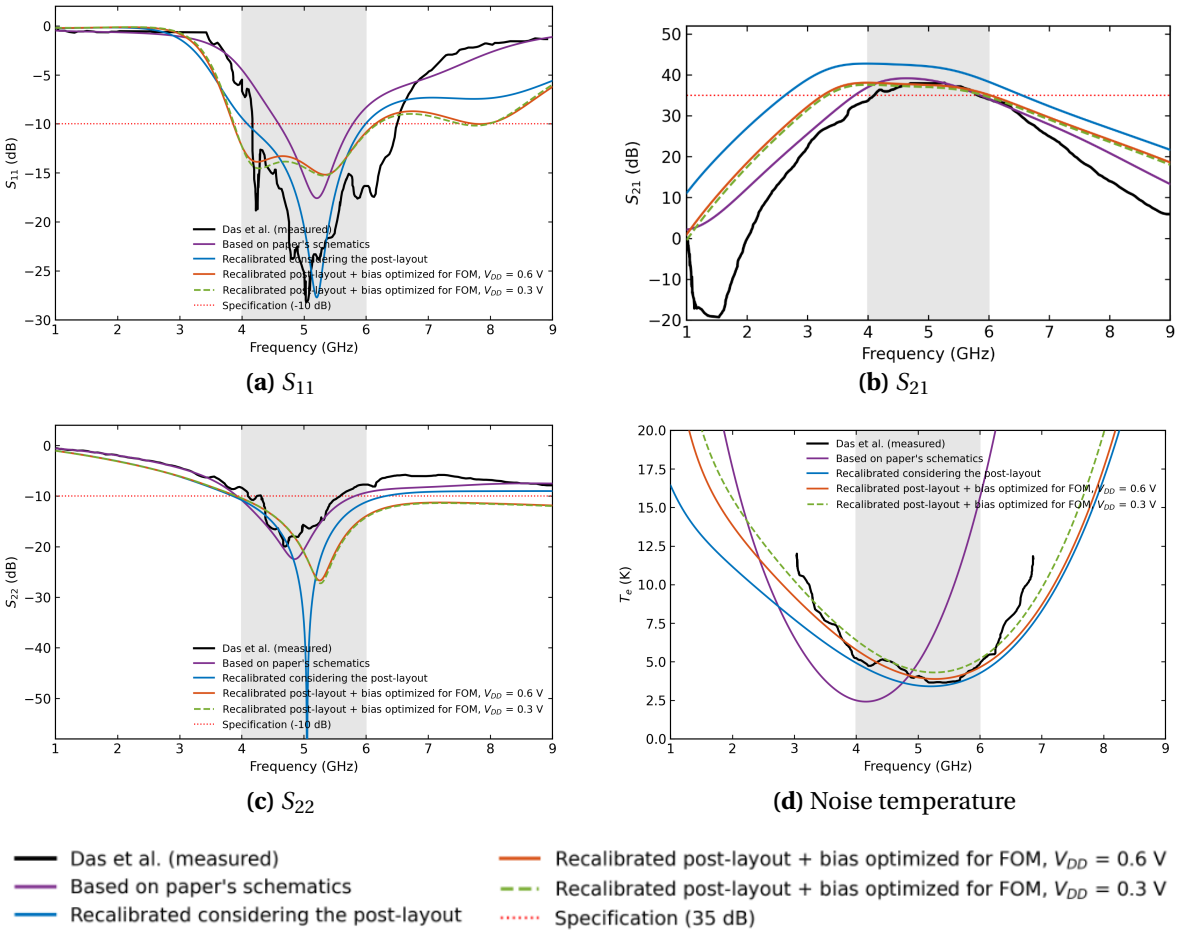
Figure 4.7: Bias grids (V_{G2}, V_{G3}) explored at the two supply voltages, at fixed V_{G1}. Discs satisfy the three specifications and are unconditionally stable, colored by the power consumption; open triangles satisfy the specifications but not the μ condition; crosses violate at least one specification. The circle marks the retained point. Note the different axis ranges of the two panels.

Stability is what limits the power reduction at $V_{DD} = 0.6$ V. All the explored points meet the three specifications, but the lowest-power corner of the grid is not unconditionally stable, and the retained point, $(V_{G2}, V_{G3}) = (416, 414)$ mV, is the first stable point encountered when walking the grid from lowest to highest power. This continues the trade-off of the passive recalibration, where the stability margin already limited how far the input match could be pushed. At $V_{DD} = 0.3$ V the compliant region is much smaller, the specifications become active in their own right, and the retained point is (430, 430) mV; a point with slightly better noise existed in the raw grid but violated S_{11} and was disqualified.

The grid was explored at fixed V_{G1}, which is acceptable at $V_{DD} = 0.6$ V but not at $V_{DD} = 0.3$ V, and this is one of the main findings of this work. Lowering the supply at constant V_{G1} reduces the drain voltage of M1 and its current density collapses from 34.25 to 23.46 mA/mm: g_m drops and the effective Fano factor rises from 0.724 to 0.771, which is exactly the double degradation of the first stage that the strategy above avoids. Raising V_{G1} to 460 mV restores $J_{D1} \approx 27$ mA/mm and improves the noise and the power balance of the low-voltage point at the same time. The bias of the first stage is therefore not transferable from one supply voltage to the other and must be co-optimized with V_{DD}, a conclusion that a single-point transistor model cannot produce.

Performance of the optimized circuit. The S-parameters of the two retained bias points are compared in figure 4.8 with the previous configurations and with the measurement of the article, and the noise temperature of all the configurations is shown in figure 4.8d. At $V_{DD} = 0.6$ V the three specifications are satisfied over the whole band ($G_{\min} = 35.14$ dB, $S_{11} < -10.85$ dB, $S_{22} < -10.45$ dB) and the noise curve is almost superimposed on the measurement of the article, at 54 % of its power. At $V_{DD} = 0.3$ V the two matches remain compliant and the average noise, 4.87 K, stays below the 5.10 K of the published low-power point, for 3.68 mW instead of 4.50 mW. The two optimized curves

keep the flat bottom obtained with the passive recalibration: the bias changes the level of the noise, not its shape, which is consistent with the fact that the bias acts on the devices while the passives set where the amplifier is matched.



(a) S_{11} **(b)** S_{21}

(c) S_{22} **(d)** Noise temperature

Figure 4.8: S-parameters and noise temperature of the simulated configurations compared with the measurement of [1]. The legend applies to the four panels.

The detailed comparison with the article is given in table 4.5.

Table 4.5: Recalibrated post-layout + bias optimized for FOM, compared with Das *et al.* [1] at both supply voltages. The reference is the nominal bias of the article at $V_{DD} = 0.6$ V and its published low-power bias at $V_{DD} = 0.3$ V. All band metrics are taken over 4–6 GHz.

	$V_{DD} = 0.6$ V			$V_{DD} = 0.3$ V		
Metric	Das *et al.*	This work	Diff.	Das *et al.*	This work	Diff.
Peak gain S_{21} (dB)	38.06	38.10	+0.1 %	34.00	37.62	+10.6 %
Minimum gain (dB)	33.73	35.14	+4.2 %	33.00	34.60	+4.8 %
Max. S_{11} (dB)	−5.52	−10.85	−96.6 %	–	−10.99	–
Max. S_{22} (dB)	−7.73	−10.45	−35.2 %	–	−10.43	–
Minimum T_e (K)	3.65	3.88	+6.3 %	–	4.31	–
Average T_e (K)	4.35	4.39	+1.0 %	5.10	4.87	−4.5 %
P_{DC} (mW)	15.80	8.46	−46.4 %	4.50	3.68	−18.2 %
FOM (dB/mW/ph)	0.177	0.340	+92.5 %	0.518	0.695	+34.3 %

Finally, table 4.6 collects the figure of merit of every configuration of this study together with the two bias points of the article, evaluated with equation (4.13).

Table 4.6: Figure of merit of all the configurations, computed with equation (4.13) and the $G_{\min}$ convention used throughout this work. With the gain convention of [1], the published values for the two bias points of the article are 0.192 and 0.517.

Configuration	$G_{\min}$ (dB)	$\overline{T_e}$ (K)	P_{DC} (mW)	FOM (dB/mW/ph)
Das *et al.*, VDD=0.6V	33.73	4.35	15.80	0.177
Das *et al.*, VDD=0.3V	33.00	5.10	4.50	0.518
Based on the paper's schematic	34.13	6.30	15.74	0.124
Recalibrated considering the post-layout	38.32	3.84	15.83	0.227
+ bias optimized, $V_{DD} = 0.6$ V	35.14	4.39	8.46	0.340
+ bias optimized, $V_{DD} = 0.3$ V	34.60	4.87	3.68	0.695

A precaution must be stated about these comparisons. The final points are simulations without layout extraction, compared with measurements of a fabricated chip, so the absolute values are optimistic and the gain in particular is systematically high. The clean comparison is the internal one, between the recalibrated circuit and the bias-optimized circuit, since both are simulated under exactly the same conditions: on that basis, the per-stage biasing enabled by the compact model buys a 47 % power reduction for 0.55 K of noise.

4.1.3 Part III: verification of the optimized circuit

Stability of the optimized circuit. An amplifier is unconditionally stable when no combination of passive source and load impedances can make it oscillate. For a two-port described by its S-parameters, this is tested with the Edwards–Sinsky factor

$$\mu = \frac{1 - |S_{11}|^2}{\left|S_{22} - \Delta\, S_{11}^*\right| + |S_{21} S_{12}|}, \qquad \Delta = S_{11} S_{22} - S_{12} S_{21}, \tag{4.14}$$

and its dual μ', obtained by exchanging the two ports. The condition is that the amplifier is unconditionally stable if and only if $\mu > 1$, equivalently $\mu' > 1$. The value itself is meaningful: μ is the distance, on the Smith chart, between the centre (the matched 50 Ω load) and the closest load impedance that would make the input reflection larger than one. It is therefore a margin and not only a pass or fail test, which is why it was preferred to the older K factor, which needs a second auxiliary condition and gives no margin.

The two factors were computed from the S-parameters of section 4.1.2, at both retained bias points, with the inductor series resistances included (figure B.2 of Appendix B). They stay above unity over the whole sweep, so the optimized circuit is unconditionally stable at both supply voltages. The minimum values, $\mu'_{\min} = 1.0042$ at $V_{DD} = 0.6$ V and 1.0012 at $V_{DD} = 0.3$ V, occur at low frequency, where the inductances behave as short circuits and the damping comes almost only from their series resistances. The margin is thin there but positive everywhere.

Time-domain verification. The μ condition is computed from the S-parameters seen at the two external ports, so it cannot detect an oscillation that would build up between the stages without reaching the accesses. A complementary verification was therefore done in the time domain: the circuit is perturbed with different signal levels and one checks that no oscillation starts between the stages. The perturbation is a current impulse injected on the drain of M1, chosen because M1 is the device with the largest noise contribution and the node the most likely to excite a strong oscillation if one exists.

A loop-gain study, with a Bode plot and a phase margin, was not possible here, because this amplifier has no global feedback loop that could be opened to define a loop gain: the instability risk comes from distributed reactive couplings between the stages and not from one identifiable loop.

The kick test is the standard alternative.

Three settings are necessary for the test to be conclusive. The integration method must be `traponly`, since the default Gear-2 integrator adds an artificial numerical damping that can hide a marginal oscillation. The time step must resolve the gigahertz dynamics, so a maximum step and a strobe period of a few picoseconds were used. The ideal bias tees used for the S-parameter analyses were replaced by realistic values, since their extreme component values prevent the transient from converging.

The result is shown in figure B.3. For impulse amplitudes of 10 µA and 1 mA, the output response is an exponentially damped ring-down with a time constant $\tau \approx 0.24$–0.26 ns, identical at both bias points, and the two responses superpose exactly after scaling by the amplitude ratio, which confirms that the circuit responds linearly to the perturbation. The spectrum of the response contains no parasitic line. The conclusion of the μ analysis is therefore confirmed in the time domain.

A standard transient simulation with a sine at the input was also run from the input to the output of the amplifier; the output is the expected amplified sine, without any anomaly, so this test is not reported in a figure. The propagation time through the amplifier was not studied, because no reference data was available to compare it with. The result is shown in figure B.3 of Appendix B.

Linearity of the optimized circuit. Reducing the drain currents to save power carries a risk on linearity, since the compression point of an amplifier decreases with its bias current. The article reports an input 1-dB compression point $P_{1\mathrm{dB,in}} > -40$ dBm at the nominal bias. The question is whether the optimized, lower-current bias points still provide enough linearity for qubit readout.

The compression was characterized with a single-tone transient simulation at 5 GHz: the input power is swept, the fundamental at the output is extracted with a coherent FFT (rectangular window over an integer number of signal periods, uniform strobe sampling), and the gain is plotted against the input power. The small-signal gain obtained at low power agrees with the S-parameter analysis, which cross-validates the two setups. The results are shown in figure 4.9: $G_0 = 38.31$ dB and $P_{1\mathrm{dB,in}} = -49.2$ dBm at $V_{DD} = 0.6$ V, and $G_0 = 37.73$ dB and $P_{1\mathrm{dB,in}} = -48.9$ dBm at $V_{DD} = 0.3$ V. The two compression points are almost identical, so halving the power consumption did not cost linearity in any significant way.

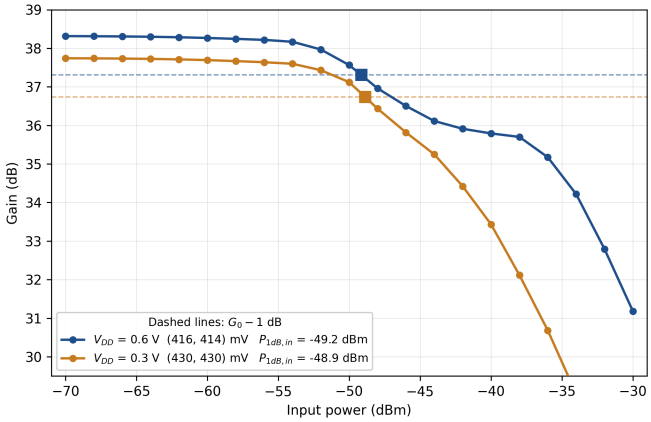


Figure 4.9: Gain compression at 5 GHz extracted from transient simulation with coherent FFT, at the two optimized bias points. The dashed lines mark $G_0 - 1$ dB.

To judge whether these values are sufficient, the relevant criterion is the one discussed by Zeng for readout LNAs [28]. Modern qubit readout is frequency-multiplexed: N tones are amplified simultaneously, so the total power at the amplifier input grows as $10\log_{10} N$ above the power of a

single tone, and the requirement becomes

$$P_{1\mathrm{dB,in}} > P_{\mathrm{tone}} + 10\log_{10} N. \tag{4.15}$$

Two hypotheses are needed and are stated explicitly here: a power per tone $P_{\mathrm{tone}} = -100$ dBm at the LNA input, a typical value for dispersive readout, and a deliberately large multiplexing factor of $N = 1000$ tones. The requirement is then $-100 + 10\log_{10}(1000) = -70$ dBm, and the worst compression point, -49.2 dBm, satisfies it with a margin of 20.8 dB. A multi-tone simulation was therefore judged unnecessary.

4.1.4 Conclusion

Four configurations were characterised under strictly identical simulation conditions, which is what makes their comparison meaningful: the two model deviations identified in Part I and the absence of layout extraction are common to all four and cancel out.

Recalibrating the passive network, using the sensitivities obtained from the analytical model, raises the figure of merit from 0.124 to 0.227 at constant power. Optimising the bias of the three stages then gives 8.46 mW for an average noise temperature of 4.39 K at $V_{DD} = 0.6$ V, that is 46% less power than the nominal point of the article for the same noise, and 3.68 mW for 4.87 K at $V_{DD} = 0.3$ V, where V_{G1} has to be re-optimised jointly with the supply. Both optimised circuits are unconditionally stable and keep an input compression point near -49 dBm, which leaves 20.8 dB of margin for a 1000-tone readout.

The result that holds is the relative one: once a compact model with a bias-dependent noise law is available, a per-stage bias optimisation delivers the same noise temperature for roughly half the power.

5 Conclusion and perspectives

5.1 Conclusion

At the beginning of this project the Rhonexum cryogenic compact model existed in DC only. The AC part of the model was developed during this period by a colleague of the laboratory, and the extraction flow that produces the model cards from the company measurement data was built by another colleague, who was also responsible for the cryogenic extraction itself and confirmed its validity at cryogenic temperature. My own contribution sits around that model and is of four kinds.

The first is the verification of the model at room temperature, in DC and in AC, against the GF22FDX PDK. In DC the model follows the PDK closely on a long channel and with bounded deviations at 20 and 40 nm, and on a published five-transistor OTA of the same technology it predicts a gain within 7 % of the measured silicon, closer to it than the foundry nominal simulation. In AC, once a small number of parameters were adjusted by hand, the transcapacitances of the model are consistent with those of the PDK, the residual differences being attributable to different charge-partitioning conventions and to the absence of an extraction dedicated to the AC regime.

The second is the device-level characterisation that this verification required. A model of the gate resistance was established for the sEKV, taken from the decomposition available in the literature and calibrated on the device with an automated extraction flow driving Cadence from Python. It reproduces the extraction to 2.2 % RMS over three geometry sweeps once a width-set access term is added, and the fitted vertical resistivity, falls in the range expected for a metal-gate stack. A measured sensitivity study then justified reducing the expression to a single term, implemented in one line of Verilog-A, at a cost below 0.3 % on the extracted capacitances. A testbench was also developed to observe the intrinsic transcapacitances of a transistor from standard AC simulations, without any access to the internals of the compact model. It exploits the different geometric scaling of the intrinsic and extrinsic charges, is solved with three geometries, and reproduces the internal partitioning of the encrypted GF22FDX model on the four front-gate transcapacitances at $V_{DS} = 0$ V.

The third is a Python tool, `gen_skillfinit.py`, which automates the creation of a Cadence library from a `.lib` file. It parses the file and writes, for every subcircuit it contains, a SKILL script that builds the library, the cell, its symbol and its CDF, and a Spectre netlist that declares the Verilog-A module and reproduces the binning rules. Both files come from the same parse, so the terminal order, the parameter names and the model names cannot disagree. For the company this removes the manual step between an extraction campaign and a usable design kit: any `.lib` file produced by a future extraction becomes a complete Cadence cryogenic library in one command, which is the practical prerequisite for distributing such libraries as a commercial product. The tool covers integration only. It does not inspect the Verilog-A source and does not check that a model card is physically meaningful.

The fourth, and the main deliverable for the company, is the first validation of the model on a complete circuit at cryogenic temperature. The 4 to 6 GHz LNA of Das et al. [1] was simulated at 4.4 K with the sEKV, after adding to the Verilog-A description a channel noise written as suppressed shot noise with a bias-dependent Fano factor. The transport and noise quantities are reproduced within a few percent of the published measurement. The AC quantities, C_{gs} and g_{ds}, deviate more and displace the resonance frequencies and the gain level of the circuit, but the simulated response keeps

the shape of the measured one, which is what makes the model usable at cryogenic temperature for this class of application. The simulations carry no layout extraction. On that basis, recalibrating the passive network raised the figure of merit from 0.124 to 0.227 at constant power, and optimising the bias of the three stages independently, which the original methodology could not do, gave 8.46 mW for an average noise temperature of 4.39 K, that is 46 % less power than the published nominal point for a comparable noise. The optimised circuits are unconditionally stable and keep an input compression point near −49 dBm. The model is therefore usable today as a cryogenic design tool, with the reservation that its AC parameters set the accuracy of the resonance frequencies and of the gain level.

Perspectives

The most immediate item is a new extraction campaign of the sEKV parameters dedicated to the AC regime. The AC behaviour currently rests on parameters that were partly adjusted by hand, and the same two quantities, C_{gs} and g_{ds}, explain the amplitude offset seen on $C_{fg,fg}$ in strong inversion, the gain level of the OTA and the frequency shift of the LNA, so no result of this report can be tightened before they are re-extracted. The transcapacitance testbench extends naturally to a fourth geometry, which would separate the source-side and drain-side extrinsic coefficients and restore the validity of the extraction at $V_{DS} \neq 0$. A second direction is optimisation by sizing. The transistor geometries of the LNA were kept at the values published in the article, and only the passive network and the bias were used as degrees of freedom. Since the sEKV is continuous in bias and scalable in W, L and n_f, the width and the finger count of each stage can be optimised jointly with the operating point, which should be particularly effective on the input stage, where the noise match, g_m and the gate resistance all depend on the geometry and which carries about 86 % of the amplifier noise. Beyond that, the LNA study should be carried to layout and post-layout extraction, which is the only way to turn the relative conclusions of this work into absolute ones, the generator should be extended to the remaining device flavours and to the PMOS versions, and a corner and statistical description of the cryogenic model is still missing, which currently prevents any variability analysis at 4 K.

A Monte-Carlo analysis of the OTA DC gain

This appendix gives the material supporting the variability argument of Section 3.1.2: the distribution of the OTA DC gain over process variability, and the two reservations attached to the comparison with silicon.

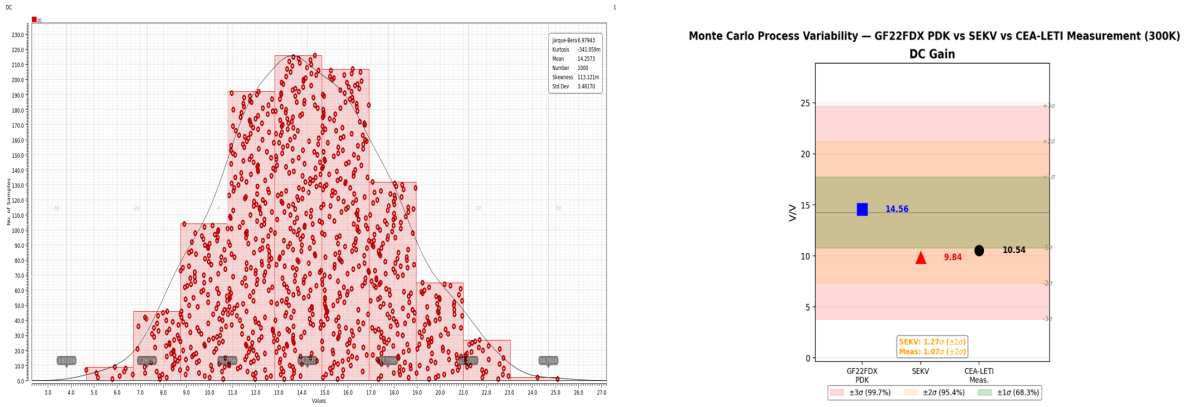


Figure A.1: Left: Monte-Carlo distribution of the OTA DC gain simulated with the GF22FDX PDK at room temperature (1000 samples, $\mu = 14.26$ V/V, $\sigma = 3.48$ V/V). Right: the three DC gain values of Table 3.1 placed on the $\pm1\sigma$, $\pm2\sigma$ and $\pm3\sigma$ bands of that distribution.

The interpretation is that the gain spread produced by process variability alone is of the same order as the gap between the sEKV and the PDK nominal value. A die drawn from this technology can plausibly deliver 9.8 V/V as well as 14.6 V/V. The measured CEA-Leti device, sitting at 1.07σ, is itself one such sample. The sEKV prediction is not an outlier of the distribution but a plausible member of it, positioned close to the measured point.

A second reservation concerns the nature of the terms compared. The value of Martinez *et al.* is measured on a fabricated die and therefore contains the layout parasitics, whereas the two simulations are pre-layout and contain none. For a DC gain the consequence is bounded: the gain is read at low frequency, so the parasitic capacitances play no role, and only the series resistances of the source and drain accesses act, by degenerating g_m. Their effect on $g_m/(g_{ds,n} + g_{ds,p})$ stays second order at the bias current used here. The direction is nevertheless unfavourable to the model: including the parasitics would lower both simulated gains, moving the PDK value towards the measurement and the sEKV value away from it. The comparison free of this bias is the one between the two simulations, which are run pre-layout under identical conditions.

B Supporting figures

This appendix collects the figures that support results already quantified in the main text: the reference curves used to set the expected shape of the intrinsic transcapacitances, and the two verifications of the optimized LNA whose numerical outcomes are given in Sections 4.1.3 and 4.1.3.

B.1 Intrinsic transcapacitances reported in the literature

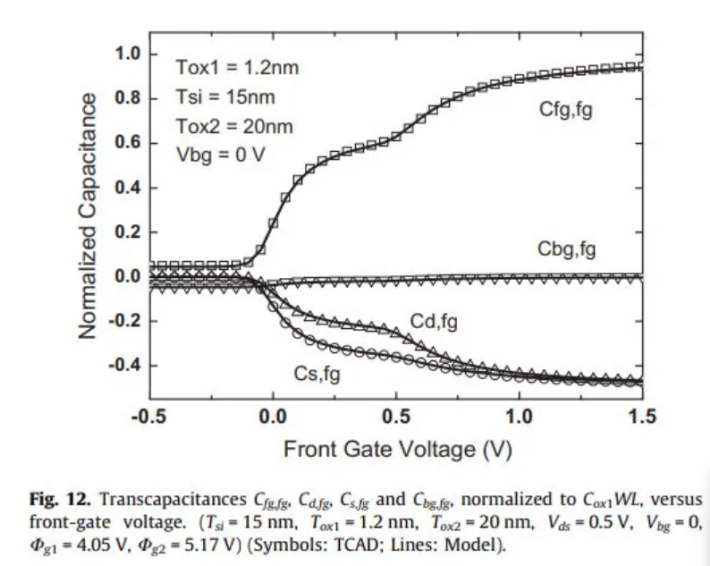


Fig. 12. Transcapacitances $C_{fg,fg}$, $C_{d,fg}$, $C_{s,fg}$ and $C_{bg,fg}$, normalized to $C_{ox1}WL$, versus front-gate voltage. (T_{si} = 15 nm, T_{ox1} = 1.2 nm, T_{ox2} = 20 nm, V_{ds} = 0.5 V, V_{bg} = 0, Φ_{g1} = 4.05 V, Φ_{g2} = 5.17 V) (Symbols: TCAD; Lines: Model).

Figure B.1: Intrinsic transcapacitances $C_{fg,fg}$, $C_{d,fg}$, $C_{s,fg}$ and $C_{bg,fg}$ normalised to $C_{ox1}WL$ versus front-gate voltage, from Lu et al. [16] ($T_{ox1} = 1.2\,\text{nm}$, $T_{si} = 15\,\text{nm}$, $T_{ox2} = 20\,\text{nm}$, $V_{ds} = 0.5\,\text{V}$, $V_{bg} = 0$). The curves are obtained from TCAD on an idealised device and are used in Section 3.2.1 only to establish the expected shape and sign of each transcapacitance.

B.2 Stability and time-domain verification of the optimized LNA

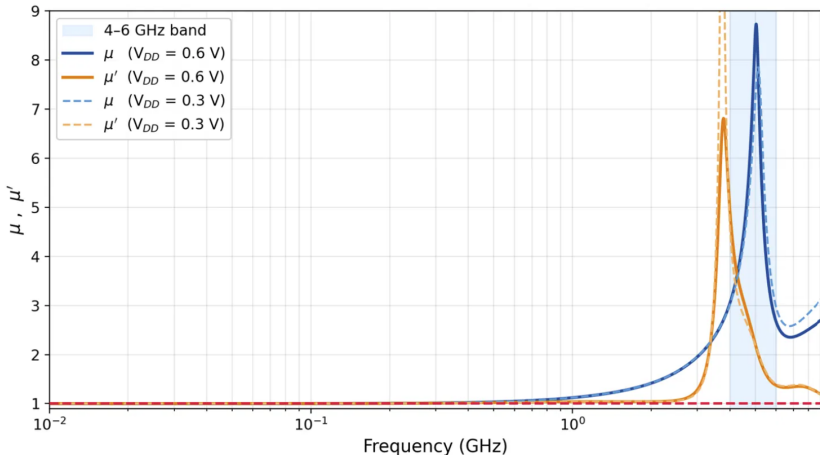


Figure B.2: Edwards–Sinsky stability factors μ and μ' of the optimized circuit at both supply voltages, computed from the S-parameters of Section 4.1.2 with the inductor series resistances included. The dashed line marks the unconditional stability limit $\mu = 1$. Both factors stay above unity over the whole sweep, the minima being $\mu'_{\min} = 1.0042$ at $V_{DD} = 0.6$ V and 1.0012 at $V_{DD} = 0.3$ V.

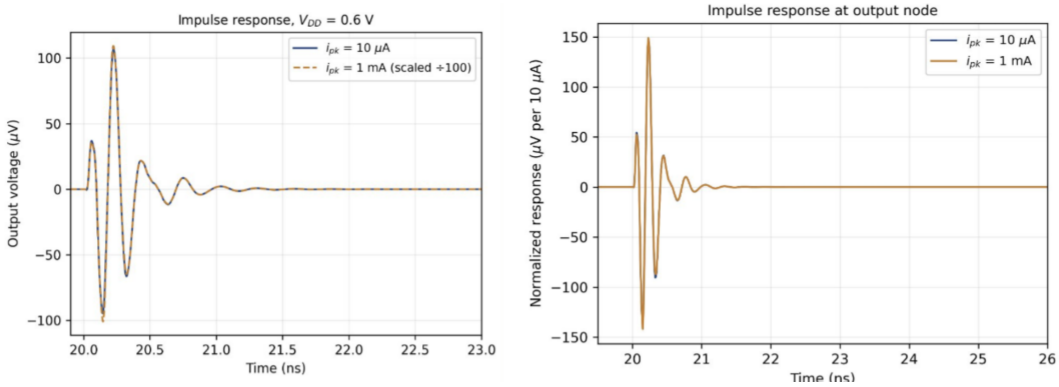


Figure B.3: Kick test: response at the output node to a current impulse injected on the drain of M1, for two impulse amplitudes. Left: raw response at $V_{DD} = 0.6$ V. Right: the two responses normalised by the impulse amplitude, which superpose and confirm that the circuit answers linearly to the perturbation.

C Derivation of the ABCD model

This appendix gives the algebra behind the model of Sec. 4.1.2. The transistor admittance parameters are taken as given from Chalkiadaki [29]; everything else – the embedding of the degeneration inductor, the conversion to chain form, the assembly and the conversion to S – is derived here. The model is small-signal and noiseless: it predicts $S(f)$ only, the noise temperature comes from Spectre.

Conventions. Port currents are counted entering the network. The admittance, impedance and chain descriptions are

$$\begin{pmatrix} I_1 \\ I_2 \end{pmatrix} = Y \begin{pmatrix} V_1 \\ V_2 \end{pmatrix}, \qquad Z = Y^{-1} = \frac{1}{\Delta_Y} \begin{pmatrix} Y_{22} & -Y_{12} \\ -Y_{21} & Y_{11} \end{pmatrix}, \qquad \begin{pmatrix} V_1 \\ I_1 \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} V_2 \\ I_2' \end{pmatrix}, \tag{C.1}$$

with $\Delta_Y \equiv \det Y = Y_{11}Y_{22} - Y_{12}Y_{21}$ and $I_2' \equiv -I_2$, the output current counted leaving the network. The scattering description is defined in $Z_0 = 50\ \Omega$ by $a_i = (V_i + Z_0 I_i)/2\sqrt{Z_0}$ and $b_i = (V_i - Z_0 I_i)/2\sqrt{Z_0}$. Each representation is used where it is natural: Y for the transistor, T for the assembly, S for the result.

Elementary blocks and cascade rule. A series impedance carries the same current at both ports ($I_1 = I_2'$) and drops $Z I_2'$ across itself ($V_1 = V_2 + Z I_2'$); a shunt admittance keeps both ports at the same potential ($V_1 = V_2$) and diverts $Y V_2$ from the input current ($I_1 = Y V_2 + I_2'$). Identifying with Eq. (C.1),

$$T_{\text{series}}(Z) = \begin{pmatrix} 1 & Z \\ 0 & 1 \end{pmatrix}, \qquad T_{\text{shunt}}(Y) = \begin{pmatrix} 1 & 0 \\ Y & 1 \end{pmatrix}. \tag{C.2}$$

The $I_2' = -I_2$ convention makes the output pair of one block equal to the input pair of the next, $(V_2^{(1)}, I_2'^{(1)}) = (V_1^{(2)}, I_1^{(2)})$, so that

$$\begin{pmatrix} V_1 \\ I_1 \end{pmatrix} = T_1 T_2 \cdots T_N \begin{pmatrix} V_2^{(N)} \\ I_2'^{(N)} \end{pmatrix}, \qquad T_{\text{total}} = \prod_k T_k, \tag{C.3}$$

the matrices being multiplied in the order of the signal path. No other representation has this property, which is the reason the amplifier is assembled in chain form.

Transistor two-port. The intrinsic device plus its distributed gate resistance R_g is described by the charge-based small-signal model of Chalkiadaki [29],

$$Y = \frac{1}{1 + \mathrm{j}\omega R_g C_{gg}} \begin{pmatrix} \mathrm{j}\omega C_{gg} & -\mathrm{j}\omega C_{gd} \\ g_m - \mathrm{j}\omega(C_{gd} + C_m) & g_{ds} + \omega^2 R_g C_{gd} C_m + \mathrm{j}\omega(C_{gd} + C_{db}) \end{pmatrix}, \tag{C.4}$$

with $C_{gg} = C_{gs} + C_{gd} + C_{gb}$ the total gate capacitance and $C_m = C_{dg} - C_{gd}$ the transcapacitance, which measures the non-reciprocity of the capacitive part. The common denominator $1 + \mathrm{j}\omega R_g C_{gg}$ is the pole introduced by R_g charging C_{gg}; expanding Eq. (C.4) to first order in $\omega R_g C_{gg}$ returns the expressions printed in [29].

Every parameter of Eq. (C.4) is read from the Spectre `dcOpInfo` record at the operating point of interest (Table C.1), so the model inherits the bias and temperature dependence of the sEKV compact model without re-deriving it. This is also why the back-gate does not appear explicitly: its decoupling capacitor is a small-signal ground (see the passive networks below), so the whole effect

of V_{BG} is carried by the operating-point values.

Table C.1: Mapping between the `dcOpInfo` record and the parameters of Eq. (C.4). Off-diagonal capacitances are taken as positive magnitudes, $C_{gd} = -\partial Q_G/\partial V_D$ and $C_{dg} = -\partial Q_D/\partial V_G$.

Symbol	`dcOpInfo`	Role
g_m	`gm`	transconductance
g_{ds}	`gds`	output conductance
C_{gg}	`cgg`	total gate capacitance
C_{gd}	`|cgd|`	feedback capacitance
C_{dg}	`|cdg|`	gives $C_m = C_{dg} - C_{gd}$
C_{db}	`|cdb|`	drain–bulk capacitance
R_g	layout (480 Ω·μm / W)	gate access resistance

The transit delay of Table I of the article is not retained: with $\tau = 100$ fs, $g_m \to g_m e^{-\mathrm{j}\omega\tau}$ adds a phase $\omega\tau = 3.1 \times 10^{-3}$ rad, i.e. 0.18° at 5 GHz.

Embedding the source degeneration. The degeneration impedance $Z_s = R_s + \mathrm{j}\omega L_s$ sits in the *common* lead, so the voltage it develops, $Z_s(I_1 + I_2)$, is added to V_1 and V_2 alike:

$$Z' = Y^{-1} + Z_s U, \qquad U = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \qquad Y' = \left[Y^{-1} + Z_s U \right]^{-1}. \tag{C.5}$$

Writing Z' entry by entry with Eq. (C.1) and taking its determinant, the Z_s^2 terms cancel because U is singular:

$$\det Z' = \frac{Y_{11}Y_{22} - Y_{12}Y_{21}}{\Delta_Y^2} + \frac{Z_s\,(Y_{11} + Y_{12} + Y_{21} + Y_{22})}{\Delta_Y} = \frac{1 + Z_s \Sigma_Y}{\Delta_Y}, \qquad \Sigma_Y \equiv \textstyle\sum_{ij} Y_{ij}. \tag{C.6}$$

Inverting Z' gives the closed form evaluated by the code,

$$\boxed{Y' = \frac{1}{1 + Z_s \Sigma_Y} \begin{pmatrix} Y_{11} + Z_s\Delta_Y & Y_{12} - Z_s\Delta_Y \\ Y_{21} - Z_s\Delta_Y & Y_{22} + Z_s\Delta_Y \end{pmatrix}.} \tag{C.7}$$

Effective transconductance. At low frequency $Y_{11}, Y_{12} \to 0$, $Y_{21} \to g_m$, $Y_{22} \to g_{ds}$, hence $\Delta_Y \to 0$ and $\Sigma_Y \to g_m + g_{ds} \equiv G_{ms}$, so Eq. (C.7) reduces to

$$g_{m,\mathrm{eff}} = \frac{g_m}{1 + G_{ms} R_s}, \qquad g_{ds,\mathrm{eff}} = \frac{g_{ds}}{1 + G_{ms} R_s}, \tag{C.8}$$

the expressions used in Sec. 4.1.2. Series feedback divides both quantities by the same loop factor, leaving the intrinsic voltage gain g_m/g_{ds} unchanged.

Input resistance created by L_s. Keeping only C_{gs} and g_m, with $V_{gs} = I_1/\mathrm{j}\omega C_{gs}$ and a source node at $V_s = Z_s(I_1 + g_m V_{gs})$,

$$Z_{in} = \frac{1}{\mathrm{j}\omega C_{gs}} + \mathrm{j}\omega L_s + \frac{g_m L_s}{C_{gs}}, \qquad \mathrm{Re}\{Z_{in}\} = \omega_T L_s. \tag{C.9}$$

The real part is set by L_s alone and dissipates no dc power, which is why the noise match and the 50 Ω match can be met simultaneously, and why L_{s1} cannot be reduced freely: lowering it weakens $\mathrm{Re}\{Z_{in}\}$ and lets the C_{gd} feedback path dominate.

From Y to the chain matrix. Substituting $I_2 = -I_2'$ in $I_2 = Y_{21}V_1 + Y_{22}V_2$ isolates V_1, and inserting the result in $I_1 = Y_{11}V_1 + Y_{12}V_2$ gives I_1:

$$V_1 = -\frac{Y_{22}}{Y_{21}} V_2 - \frac{1}{Y_{21}} I_2', \qquad I_1 = -\frac{\Delta_Y}{Y_{21}} V_2 - \frac{Y_{11}}{Y_{21}} I_2', \tag{C.10}$$

that is

$$A=-\frac{Y_{22}}{Y_{21}},\quad B=-\frac{1}{Y_{21}},\quad C=-\frac{\Delta_Y}{Y_{21}},\quad D=-\frac{Y_{11}}{Y_{21}},\qquad \text{and}\qquad AD-BC=\frac{Y_{11}Y_{22}-\Delta_Y}{Y_{21}^2}=\frac{Y_{12}}{Y_{21}}, \tag{C.11}$$

the last identity equalling 1 for a reciprocal block and serving as a numerical check.

Passive networks. Inductors are entered with their series loss, $Z_L(L)=\omega L/Q(\omega)+\mathrm{j}\omega L$; this is the term recalibrated on the post-layout extraction in configuration F2. The supply decoupling cell contributes $Z_f(\omega)=\left(1/R_f+\mathrm{j}\omega C_f\right)^{-1}$ in series with the drain inductor, so the three drain loads are

$$Z_{L1}=Z_L(L_{d1})+Z_f(10\,\Omega,15\,\mathrm{pF}),\quad Z_{L2}=\Big[Z_L(L_{d2})\|50\,\Omega\Big]+Z_f(10\,\Omega,10\,\mathrm{pF}),\quad Z_{L3}=\Big[Z_L(L_{d3})\|80\,\Omega\Big]+Z_f(10\,\Omega,10\,\mathrm{pF} \tag{C.12}$$

each entering the cascade as a shunt admittance $Y_{Lk}=1/Z_{Lk}$ at the corresponding drain node. Coupling capacitors are series elements $Z=1/\mathrm{j}\omega C$.

Each gate is biased through 100 kΩ decoupled by 5 pF. At 5 GHz that capacitor presents $|Z_C|=1/2\pi fC=6.4\,\Omega$, negligible in series with 100 kΩ, so the branch is a pure shunt conductance $Y_{Gk}\simeq 10\,\mu$S. The same argument applied to the 10 pF back-gate capacitor gives 3.2 Ω: the back-gate is a small-signal ground and adds no branch to the model.

Assembly. The amplifier is cut into the blocks of Table C.2, taken in the order of the signal path, and

$$T_{\mathrm{LNA}}(f)=\prod_{k=0}^{16}T_k(f). \tag{C.13}$$

The complete $S(f)$ of the three-stage amplifier therefore costs seventeen 2×2 complex products per frequency point, which is what makes the model fast enough to be swept over design parameters.

Table C.2: Decomposition of the LNA into elementary two-ports. Y_{Gk} is the gate bias conductance, $Y_{Lk}=1/Z_{Lk}$ the drain load admittance. Numerical values are those of the netlist; L_{s1} is the degeneration inductor of M_1.

k	Block	Matrix	Element
0	input inductor loss	series	R_{Lg}
1	input gate inductor	series	$\mathrm{j}\omega L_g$
2	gate 1 bias network	shunt	Y_{G1}
3	M_1 with L_{s1}	transistor	Eqs. (C.7), (C.11)
4	stage 1 drain load	shunt	Y_{L1}
5	interstage inductor 1–2	series	$Z_L(L_{12})$
6	coupling capacitor 1–2	series	$1/\mathrm{j}\omega C_{12}$
7	gate 2 bias network	shunt	Y_{G2}
8	M_2	transistor	Eqs. (C.4), (C.11)
9	stage 2 drain load	shunt	Y_{L2}
10	interstage inductor 2–3	series	$Z_L(L_{23})$
11	coupling capacitor 2–3	series	$1/\mathrm{j}\omega C_{23}$
12	gate 3 bias network	shunt	Y_{G3}
13	M_3	transistor	Eqs. (C.4), (C.11)
14	stage 3 drain load	shunt	Y_{L3}
15	output series resistor	series	R_{out}
16	output coupling capacitor	series	$1/\mathrm{j}\omega C_{\mathrm{out}}$

From the chain matrix to the S-parameters.

Port 2 matched ($a_2 = 0$). $a_2 = 0$ means $V_2 = Z_0 I_2'$, i.e. port 2 closed on Z_0. Equation (C.1) then gives $V_1 = (AZ_0 + B) I_2'$ and $I_1 = (CZ_0 + D) I_2'$, hence

$$Z_{in} = \frac{AZ_0 + B}{CZ_0 + D}, \qquad S_{11} = \frac{Z_{in} - Z_0}{Z_{in} + Z_0} = \frac{A + B/Z_0 - CZ_0 - D}{\Delta}, \qquad \Delta \equiv A + \frac{B}{Z_0} + CZ_0 + D, \quad \text{(C.14)}$$

after dividing numerator and denominator by Z_0. The waves are $b_2 = (V_2 + Z_0 I_2')/2\sqrt{Z_0} = \sqrt{Z_0} I_2'$ and $a_1 = (V_1 + Z_0 I_1)/2\sqrt{Z_0} = \sqrt{Z_0}\,\Delta\, I_2'/2$, so $S_{21} = b_2/a_1 = 2/\Delta$.

Port 1 matched ($a_1 = 0$). Now $V_1 = -Z_0 I_1$, which inserted in Eq. (C.1) gives $V_2 = -I_2'(B + Z_0 D)/(A + Z_0 C)$, hence

$$Z_{out} = \frac{V_2}{I_2'} = \frac{B + Z_0 D}{A + Z_0 C}, \qquad S_{22} = \frac{Z_{out} - Z_0}{Z_{out} + Z_0} = \frac{-A + B/Z_0 - CZ_0 + D}{\Delta}. \quad \text{(C.15)}$$

The port-1 current is $I_1 = CV_2 + DI_2' = (AD - BC) I_2'/(A + Z_0 C)$, so $b_1 = -\sqrt{Z_0} I_1 = -\sqrt{Z_0}(AD - BC) I_2'/(A + Z_0 C)$, while $a_2 = (V_2 - Z_0 I_2')/2\sqrt{Z_0} = -\sqrt{Z_0}\,\Delta\, I_2'/2(A + Z_0 C)$, giving $S_{12} = 2(AD - BC)/\Delta$. Collecting,

$$S_{11} = \frac{A + B/Z_0 - CZ_0 - D}{\Delta}, \quad S_{12} = \frac{2(AD - BC)}{\Delta}, \quad S_{21} = \frac{2}{\Delta}, \quad S_{22} = \frac{-A + B/Z_0 - CZ_0 + D}{\Delta}, \quad \text{(C.16)}$$

with $\Delta = A + B/Z_0 + CZ_0 + D$ and $Z_0 = 50\ \Omega$. Since $S_{12}/S_{21} = AD - BC = Y_{12}/Y_{21}$ for a single device, the reverse isolation of the model comes entirely from the C_{gd} feedback path.

Consistency checks. Three checks were run on the code before use: for a cascade of passive blocks only, $AD - BC = 1$ to machine precision and $S_{12} = S_{21}$; with all transistors switched off ($g_m = 0$), $|S_{21}| \le 1$ over the whole band; and as $\omega \to 0$ a single stage tends to the textbook gain $-g_m Z_L$. The model is used to identify which element controls which metric and to scan design parameters quickly, but the numbers reported in the main text always come from Spectre.

D Project schedule

The Master's thesis started on 15 February 2026 and the final report was submitted at the end of August 2026, which corresponds to a total duration of 28 weeks. Figure D.1 summarises the initially planned schedule and the schedule that was effectively followed. The three phases of the project follow the organisation of the manuscript: the generation and integration of the cryogenic PDK (Chapter 1), its validation against the foundry PDK (Chapter 2), and the design and optimisation of the cryogenic LNA (Chapter 3). The main deviations from the initial plan concern the transcapacitance and gate resistance studies, which required more iterations than anticipated, and the writing of the report, which was started earlier and carried out in parallel with the design work.

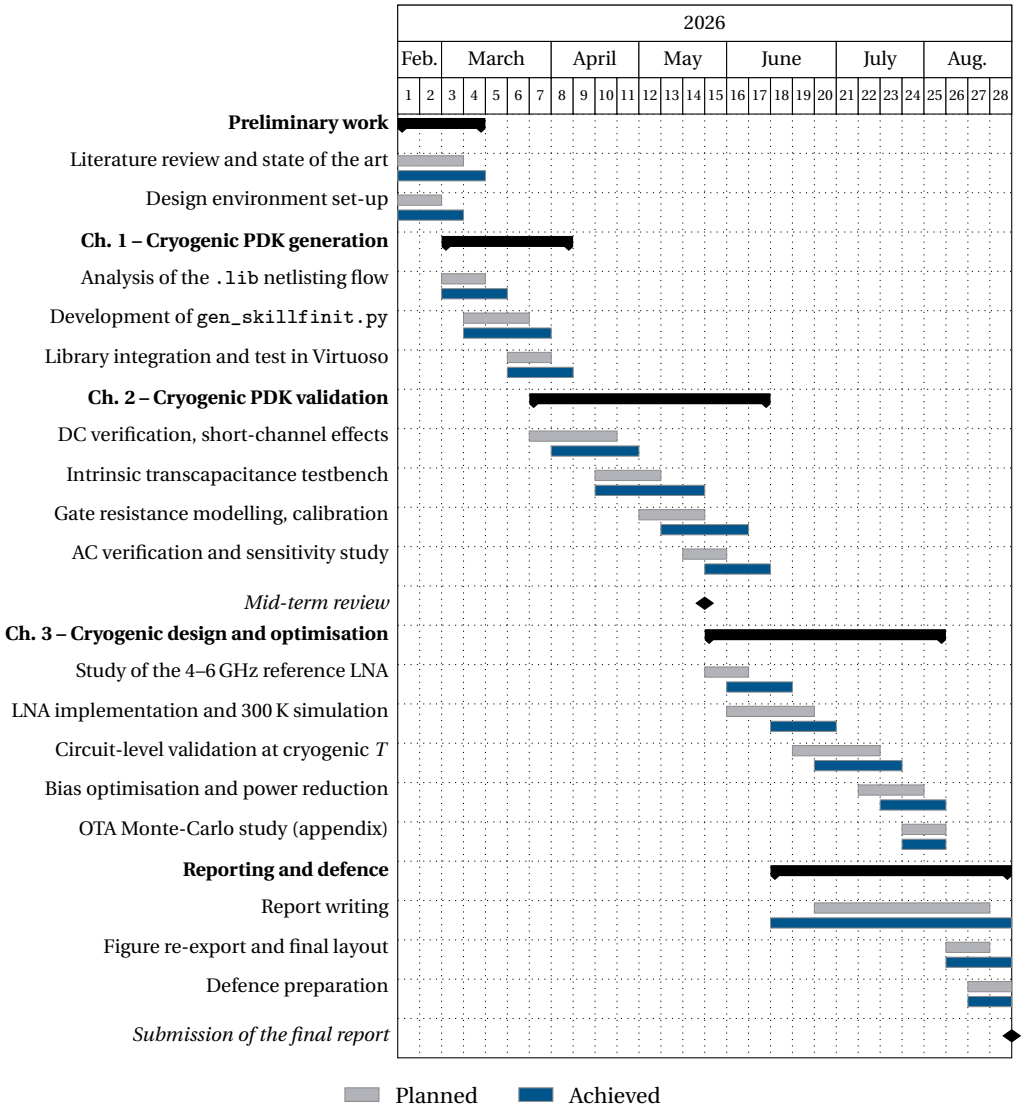


Figure D.1: Planned and achieved schedule of the Master's thesis, from the official start date (15 February 2026) to the submission of the final report. The time axis is graduated in weeks.

Resumé Français

Résumé

Le passage à l'échelle des ordinateurs quantiques impose de rapprocher l'électronique de commande des qubits, donc de concevoir des circuits CMOS fonctionnant à quelques kelvins. Les modèles compacts fournis par les fondeurs ne sont caractérisés que jusqu'à 233 K, ce qui interdit toute vérification des performances avant le tape-out. Ce travail de master, réalisé chez Rhonexum au sein du laboratoire AQUA de l'EPFL, met en place et valide une chaîne de conception cryogénique complète en technologie GF22FDX (22 nm FDSOI), fondée sur le modèle compact sEKV développé par l'entreprise.

La première partie automatise la génération du PDK. Un outil Python produit les scripts SKILL qui construisent la bibliothèque Cadence Virtuoso à partir des fichiers de modèle, avec les symboles, les vues et les interfaces de paramètres associées. Le flot supprime l'étape manuelle qui séparait l'extraction de paramètres de l'environnement de simulation.

La deuxième partie valide le modèle à température ambiante, où des références sont disponibles. Le comportement en canal court est vérifié en DC et en AC. Les transcapacités intrinsèques sont extraites par une méthode bilinéaire à trois transistors qui sépare les contributions intrinsèques et extrinsèques. La résistance de grille est ramenée à une expression à terme unique qui reproduit la valeur du PDK avec 2.2 % d'erreur RMS sur les balayages en largeur, en longueur et en nombre de doigts.

La troisième partie porte sur le circuit. Un amplificateur faible bruit 4 to 6 GHz publié et mesuré à 8 K est re-simulé avec le PDK cryogénique. La consommation reproduit la référence à 0.2 % et le gain moyen dans la bande à environ 2 %. Les écarts restants sur la fréquence centrale et la bande passante sont attribués à la surestimation de C_{gs} et à la sous-estimation de g_{ds}. Une optimisation du circuit ,à température cryogénique, de la polarisation réduit ensuite la consommation d'environ 46 % à bruit équivalent.

English Summary

Abstract

Scaling quantum computers requires moving the control electronics close to the qubits, hence designing CMOS circuits that operate at a few kelvin. Foundry compact models are characterised only down to 233 K, which prevents any performance verification before tape-out. This master's thesis, carried out at Rhonexum within the AQUA laboratory at EPFL, builds and validates a complete cryogenic design flow in GF22FDX (22 nm FDSOI), based on the sEKV compact model developed by the company.

The first part automates PDK generation. A Python tool produces the SKILL scripts that build the Cadence Virtuoso library from the model files, together with the symbols, views and parameter interfaces. The flow removes the manual step that separated parameter extraction from the simulation environment.

The second part validates the model at room temperature, where references are available. Short-channel behaviour is verified in DC and AC. The intrinsic transcapacitances are extracted with a three-transistor bilinear method that separates intrinsic from extrinsic contributions. The gate resistance is reduced to a single-term expression that reproduces the PDK value with 2.2 % RMS error across width, length and finger-count sweeps.

The third part addresses the circuit level. A published 4 to 6 GHz low-noise amplifier, measured at 8 K, is re-simulated with the cryogenic PDK. The DC power matches the reference within 0.2 % and the average in-band gain within about 2 %. The remaining deviations in centre frequency and bandwidth are traced to the overestimation of C_{gs} and the underestimation of g_{ds}. A bias optimisation then reduces the power consumption by about 46 % at equivalent noise performance.

Bibliography

[1] S. Das, S. Raman, and J. C. Bardin, "A 4-to-6-GHz cryogenic CMOS LNA with 4.4-K average noise temperature in 22-nm FDSOI," *IEEE Microwave and Wireless Technology Letters*, vol. 34, no. 4, pp. 411–414, 2024.

[2] S. Krinner et al., "Engineering cryogenic setups for 100-qubit scale superconducting circuit systems," *EPJ Quantum Technology*, vol. 6, no. 1, p. 2, 2019. DOI: 10.1140/epjqt/s40507-019-0072-0.

[3] B. Patra et al., "Cryo-CMOS circuits and systems for quantum computing applications," *IEEE Journal of Solid-State Circuits*, vol. 53, no. 1, pp. 309–321, 2018. DOI: 10.1109/JSSC.2017.2737549.

[4] J. C. Bardin, D. Sank, O. Naaman, and E. Jeffrey, "Quantum computing: An introduction for microwave engineers," *IEEE Microwave Magazine*, vol. 21, no. 8, pp. 24–44, 2020. DOI: 10.1109/MMM.2020.2993475.

[5] J. C. Bardin, "Cryogenic low-noise amplifiers: Noise performance and power dissipation," *IEEE Solid-State Circuits Magazine*, vol. 13, no. 2, pp. 22–35, 2021. DOI: 10.1109/MSSC.2021.3072803.

[6] Y. Peng, A. Ruffino, and E. Charbon, "A cryogenic broadband sub-1-dB NF CMOS low noise amplifier for quantum applications," *IEEE Journal of Solid-State Circuits*, vol. 56, no. 7, pp. 2040–2053, 2021. DOI: 10.1109/JSSC.2021.3073068.

[7] F. Balestra and G. Ghibaudo, "Physics and performance of nanoscale semiconductor devices at cryogenic temperatures," *Semiconductor Science and Technology*, vol. 32, no. 2, p. 023 002, 2017. DOI: 10.1088/1361-6641/32/2/023002.

[8] A. Beckers, F. Jazaeri, and C. Enz, "Characterization and modeling of 28-nm bulk CMOS technology down to 4.2 K," *IEEE Journal of the Electron Devices Society*, vol. 6, pp. 1007–1018, 2018. DOI: 10.1109/JEDS.2018.2817458.

[9] R. M. Incandela, L. Song, H. Homulle, E. Charbon, A. Vladimirescu, and F. Sebastiano, "Characterization and compact modeling of nanometer CMOS transistors at deep-cryogenic temperatures," *IEEE Journal of the Electron Devices Society*, vol. 6, pp. 996–1006, 2018. DOI: 10.1109/JEDS.2018.2821763.

[10] C. C. Enz and E. A. Vittoz, *Charge-Based MOS Transistor Modeling: The EKV Model for Low-Power and RF IC Design*. Chichester, UK: John Wiley & Sons, 2006.

[11] C. Enz, F. Chicco, and A. Pezzotta, "Nanoscale MOSFET modeling: Part 1: The simplified EKV model for the design of low-power analog circuits," *IEEE Solid-State Circuits Magazine*, vol. 9, no. 3, pp. 26–35, 2017. DOI: 10.1109/MSSC.2017.2712318.

[12] C. Enz, F. Chicco, and A. Pezzotta, "Nanoscale MOSFET modeling: Part 2: Using the inversion coefficient as the primary design parameter," *IEEE Solid-State Circuits Magazine*, vol. 9, no. 4, pp. 73–81, 2017. DOI: 10.1109/MSSC.2017.2745838.

[13] B. J. Sheu, D. L. Scharfetter, P.-K. Ko, and M.-C. Jeng, "BSIM: Berkeley short-channel IGFET model for MOS transistors," *IEEE Journal of Solid-State Circuits*, vol. 22, no. 4, pp. 558–566, 1987. DOI: 10.1109/JSSC.1987.1052773.

[14] D. Lu et al., *BSIM-IMG 102.8.0 independent multi-gate MOSFET compact model technical manual*, BSIM Group, University of California, Berkeley.

[15] B. Martinez et al., "Cryogenic circuit performance prediction using design-oriented model (SEKV) on 22nm FDSOI," in *2025 IEEE International Symposium on Circuits and Systems (ISCAS)*, IEEE, 2025. DOI: 10.1109/ISCAS56072.2025.11043615.

[16] D. D. Lu, M. V. Dunga, C.-H. Lin, A. M. Niknejad, and C. Hu, "A computationally efficient compact model for fully-depleted SOI MOSFETs with independently-controlled front- and back-gates," *Solid-State Electronics*, vol. 62, no. 1, pp. 31–39, 2011.

[17] D. E. Ward and R. W. Dutton, "A charge-oriented model for MOS transistor capacitances," *IEEE Journal of Solid-State Circuits*, vol. 13, no. 5, pp. 703–708, 1978. DOI: 10.1109/JSSC.1978.1051123.

[18] Y. Tsividis and C. McAndrew, *Operation and Modeling of the MOS Transistor*, 3rd ed. New York, NY: Oxford University Press, 2011.

[19] B. Razavi, R.-H. Yan, and K. F. Lee, "Impact of distributed gate resistance on the performance of MOS devices," *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, vol. 41, no. 11, pp. 750–754, 1994.

[20] C. C. Enz and Y. Cheng, "MOS transistor modeling for RF IC design," *IEEE Journal of Solid-State Circuits*, vol. 35, no. 2, pp. 186–201, 2000.

[21] A. Litwin, "Overlooked interfacial silicide-polysilicon gate resistance in MOS transistors," *IEEE Transactions on Electron Devices*, vol. 48, no. 9, pp. 2179–2181, 2001.

[22] I. Jo and T. Matsuoka, "Accurate extraction of effective gate resistance in RF MOSFET," *Circuits and Systems*, vol. 6, no. 5, pp. 143–151, 2015.

[23] B. Dormieu, P. Scheer, C. Charbuillet, H. Jaouen, and F. Danneville, "Revisited RF compact model of gate resistance suitable for high-k/metal gate technology," *IEEE Transactions on Electron Devices*, vol. 60, no. 1, pp. 13–19, 2013.

[24] M. Lauritano, P. Baumgartner, and A. Ç. Ulusoy, "Test structures for the characterization of the gate resistance in 16 nm FinFET RF transistors," *Electronics*, vol. 12, no. 14, p. 3011, 2023. DOI: 10.3390/electronics12143011.

[25] T. Sonehara et al., "Mechanism of contact resistance reduction in nickel silicide films by Pt incorporation," *IEEE Transactions on Electron Devices*, vol. 58, no. 11, pp. 3778–3786, 2011.

[26] Y. Cheng and M. Matloubian, "High-frequency characterization of gate resistance in RF MOSFETs," *IEEE Electron Device Letters*, vol. 22, no. 2, pp. 98–100, 2001.

[27] S. Das, S. Raman, and J. C. Bardin, "Experimental characterization of the MOSFET Fano factor at cryogenic temperatures for accurate cryo-CMOS RF modeling," *IEEE Transactions on Microwave Theory and Techniques*, 2025, Early access.

[28] Y. Zeng, "Low-power HEMT LNAs for quantum computing," Licentiate thesis, Chalmers University of Technology, Gothenburg, Sweden, 2023.

[29] M.-A. Chalkiadaki, "Characterization and modeling of nanoscale MOSFET for ultra-low power RF IC design," PhD thesis No. 7030, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland, 2016.

Competency self-assessment report

Nanotech programme, Grenoble INP Phelma, UGA, academic year 2025–2026

Final-year project: *Cryo-CMOS Design and Validation for Quantum Computing Applications*, February to August 2026

Host organisation: Rhonexum, a start-up project incubated within the AQUA laboratory, EPFL

Industrial supervisor: Hung-Chi Han. Master's thesis director: Prof. Edoardo Charbon. Academic tutor: Maryline Bawedin.

D.1 Background and self-assessment approach

This report presents the self-assessment of the five competencies of the Phelma framework, as I developed them over my entire education and not only during my final-year project. Each section follows the same organisation: the context in which the competency was built, the argument supported by specific experiences and by the resources drawn on, the evidence that documents those experiences, and finally an assessment of whether the target level is fully or partially achieved. The intended learning outcomes of the framework are addressed one by one, and whenever one of them is only partially reached, I say so.

The background below serves as the backdrop for the whole report, and I prefer to set it out at once rather than scatter it across the five sections.

I arrived at Phelma after three years of preparatory classes. I was not a student to whom science came easily at high school, and my second preparatory year did not open the doors of the schools I was aiming at, which led me to repeat that year. Entering a school such as Phelma therefore represented a considerable achievement for me, obtained through long work and without shortcuts. I entered the PET track although I initially wanted the GEN programme, and the administration told me from the outset that this transfer would not be possible from my track. The only other programme that genuinely attracted me was Nanotech, whose selectivity and number of applicants I had understood. I then made a decision that shaped everything that followed: not to coast on what I had learnt in preparatory classes, as is common when arriving at an engineering school, but to treat the first year as a decisive one. I knew that ending up in a programme that did not interest me would have been a collapse. I therefore worked without neglecting a single point, and I obtained Nanotech without a single resit.

That first year was also my first experience of living far from my home town and of living on my own. I made friendships that I still maintain, I joined a student union list, and I discovered associative life as well as an artistic field to which I had never devoted myself. I also took up football again, which I had set aside.

The second year opened with a semester at Politecnico di Torino, followed by a return to Grenoble and a laboratory internship in Paris. The third year took place entirely at EPFL, with a semester of coursework and then the final-year project within Rhonexum, a young company incubated in Prof. Edoardo Charbon's AQUA laboratory. Three countries, three education systems, and at every step a level of demand higher than what I believed I could sustain.

One point of honesty is required about the final-year project. The Rhonexum project is collective

work. The small-signal extension of the sEKV compact model was developed during the same period by a colleague of the laboratory, and the extraction flow that produces the model cards from the company's measurements was built by another colleague, who was also responsible for the cryogenic extraction itself and for confirming its validity at low temperature. My own contributions, the only ones used as evidence in this report, are those summarised in the conclusion of my thesis: the room-temperature verification of the model in the static and small-signal regimes, the device-level characterisation that this verification made necessary, the Python tool that automatically generates the design kit, and the first validation of the model on a complete circuit at cryogenic temperature.

The technical evidence is provided as precise cross-references to the chapters, sections and figures of the thesis, a document the jury members already have, rather than by duplication.

D.2 C1: Designing or implementing engineering solutions to meet a specification

D.2.1 Level 1: follow an established approach or protocol

D.2.1.1 Context

Level 1 was built progressively from the practical courses of the first year at Phelma, then during the semester spent at EPFL, where the design and characterisation of printed circuit boards allowed me to connect, for the first time, the circuits seen in lectures to an object that could actually be measured. It was consolidated during my second-year internship in a Paris laboratory, on an optical bench dedicated to the detection of chiral molecules, and then during the first part of my final-year project.

D.2.1.2 Argument

Following an established protocol requires understanding what each step guarantees, otherwise the slightest deviation becomes invisible.

Implementing a solution from a given set of specifications. At EPFL, the design of printed circuit boards imposed a strict sequence: schematic, component selection, routing, fabrication, assembly, then measurement. That is the moment when I really discovered how electronics is done, as opposed to the electronics one studies. Measuring and designing many of the circuits seen in lectures made me understand concretely the gap between the idealised circuit and the real one, where routing parasitics and component tolerances move the result. That lesson proved directly reusable three years later, when I had to separate, among the deviations of the cryogenic LNA simulation, what came from the transistor model and what came from the absence of post-layout extraction.

Implementing an appropriate validation methodology. The experimental work of my second-year internship rested on fixed procedures, each of which had a precise reason. The experimental room was permanently held at 22 °C in order to guarantee the stability of the setups and avoid thermal expansion effects, an essential constraint when an experiment is prepared one day and measured the next. Calibration was systematically performed with a spectropolarimeter. The use of the xenon lamp, which is very power hungry, was restricted to strictly necessary periods. Respecting these constraints was not a formality: the sensitivity sought on circular dichroism measurements made any thermal drift unacceptable.

During the final-year project, reproducing at 4.4 K the 4 to 6 GHz low-noise amplifier published by Das *et al.* in GF22FDX technology required following the topology, the three transistor sizing points and the bias conditions of the article to the letter, failing which the comparison with the published measurements would have lost all meaning. It also required respecting the model selection rule,

known as the binning rule, which chooses the parameter card applied to a device according to its width and its length. This constraint is not trivial: an inconsistency between what the schematic believes it knows about a device and what the simulator knows about it produces no error message. A terminal order declared in the cell description that does not match the port order of the subcircuit, or a parameter name differing by one character, still converges and gives a wrong result. What I take from this is that the most dangerous part of a protocol is the part one believes is too simple to be checked.

Using the knowledge gained during my education. The final-year project did not mobilise new knowledge in the strict sense; it mobilised course knowledge applied to an unusual object. The device physics and the charge model come from the Phelma and Nanotech curricula, the theory of radio-frequency circuits and impedance matching from the courses taken at Politecnico and EPFL, the theory of two-ports and the conversion between representations from a high-frequency electronics course, and the numerical methods and scientific programming from the applied computing courses. The analytical chain-matrix model described below is, quite literally, a textbook exercise scaled up to a real circuit.

Reporting the results in a structured manner. The thesis is organised to be readable by a general engineer: a glossary resolves the notation conflicts, the structure follows the three phases of the project, every figure and every table is numbered and called in the text, and every result is presented together with the simulation conditions that produced it and its domain of validity. The numerical comparisons with the reference are gathered in tables rather than scattered through the text, so that a reader can check a value without rereading the demonstration. The imposed format, forty pages from the introduction to the conclusion, was respected.

Complying with deadlines and applicable regulations. The project ran over six months with a planned schedule built at the start and tracked to the end, whose deviation from the actual course appears in the appendix of the thesis. The external deadlines, submission of the report one week before the defence and intermediate milestones with the industrial supervisor, were met.

The regulatory constraints were real and conditioned the way I worked. The GF22FDX design kit is covered by a non-disclosure agreement with the foundry, its models are encrypted, and access to the Cadence tools is subject to licence. The model cards produced by the company are its intellectual property. I therefore designed my work so that I would never need to circumvent these protections: the transcapacitance measurement method described below works precisely *without* access to the internals of the encrypted model. Writing a report that is both complete for the jury and compatible with the requested confidentiality was part of the constraint.

D.2.1.3 Resources drawn on

- **Knowledge**: analogue electronics, MOSFET device physics, experimental optics, theory of radio-frequency circuits and impedance matching.
- **Skills**: design, routing and measurement of printed circuit boards, operation of an experimental bench, schematic entry in Cadence Virtuoso, Spectre simulation, reading of operating-point variables.
- **Attitudes**: rigour in respecting experimental conditions, vigilance on apparently secondary parameters, acceptance of redundant verification, respect for third-party intellectual property.

D.2.1.4 Evidence

Appendix D.8.2, the second-year impact analysis document describing the bench procedures. Appendix D.8.3, planned and actual schedule. Thesis, sections 2.1.1 and 4.1.5, glossary and overall

structure.

D.2.2 Level 2: propose an approach or a protocol

D.2.2.1 Context

Level 2 developed during the final-year project, in a context that made it almost compulsory. Rhonexum is a very small structure in a growth phase. Supervision there is available intermittently, and the expectation is concrete results delivered within short deadlines, because the company has to demonstrate the value of its offering in order to grow. I therefore spent a large part of the internship defining the approach myself, for want of a pre-existing protocol applicable to the problem at hand.

One precedent had prepared me for this. The group project carried out in the second year on an energy-harvesting module based on nanogenerators for bridge monitoring was the first exercise in which I was asked to design a device that did not exist, rather than to apply a known method. That project required calling on imagination while applying scientific knowledge, a combination I had never had to produce before.

D.2.2.2 Argument

Five achievements document this level.

Automating the generation of the cryogenic design kit. The model file produced by the extraction contains all the physics but describes model cards, not devices that can be placed on a schematic. To be usable, the model has to be presented to Cadence exactly like a foundry kit device, through three objects that must agree with one another: a symbol, a cell description that declares the editable parameters, their default values and the terminal order, and a Spectre netlist that binds the instance to the Verilog-A module. Building these by hand is feasible for one device and impractical for a complete kit, all the more so since extraction is an ongoing activity and every campaign produces an updated file.

I developed a Python tool, `gen_skillfinit.py`, which takes the model file as the single source of truth and automatically derives everything Cadence needs: a SKILL script that builds the library, the cell, its symbol and its description, and a Spectre netlist that declares the Verilog-A module and reproduces the binning logic. The design point that matters is that both files come from the same parse of the same input file: the consistency between what the schematic knows about the device and what the simulator knows about it is obtained by construction rather than by proofreading. For the company, this removes the manual step between an extraction campaign and a usable design kit, which is the practical prerequisite for distributing such libraries as a commercial product. I also delimited the scope of the tool in my report: it covers integration only, it does not inspect the Verilog-A source and it does not check that a model card is physically meaningful. Stating what a deliverable does not do is part of the deliverable.

Designing a testbench for intrinsic transcapacitances. Separating the intrinsic charges of a transistor from the extrinsic parasitics normally requires de-embedding structures fabricated next to the device, which do not exist in a simulation environment. Instead I built a testbench that exploits the fact that intrinsic and extrinsic charges do not follow the same geometric scaling law, which allows them to be separated by solving the system over three geometries. The method works without any access to the internals of the compact model, which matters on two counts: it made it possible to reproduce the internal partitioning of the GF22FDX model, although that model is encrypted, over the four front-gate transcapacitances at zero drain-source voltage, and it does so without ever undermining the protection the foundry has placed on its model. I also stated the limit of the method and the extension to a fourth geometry that would separate the source-side and drain-side

extrinsic coefficients, which is the condition of its validity away from that bias point.

Gate resistance model and the decision to simplify. I took from the literature the physical decomposition of the gate electrode resistance and calibrated it on the device using an automated extraction flow driving Cadence from Python. The model reproduces the extraction to 2.2 % RMS over three geometry sweeps, once a width-dependent access term is added, and the fitted vertical resistivity, $4.14\,\Omega\mu m^2$, falls within the range expected for a metal gate stack.

The part I am most satisfied with is not the calibration but the decision that followed. A four-term model is heavy to implement in a Verilog-A description meant to be simulated thousands of times. Rather than deciding in the name of an arbitrary tolerance, I measured the actual sensitivity of the extracted quantities to an error on the gate resistance, and that measurement justified reducing the expression to a single term, implemented in one line, at a cost below 0.3 % on the extracted capacitances. Deciding a simplification from a measured sensitivity rather than from a tolerance argument is, in my view, the difference between applying a method and taking responsibility for one.

Developing an analytical model of the amplifier. To optimise the passive network of the LNA, I needed the sensitivity of every metric to every element, information that simulation alone does not give in a readable form. I therefore established an analytical chain-matrix model of the complete amplifier, deriving the transistor two-port, the insertion of the source degeneration and the conversion from the chain matrix to the S-parameters, with consistency checks in the limiting cases. This model revealed a clear division of labour among the elements: the gate inductance controls the input matching frequency with almost no effect on its depth, the degeneration inductances control the depth of the dip, and the three drain inductances each shift the gain peak by roughly the same amount, which is exactly what is needed to detune the three resonances independently. That readability is what made a guided manual adjustment possible instead of a blind search.

Designing to a specification, with resources, schedule and compliance checks. The final optimisation was carried out under explicit constraints, with resources and a schedule I had to plan myself: Cadence licence time, sweep duration, and a two-month window for the circuit phase. I verified that the optimised circuit met every constraint. The recalibration of the passive network moves the figure of merit from 0.124 to 0.227 at constant power consumption, by acting on the shape of the response rather than on the power. The bias optimisation of the three stages, which the original methodology could not explore because it relied on a small-signal model extracted at a single operating point, then yields 8.46 mW for an average noise temperature of 4.39 K, that is 46 % less power than the nominal point of the article for comparable noise. I then verified that both optimised circuits remain unconditionally stable, and I characterised their linearity by time-domain simulation with coherent Fourier transform: the input compression point remains close to -49 dBm, which leaves 20.8 dB of margin for a thousand-tone multiplexed readout. Establishing this last criterion required going to the literature for the actual constraint of qubit readout and stating explicitly the two assumptions on which it rests, rather than concluding on an isolated number. A solution that improves one metric while silently degrading the others is not a solution.

D.2.2.3 Resources drawn on

- **Knowledge**: simplified EKV charge model, short-channel FDSOI device physics, theory of transcapacitances, impedance matching and inductive degeneration, two-port theory, amplifier stability and linearity.
- **Skills**: Python, SKILL, Verilog-A, SPICE and ngspice, Cadence Virtuoso, Spectre and ADE Maestro, OCEAN scripts, automated driving of simulations from Python, data processing and

plotting with matplotlib, Git.

- **Attitudes**: autonomy in a lightly supervised environment, ability to formulate a specification when nobody provides one, concern for explicitly delimiting the scope of what is delivered.

D.2.2.4 Evidence

Thesis: chapter 2 and figure 2.1 for the generator, sections 3.2.4 and 3.2.5 for the transcapacitance testbench, sections 3.3.7 to 3.3.10 for the gate resistance and the simplification decision, section 4.1.7 and appendix C for the analytical model, sections 4.1.8 to 4.1.13 for the optimisation and the verifications. Appendix D.8.3, planned and actual schedule.

D.2.3 Progression radars

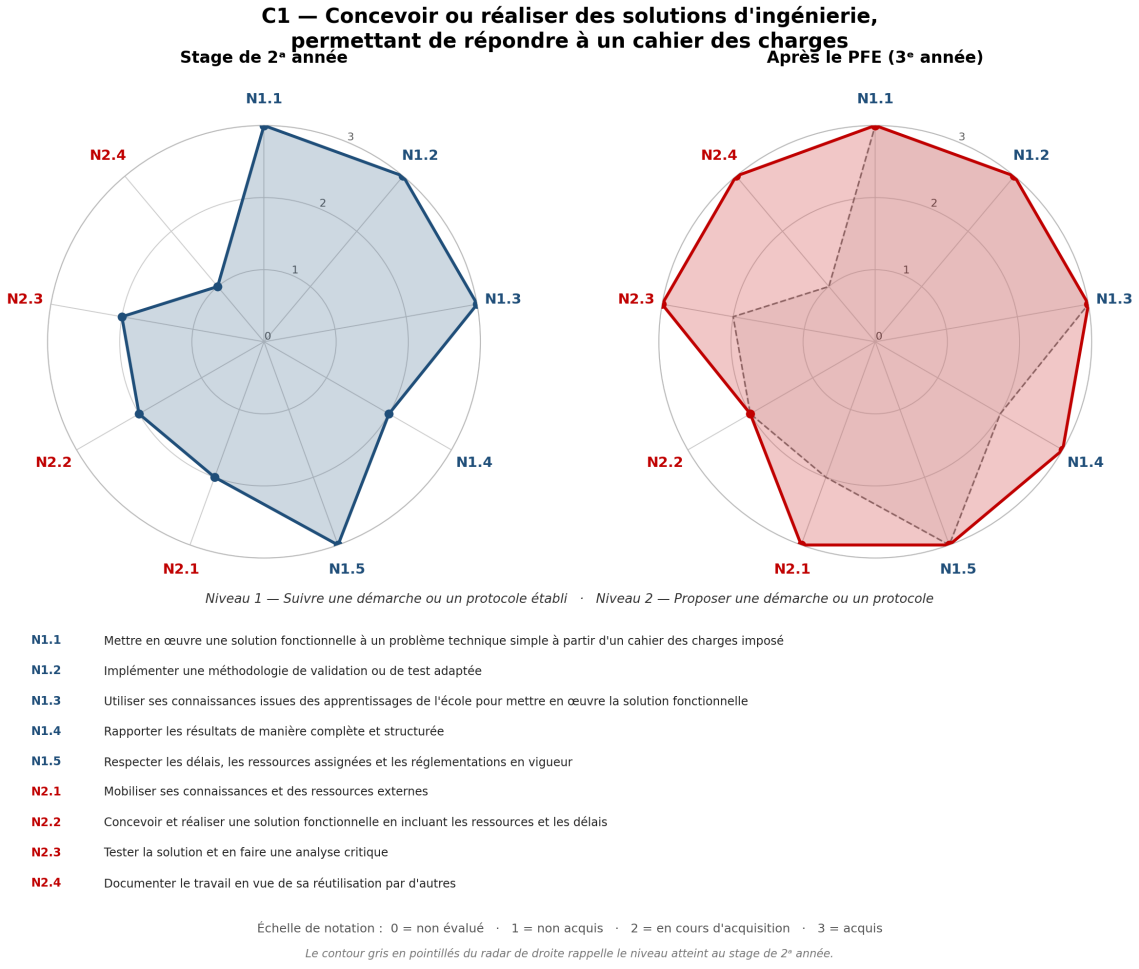


Figure D.2: Competency C1: radar filled in during the second-year internship, on the left, and radar updated after the final-year project, on the right. The dashed grey outline on the right recalls the level reached in the second year.

D.2.4 Final assessment

Level 1 is achieved, on all five intended learning outcomes. Level 2, the third-year target, is achieved as well: of the five achievements described, three consist in building a method that was absent or inapplicable as it stood, one produces a tool that is reusable and documented beyond my presence in the company, and the last conducts an optimisation under verified constraints. I consider that the point still to improve is my ability to estimate in advance the time a methodological adaptation

will take: calibrating the gate resistance model took me appreciably longer than I had announced.

D.3 C2: Implementing a fundamental research approach or an applied research approach for innovation purposes

D.3.1 Level 1: analyse and use an innovation or research approach

D.3.1.1 Context

My first real contact with research took place during my second-year internship, in a Paris laboratory working on the detection of chiral molecules at low concentration by means of nanoapertures. That internship played a demystifying role: I understood the codes of research work, its pace, its irreducible share of uncertainty, and I confirmed the importance of working on a subject aligned with my values, combining technique with a form of societal responsibility.

D.3.1.2 Argument

Contributing to a research project in the second year. My participation was not merely observational. I proposed a calibration protocol for circular dichroism, carried out sensitivity measurements in a microscopy configuration, and designed a sample holder that proved essential to the alignment of the following experiments. That is when I understood the difference between a result and a usable result: a measurement only has value if the chain that produced it is reproducible by somebody else.

Understanding the existing situation that calls for a research approach. The most useful analysis of the final-year project concerned the context, and it opens my thesis. At room temperature, an analogue designer works in a closed loop: they enter a schematic, simulate it with the compact models of the design kit, and trust the predicted performance enough to commit to a fabrication run. At cryogenic temperature that loop is broken. The models delivered with an industrial kit are calibrated over a temperature range that stops well above 4 K, and nothing guarantees their extrapolation below, while the device physics itself changes: the threshold voltage increases, the subthreshold slope stops following the thermal law and saturates, the mobility is modified. The performance of a cryogenic circuit therefore cannot be verified before fabrication, and every design iteration costs a fabrication run followed by a cool-down.

Identifying reliable and applicable sources. The project rested on a restricted but structuring corpus. Every reference retained answered a question I could not settle on my own: the article by Das *et al.* for the topology and the reference measurements, the literature on the gate electrode for the form of the resistance model, the published work on FDSOI transcapacitances for the expected shape and sign of each curve, and the simplified EKV formulation for the model itself. The selection was made by need rather than by accumulation, and I favoured peer-reviewed publications and silicon-measured data over unvalidated simulation results.

Interpreting the state of the art and questioning it against the existing situation. Reading the state of the art is not enough; one has to determine what it allows in one's own context. Three examples of that questioning. The radio-frequency characterisation literature proposes a reference method to separate intrinsic and extrinsic charges, based on fabricated de-embedding structures; I established that it was inapplicable in a simulation environment and drew from that the need for a different method. The literature on gate electrode resistance proposes a physical decomposition, but its published forms contained no width-dependent access term, which calibration showed to be necessary; I therefore extended the form instead of forcing the data into the existing mould. Finally, published cryogenic amplifiers are optimised from small-signal models extracted at a single operating point; it is precisely that limit of the state of the art which defines the space left for my

work.

Identifying the barriers. I identified four types of barrier and treated them differently. The theoretical barrier is the absence of a charge model validated at cryogenic temperatures, a barrier my work does not lift but does document precisely. The technical barrier is twofold: an encrypted foundry model and the absence of de-embedding structures in simulation, circumvented by the three-geometry method. The organisational barrier lies in the fact that parameter extraction belongs to a colleague: my work consumes his deliverable, which imposed synchronisation points and a systematic check of the versions of the files exchanged. The human barrier, finally, is the intermittent supervision in a very small structure hosting a master's thesis student for the first time, to which I responded by preparing every exchange so as to get the most out of it in a short time.

D.3.1.3 Resources drawn on

- **Knowledge**: critical reading of scientific articles, physics of compact modelling, low-temperature device physics, technological context of quantum computing.
- **Skills**: bibliographic search and selection, reference management, reproduction of published results, design of an experimental setup.
- **Attitudes**: curiosity, acceptance of the uncertainty inherent to research, ability to work without knowing whether the outcome will be positive.

D.3.1.4 Evidence

Appendix D.8.2, analysis of the second-year research project and of its usefulness. Thesis, chapter 1 and bibliography.

D.3.2 Level 2: carry out a research or innovation process

D.3.2.1 Context

The final-year project is itself an application, since its object is the validation of a model against independent references. The difficulty was not to simulate, but to define in advance what would count as a success, and then to build the comparisons so that they would be interpretable.

D.3.2.2 Argument

Suggesting and developing new approaches. Three new approaches came out of this work: the differential three-geometry method to observe the transcapacitances of an encrypted model, the automatic design-kit generator, and the analytical chain-matrix model of the complete amplifier. None of the three was requested in the initial subject; each was born of an obstacle encountered and of the observation that no available method removed it.

Evaluating feasibility and arguing the choices. Each of these approaches was preceded by a reasoned feasibility assessment, and that is what distinguishes a proposal from an improvisation. Before building the three-geometry testbench, I assessed the S-parameter route with de-embedding and argued that it required structures that do not exist in simulation and brought in non-quasi-static effects at the useful frequency scale. Before running a full two-dimensional grid for the bias optimisation, I tested one-dimensional sweeps which showed that the in-band parameters are not monotonic with bias and that the constraints become active in different regions of the plane, which made a local optimiser unreliable: the choice of exploration method is therefore justified by a prior observation and not by habit. Before writing the analytical model, I estimated the number of simulations the same sensitivity map would have required in Cadence and concluded that it was not reachable within the available time.

Choosing a comparison point that makes the exercise conclusive. The verification of the model was carried out at room temperature against the GF22FDX foundry kit, which may seem paradoxical for a cryogenic model. It is on the contrary the only choice that makes the exercise conclusive: room temperature is the only operating point where an independent reference exists on both sides of the comparison. At low temperature there is no reference model to measure against, and a cross-validation there would be circular.

Bottom-up approach and traceability of the deviations. The validation of the circuit was built in three tiers: verification at transistor level against the published parameters, then on the complete published circuit, then optimisation. Each tier only opened once the previous one had been validated, and the success criterion was not perfect agreement, which is impossible with a non-exhaustive compact model, but the ability to trace every deviation back to an identified mechanism.

The clearest demonstration lies in a precise chain of events. The transistor-level verification had established that the transport and noise quantities were reproduced to within a few percent, but that the small-signal quantities deviated markedly: gate-source capacitance overestimated by 88 % and output conductance underestimated by 34 %. These two deviations predicted a shift of the resonance frequencies and an excess of gain, and the simulation of the complete circuit confirmed both. The fine analysis is more interesting still than the prediction: the capacitance overestimation taken alone should have produced a shift well above the 10 % observed, which shows that this capacitance is only one contribution among others to the resonant network. A confirmed deviation is therefore not necessarily an understood deviation, and that is the distinction I have learnt to make.

Building a comparison in which the biases cancel. The main result of the fourth part rests on a methodological precaution that I regard as my most solid contribution. Four configurations were characterised under strictly identical simulation conditions, and that is what makes their comparison interpretable: the two model deviations identified upstream and the absence of post-layout extraction are common to all four and cancel in the comparison. The result I claim is therefore explicitly the relative one, stated as such in my conclusion: once a compact model with a bias-dependent noise law is available, a stage-by-stage bias optimisation delivers the same noise temperature for about half the power. The absolute values, on the other hand, remain conditioned by the small-signal parameters of the model, and I say so in my general conclusion rather than leaving the reader to discover it.

Statistical defence of a comparison against silicon. The model predicts the gain of a published operational transconductance amplifier to within 7 % of the measured silicon, closer than the foundry's nominal simulation. That favourable result could have been presented as it stood. I preferred to submit it to a variability analysis: over one thousand samples, the spread of the gain due to process alone is of the same order as the gap between the model and the nominal value of the kit, and the measured device itself lies 1.07 standard deviations from the mean. The honest conclusion is that the model's prediction is not an outlier of the distribution but a plausible member of it, close to the measured point. I also stated the second reservation, which works against me: the measurement contains the routing parasitics and the simulations do not, and including them would move the kit's value towards the measurement while moving the model's value away from it.

Autonomy in goal setting and decision making. The initial subject set a direction, not a method. I set the intermediate objectives of each phase myself and made the technical decisions that conditioned them, some of which committed everything that followed. The heaviest was locking an extraction configuration for which the transport and charge models cannot be satisfied simulta-

neously: I decided to freeze that configuration, to document it as a known bias, and to carry that bias through all subsequent comparisons, rather than waiting for a re-extraction that was neither possible in the available time nor within my scope.

Reasoned presentation, with perspective and critical thinking. The results were presented at two levels: orally, in the laboratory group meeting, in front of researchers who did not follow the project day to day, which required rebuilding the context in a few minutes; and in writing, in a report where every result comes with its domain of validity and where the reservations are stated in the same place as the figures they limit. Finally I listed in my perspectives what the work calls for: a dedicated small-signal extraction campaign, without which no result of this report can be tightened, a fourth geometry on the transcapacitance testbench, a sizing optimisation of the input stage which alone carries about 86 % of the amplifier noise, and a statistical description of the cryogenic model which is still missing. Stating precisely what comes next is, in my view, the last step of an applied research process.

D.3.2.3 Resources drawn on

- **Knowledge**: experimental method, propagation of deviations from one modelling level to the next, noise physics, variability analysis.
- **Skills**: design of simulation testbenches, sensitivity analysis, Monte-Carlo simulation, data processing and presentation, scientific writing in English.
- **Attitudes**: intellectual honesty, resistance to the temptation of presenting a result as more solid than it is, discipline in attributing contributions.

D.3.2.4 Evidence

Thesis: section 3.1.2 and appendix A for the statistical defence, sections 4.1.2 to 4.1.5 for the verification and the traceability of the deviations, section 4.1.3 for the noise model, section 4.1.14 and chapter 5 for the statement of the relative result and the perspectives.

D.3.3 Progression radars

D.3.4 Final assessment

Level 1 has been achieved since the second-year internship, on all five intended learning outcomes, including the identification of barriers and the questioning of the state of the art. Level 2, the third-year target, is achieved: the approach was defined before being applied, the feasibility of each new approach was argued before being committed to, and the favourable results were subjected to an explicit critique of their own validity.

D.4 C3: Collaborate within a team and in project mode

D.4.1 Level 1: participate in a team

D.4.1.1 Context

Participation in collective work began in the first year at Grenoble with my involvement in a student union list, continued in the second year with a technical group project, and then in the third year with integration into a company team.

D.4.1.2 Argument

Assuming a clear role and executing identified tasks. My involvement in a student union list was my discovery of associative life. The work there was collective, subject to a firm deadline and to an obligation of a result visible to the whole school. I contributed to the production of a piece of music,

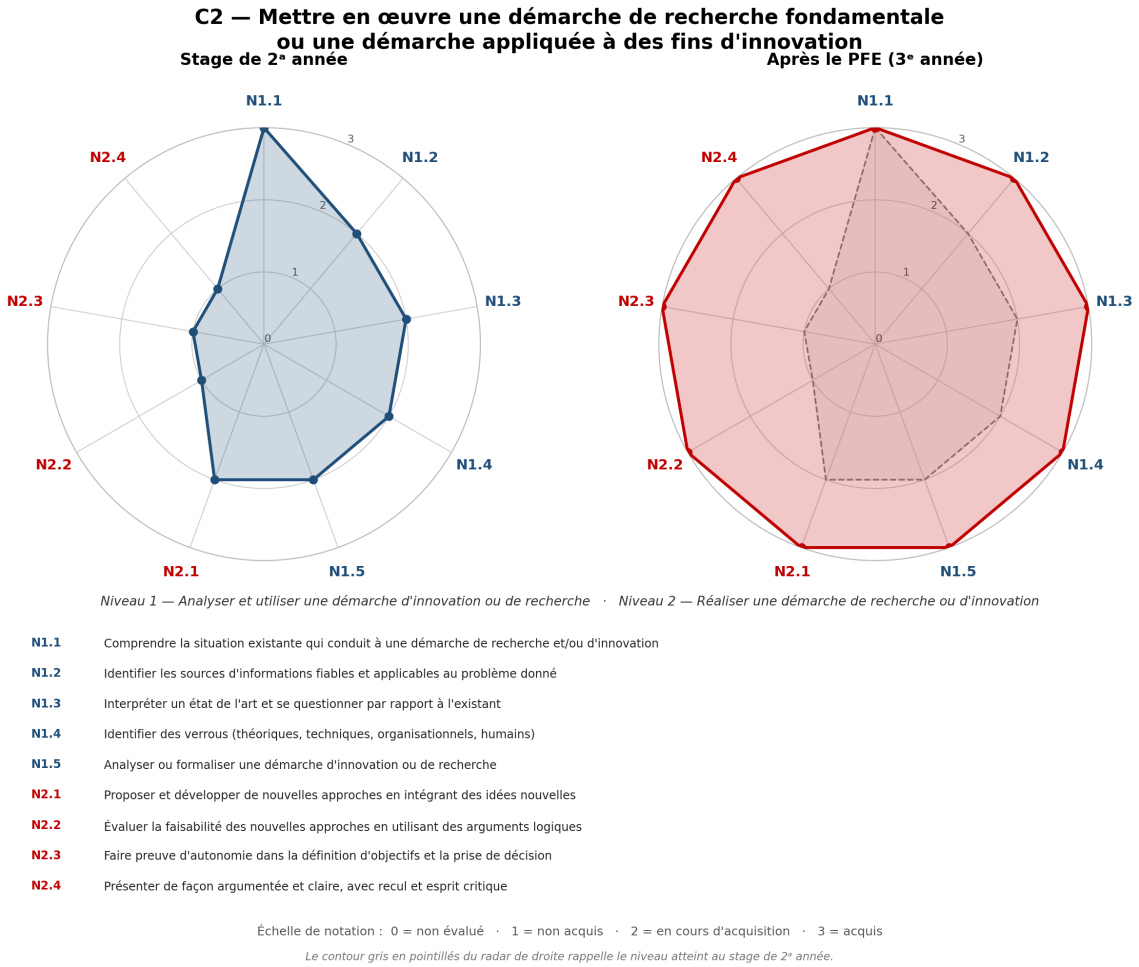


Figure D.3: Competency C2: radar of the second year and radar updated after the final-year project.

turned into a video clip, which reached a large audience. That experience had two lasting effects. It made me discover an artistic field to which I had never devoted myself and to which I would have liked to give more time. It put me in contact with students from other schools, which broadened my horizon well beyond Phelma. It also came at a specific moment of my path, my first year away from my home town and my first experience of living on my own, where collective life played a role that went well beyond the academic setting.

Achieving objectives independently within the team's organisation. Within Rhonexum, collaboration concerned precise technical interfaces. At the start of the project, the company's cryogenic compact model existed only in the static regime. The small-signal extension was developed during the same period by a colleague of the laboratory, and the extraction flow that produces the model cards from the measurements was built by another colleague, who was also responsible for the cryogenic extraction and for confirming its validity at low temperature. My work sits around that model: I consume their deliverable and I produce mine.

Collaborating meant here making sure that the model files exchanged were indeed the expected version, that the simulation conditions were shared and documented, and that the deviations observed were attributed to the right step of the chain. A large part of my validation work consisted precisely in not attributing to my scope a deviation coming from someone else, nor the reverse. This attribution discipline is not an administrative formality: in a chain where each person depends on the deliverable of the previous one, an attribution error sends everybody looking in the wrong place.

It is also the reason why the conclusion of my thesis begins by stating who did what before stating what I did.

Incorporating the opinions and critiques of others, while respecting diversity. The work was punctuated by corrections that I incorporated rather than defended. On substance, my industrial supervisor challenged several of my early interpretations of deviations, and in at least one case the correction led me to redo a complete simulation campaign. On form, a supervisory directive asked me to flatten the structure of my report, judged too finely subdivided: I rebuilt the hierarchy of five chapters rather than argue to keep it, and the document gained in readability. I also made and corrected my own mistakes, notably an error in the simulation temperature at the start of the project, which requires accepting to go back over work already done.

On respect for diversity, the nanogenerator project brought together seven students of the same master's programme from different countries and therefore from different educational backgrounds. Taking part meant dealing with ways of working that were not mine, a point I develop in the section devoted to the international competency. The final outcome was a relevant and ingenious product, with limits that we identified and stated rather than concealed.

Implementing project management tools. Three tools structured my work. A Gantt-type schedule, built at the beginning of the project and compared with the actual course, whose result appears in the appendix. A Git repository for the code of the generator, the extraction and plotting scripts, which served both as a backup and as a log of decisions. And regular reporting to my industrial supervisor, in the form of prepared meetings whose agenda I set, a format made necessary by the intermittent availability of supervision.

D.4.1.3 Resources drawn on

- **Knowledge**: organisation of a technical project, version control concepts.
- **Skills**: Git, Gantt chart, oral presentation in group meetings, writing of technical documents usable by a third party.
- **Attitudes**: reliability on commitments made, listening, willingness to go back over one's own work after a critique, explicit recognition of the contribution of others.

D.4.2 Level 2: lead or facilitate a project

D.4.2.1 Context

The framework formulates this level as leading a team *or* facilitating a project. I have never supervised anybody in the hierarchical sense of the term. It is therefore mainly through the second route, project leadership, that I reach this level, with a first partial experience of team facilitation during the second-year project.

D.4.2.2 Argument

Facilitating a subgroup on the nanogenerator project. The subject required designing a device that did not exist, which was a first for me: until then I had been asked to apply methods, not to imagine them. I led the French part of the group. That responsibility made me discover concretely what project leadership is: creating momentum in a group that has no appointed leader, distributing work whose outline nobody knows in advance, and accepting that an exploratory project comes with acknowledged limits rather than aiming at an illusory completeness. In particular I had to make the intermediate objectives explicit more than I would have done spontaneously, in order to make it acceptable to propose a path with no guarantee of success.

Structuring the final-year project and mobilising resources. My project comprised three phases with very unequal weights, of the order of ten, thirty and sixty percent, to be carried out over six months. Coordination consisted in breaking them down into subtasks, planning them and arbitrating between them. Generating the design kit is a technical prerequisite without which nothing else is possible, but it carried only a tenth of the assessment, which forbade devoting to it a time proportional to its actual difficulty. The validation of the model produces intermediate results that are unspectacular but indispensable to the interpretation of the third phase: without the small-signal deviations established upstream, none of the deviations observed on the circuit would have been interpretable, and the main part of the work would have collapsed. Mobilising resources was part of the exercise: access to licences and simulation servers, compute time for the sweeps, and above all access to people, since the deliverables I depended on were produced by colleagues for whom that was not the only task.

Monitoring progress, handling the unexpected and making decisions. The schedule was tracked and re-arbitrated. The main unexpected event occurred at the end of the second phase: the small-signal deviations turned out to be far larger than could be anticipated, which called into question the feasibility of the circuit phase as it had been planned. Two options existed: request a new extraction campaign, which was neither achievable in the remaining time nor within my scope, or lock the configuration, document the bias and design the rest so that this bias cancels in the comparisons. I chose the second, and that decision gave rise to the protocol of four configurations simulated under identical conditions. The time thus saved was reallocated to the analytical model, which had not been planned at the outset. A decision of that kind is not taken on a preference; it is taken by observing that one of the two routes is closed and rebuilding the method around that observation.

Communicating the results to decision makers. The results were communicated to three different audiences, with a level and an angle adapted to each. In the laboratory group meeting, in front of researchers and the head of the laboratory, the presentation dealt with the method and its validity, with the constraint of rebuilding the context for an audience that did not follow the project day to day. For the company, the question to be answered was not academic but operational: is the model usable today for design, and if so within what limits. Finally, the conclusion of the thesis is written as a prioritised recommendation addressed to those who decide the allocation of resources: the extraction campaign dedicated to the small-signal regime is priority number one, because no other result of the report can be tightened before it. Stating a priority rather than a list of wishes is what makes a conclusion useful to a decision.

Project leadership in a very small structure. This point deserves to be made explicit because it strongly conditioned the way I worked. Rhonexum is a young incubated company, very small, growing quickly, and I was the first master's thesis student it hosted. Two constraints follow. First, supervision is available only intermittently: over long periods there is nobody to ask a question to, and one has to decide alone, at the risk of reversing the decision later. Second, the expectation is concrete and rapid results, because the structure has to demonstrate its value in order to grow.

Leading my own project under those conditions meant setting my priorities without immediate validation, preparing every meeting with my industrial supervisor so as to get the most out of a limited time, documenting my decisions at the moment I took them so as to be able to defend them later, and accepting to move forward with a level of uncertainty that the academic setting had never imposed on me. In hindsight, it is the most transferable professional lesson of this internship, and the one I would not have learnt in a large structure where supervision is permanent.

D.4.2.3 Resources drawn on

- **Knowledge**: project leadership, planning and arbitration of priorities.
- **Skills**: Gantt chart, preparation and running of progress meetings, presentation of results in laboratory meetings, adaptation of the message to the audience.
- **Attitudes**: autonomy, decision making under uncertainty, ability to carry a group along without formal authority.

D.4.2.4 Evidence

Appendix D.8.3, planned and actual schedule of the project. Appendix D.8.4, nanogenerator group project. Thesis, chapter 5, contribution attribution paragraph and prioritised recommendations.

D.4.3 Progression radars

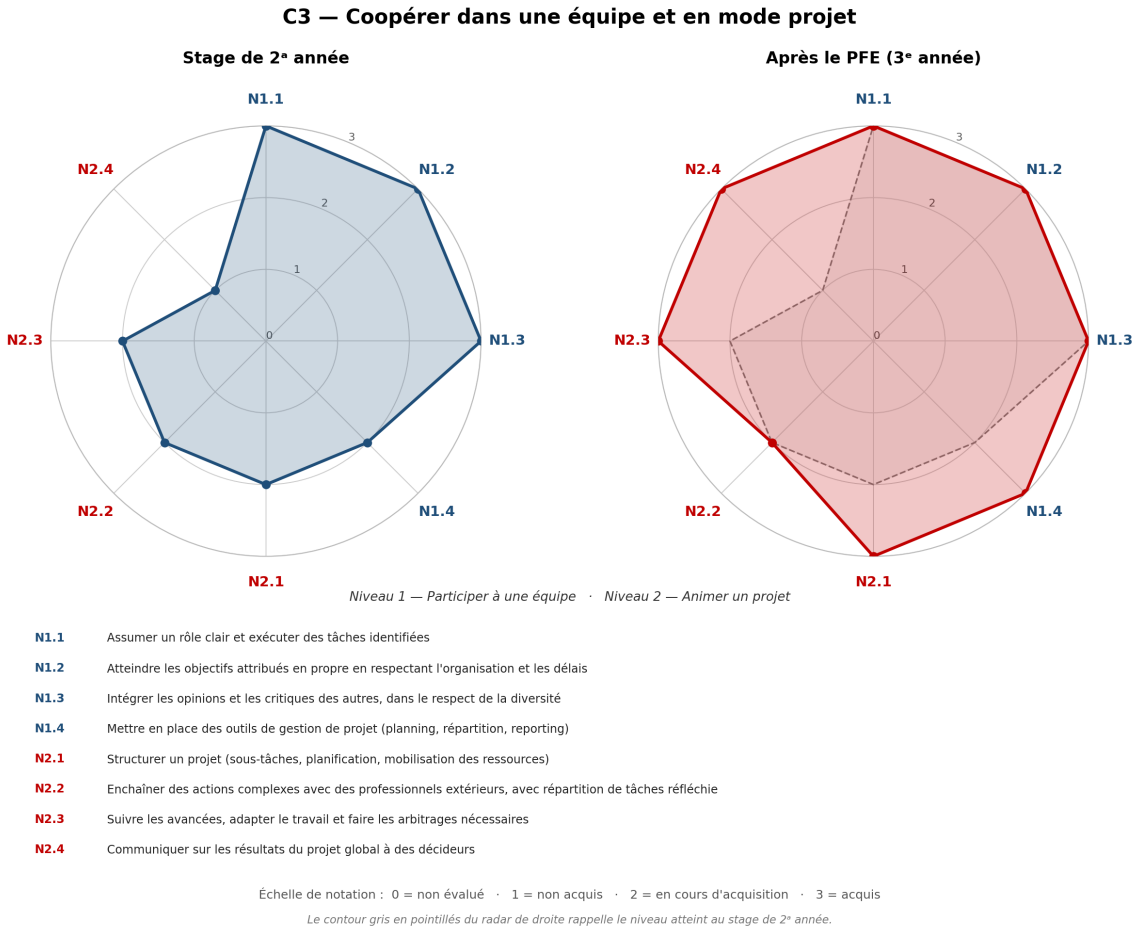


Figure D.4: Competency C3: radar of the second year and radar updated after the final-year project.

D.4.4 Final assessment

Level 1 is fully achieved on the four intended learning outcomes. Level 2 is achieved through the project leadership route: structuring into subtasks, mobilisation of resources, monitoring and re-arbitration in the face of a major unexpected event, and communication of results to decision makers. I consider, on the other hand, that team leadership in the sense of supervising people is still ahead of me: I have never had a colleague under my responsibility, and I do not wish to present as achieved something I have not done. That is the step I identify as the next one in my professional development.

D.5 C4: Work and communicate in an international and intercultural environment

D.5.1 Part A: justify international mobility

D.5.1.1 Context

My education took place in three countries: an undergraduate cycle at Grenoble, a semester at Politecnico di Torino in Italy, then a complete third year at EPFL in Switzerland, coursework and final-year project included. The two mobilities total about eighteen months, including six months of professional experience abroad, a duration well above the expected minimum. To this is added an international working environment within the AQUA laboratory and Rhonexum, as well as within the Paris laboratory of my second-year internship.

D.5.1.2 Comparative analysis of work and training habits

The framework asks for a comparative analysis between France and the mobility locations. I conduct it along six dimensions, based on what I observed rather than on what is generally said about each country.

Relationship to assessment. In France, at Phelma, assessment combines final examinations, written reports and projects, with a share of continuous assessment and a relative tolerance for imperfection in projects. In Italy, at Politecnico, the syllabus is very precisely prescribed and the exact restitution of the content gives the maximum mark; oral examinations hold a place there that the French system gives them little. In Switzerland, at EPFL, assessment rests on weekly exercise sets and examinations whose level of demand assumes continuous work, without resits of the same nature as in France.

Expected autonomy. This is the most striking difference. The Italian system I experienced expects compliant execution and leaves little room for initiative in the assessment. The Swiss system expects the opposite: the student is assumed to close the gap between the lecture and the examination alone, and nobody comes to check that they do. The French system sits between the two.

Density and pace of work. The workload in Turin was higher than what I had known at Grenoble, and that of Lausanne was beyond comparison with either. The second Grenoble semester had its own intensity, with an examination almost every Monday over the last two months, but the Lausanne pace was continuous rather than in peaks.

Relationship to failure and incompleteness. In the Italian context, handing in incomplete work is costly and therefore avoided, which discourages exploration. In the Swiss context, incompleteness is admitted if it is argued, which is also found in the research culture of the laboratory. That difference has direct consequences on how a group project is run.

Relationship between students and teaching staff. It was markedly more formal in Italy, more direct in France, and at EPFL strongly marked by the proximity between teaching and research, the teaching staff often being the researchers whose work one reads.

Language and environment. The Italian semester took place in an environment where English was the common language of the international students, Italian remaining the language of the country but not of the work. At EPFL and at Rhonexum, English is the default working language, including between French speakers as soon as a non-French speaker is present. In the company, finally, the logic changes in nature: what is expected is not a mark but a deliverable usable by somebody else, which is a third regime, distinct from the two academic ones.

D.5.1.3 Reflective argumentation on the competency criteria

Open-mindedness and tolerance. The nanogenerator project produced in me a judgement that I had to revise. I had first interpreted the reluctance of my classmates from the Italian system to take initiative as a lack of commitment. I later understood that it was a rational response to an assessment system that rewards exact restitution and penalises incompleteness: within their frame of reference, taking an unprescribed initiative is a risk with no counterpart. Revising that judgement is precisely what this criterion asks for, and I mention it because the first judgement was wrong.

Adaptability. I changed working system three times in three years. In Turin, I had to abandon my French revision method, built around understanding the reasoning, for a method suited to assessments requiring the precise restitution of a defined content. In Lausanne, I had to move from working in peaks before deadlines to working continuously, failing which the pace was unsustainable. At Rhonexum, I had to move from a logic of assessed results to a logic of usable deliverables, which changes what counts as finished work.

Ambiguity in intercultural relations. In a team where everyone works in a language that is nobody's mother tongue, agreement is often more fragile than it appears. A nod may mean that the other person has understood the words without sharing the same understanding of the thing. I learnt to reformulate what I believed I had understood and to have the points where ambiguity is costly confirmed in writing. That habit became a technical practice in my project: the systematic check of the version of the model files exchanged with my colleagues was born exactly from that observation.

Respect for cultural diversity. The AQUA laboratory brings together researchers and doctoral students of varied nationalities and different statuses: doctoral students, postdoctoral researchers, interns. The Paris laboratory of my second-year internship showed the same diversity of ages and cultures, with many international doctoral students and interns, and mutual respect and inclusion were part of the daily functioning there. Respect does not show only in behaviour: it also shows in the correct attribution of contributions, which is the technical form of recognising the work of others, and that is why the conclusion of my thesis begins there.

D.5.1.4 A reservation I state

I will not present these mobilities as uniformly successful experiences. The semester in Turin was difficult: distance from those close to me, a very high workload and a system I did not understand meant that I did not experience the city nor meet as many locals as I would have liked. What made me see it through is the degree I was aiming at and the refusal to abandon a step I had chosen. The first semester in Lausanne was the most demanding of my entire education, compounded by an internship search and by the disappointment of not obtaining the Paris position I was aiming at. From those two periods I take away something I would not have learnt otherwise: an international experience does not need to be pleasant in order to be formative, and the ability to hold on in an environment where one has neither one's landmarks nor one's network is a competency in itself. The second part of the year was in fact very different, the project at Rhonexum having been an excellent experience, through the people I met there and through the working environment.

D.5.2 Part B: master one or more foreign languages

D.5.2.1 Argument

English. The intended learning outcome is the certification of a B2 level in English. That certification has been obtained and the requirement is therefore met. My actual level today is beyond it, and that progress did not come from language classes but from use. English was my effective

working language throughout the third year: the courses at EPFL, the laboratory group meetings, the scientific literature, the code, the technical documentation and the entirety of my thesis. It was also my everyday language, since the people I met in Turin and then in Lausanne came from different countries and English was our only common ground.

That double exposure, professional and personal, made me cross a threshold I had not crossed in class. I can today hold real conversations in English, technical as well as personal, without preparation and without a conscious effort of translation. The most direct evidence is the writing in English of a forty-page technical thesis accompanied by a glossary, three appendices and a bibliography. This is not only a comprehension skill: it required writing in a scientific register, with the terminological precision demanded by a field where a badly chosen word changes the physical meaning of a sentence, and where notations that conflict from one chapter to the next had to be defined explicitly.

Italian. I am, on the other hand, honest on this point: my Italian did not progress during my semester in Turin and remains weak. The reason is simple and slightly paradoxical. The environment was so international that all my interactions, in class and outside, took place in English, and I almost never needed to resort to Italian. A mobility in a foreign country therefore does not automatically produce the acquisition of the language of that country. It produces the acquisition of the language actually practised, which in my case was English. I should have forced that use deliberately instead of waiting for the context to make it necessary, and that is a regret I state clearly.

D.5.2.2 Evidence

Thesis, written entirely in English. Appendix D.8.5, transcripts from Politecnico di Torino and EPFL, B2 English certification.

D.5.3 Final assessment

Part A is achieved on the three intended learning outcomes: two long mobilities in two different countries including six months of professional experience, a comparative analysis of work and training habits conducted along six dimensions, and a reflective argumentation on the four criteria of open-mindedness, adaptability, intercultural ambiguity and respect for diversity. Part B is achieved: the B2 certification in English has been obtained, and the actual professional use goes beyond that level. It is not achieved for Italian, and I take responsibility for the explanation.

D.6 C5: Integrate technological, environmental and societal transitions into the work

D.6.1 Level 1: analyse an existing situation

D.6.1.1 Context

Impact analysis was the object of formalised work during my second-year internship, in the form of a dedicated document analysing the environmental footprint and the societal stakes of the project I was taking part in. That document is the main evidence for this level. The analysis then dealt, during the final-year project, with the host organisation and with a technical system of a completely different nature.

D.6.1.2 Argument

Analysis of the carbon footprint and impacts of a project, second-year internship. I analysed the footprint of the project along several axes. On travel, the home-to-work journeys were made by shared bicycle or by metro, with a subscription partly subsidised by the laboratory, and no business

travel was necessary. On energy, I identified that the main source of impact was not the measurement equipment but the permanent thermal regulation of the experimental room, essential to the stability of the setups. That result was not intuitive and illustrates the value of the exercise: the dominant impact is not always the one suspected.

On resources, I noted the pooling of scientific equipment through a central platform, which avoids redundant purchases, the sorting of waste, the systematic recycling of the solvents used to clean the samples, the complete absence of paper printing, and even day-to-day details such as the use of coffee beans rather than capsules. I also identified the negative points without minimising them: cleanroom fabrication, power consumption of the optical equipment, and the use of precious metals such as gold and titanium.

The conclusion of that analysis was a balance. The project contributes positively to medical stakes, the detection of chiral biomolecules at very low concentration being a key element for applications in virology and biochemistry, and the use of nanoapertures reduces the amount of matter needed for the experiments. Against that, the processes involved generate a non-negligible environmental cost. Stating that balance rather than a verdict seems to me the only honest way to conclude an impact analysis.

Analysis of the societal and environmental aspects of the host organisation. The final-year project took place in an academic laboratory hosting a young company, a configuration that has its own characteristics.

On professional ethics, the dominant practice there is that of research: explicit attribution of contributions, traceability of data, and refusal to present a simulation result as a measurement. That requirement is not symbolic in a context where the company has a commercial interest in its model appearing good, and where the intern is the person producing the figures. That is the main ethical tension I encountered, and I responded to it by stating the reservations in the same place as the results, including where this weakened the most favourable figure.

On diversity, the laboratory brings together people of varied nationalities, ages and statuses, and English as a common language plays an equalising role in technical exchanges. On socio-ecological responsibility, the dominant characteristic is pooling by necessity: the cryostats, the microwave measurement equipment and the compute servers are shared resources, which mechanically limits the redundancy of equipment that is very costly in energy and materials. On waste reduction, the most notable fact concerning my own work is that it consumed no silicon wafer at all: all the results are obtained in simulation, and that is precisely the object of the project.

Footprint of the final-year project itself. My work was entirely computational. No business travel was necessary, the home-to-work journey was local, and no printing was done. The dominant item is therefore the electrical consumption of the simulation servers, notably during the two-dimensional sweeps and the one-thousand-sample variability analysis, which are the heaviest computations I ran. I note that this item, real as it is, must be compared with what the work makes it possible to avoid, which I develop at level 2.

Analysis of the systemic impact. The introduction of my thesis is built as an analysis of the existing situation at system scale. A superconducting quantum processor operates at a few tens of millikelvins, while the electronics that generates its control pulses and reads its state sits at room temperature. Each qubit therefore requires its own line going down into the cryostat. That architecture works for the few tens of qubits of current processors but does not scale to the thousands of physical qubits that fault-tolerant architectures demand: the cabling occupies the available volume, brings a thermal load at every stage, and limits the number of channels a dilution refrigerator can

host.

The approach studied to lift that limit consists in moving the control and readout electronics down into the cryostat, typically at the 4 K stage. And that is where the number governing everything else lies: the cooling power available at that stage is of the order of one watt. The first block of the readout chain, the low-noise amplifier, must deliver more than 35 dB of gain with a noise temperature of a few kelvins over the 4 to 8 GHz band, all within that budget. This analysis produces a conclusion that is unusual for an electronics project: the critical resource is neither silicon nor compute time, but one watt of cold.

D.6.1.3 Evidence

Appendix D.8.2, complete environmental and societal impact analysis document from the second-year internship. Thesis, chapter 1.

D.6.2 Level 2: propose solutions

D.6.2.1 Argument

Treating dissipated power as the structuring constraint. Since the cooling budget of the stage is of the order of one watt, the power consumed per amplifier does not merely determine an efficiency: it determines how many readout channels can fit in the cryostat, and therefore how many qubits the machine can address. That is the reformulation I adopted, and it changes the hierarchy of design criteria. A power saving that would be secondary at room temperature becomes here the limiting factor of scaling.

The main result of my work reads in that perspective. The bias optimisation of the three stages brings the consumption of the amplifier from 15.74 mW at the published nominal point down to 8.46 mW, for an average noise temperature of 4.39 K comparable to the reference, that is 46 % less power. A second operating point, at reduced supply voltage, goes down to 3.68 mW for 4.87 K. At constant cooling budget, this means roughly twice as many readout channels for the first point, and more still for the second at the price of slightly degraded noise. I verified that this saving is paid neither in stability, both circuits remaining unconditionally stable, nor in linearity, the compression point remaining close to -49 dBm with more than 20 dB of margin for a thousand-tone readout.

Making verifiable what was not, in order to remove physical iterations. The second proposal is the one carried by the company and to which my work contributes directly. Today, the performance of a cryogenic circuit cannot be verified before fabrication, and each design iteration costs a complete fabrication run followed by a cool-down. Each of those iterations consumes silicon wafers, chemicals, cleanroom process energy, masks, and several months of delay.

Providing a compact model valid at low temperature, usable in a standard simulator and placeable on a schematic like a foundry device, moves the verification from silicon to simulation. My contribution to that proposal is twofold. The automatic generation tool removes the manual step between an extraction campaign and a usable kit, which makes the cycle short enough to actually be practised. And the validation work answers the question that conditions everything else: to what extent the confidence placed in this model is justified, and on which parameters it is not yet. A model whose limits are unknown does not reduce the number of iterations, it merely displaces their cause.

From design to end of life. The framework asks for reasoning through to end of life, and that is the point on which my proposal is strongest, because it acts upstream rather than downstream. An integrated circuit in an advanced technology has no recycling channel: the quantities of metals it contains are too small and too intertwined to be recovered economically, and a research prototype

ends up as electronic waste after a few measurements. A cryostat poses a resource problem of another order, since dilution refrigerators consume helium, a finite resource with no substitute for that use. The conclusion I draw is that the most effective lever is not to improve end-of-life treatment, it is to reduce the number of objects manufactured. Removing one prototyping iteration removes its entire life cycle, from wafer to waste, which is structurally more effective than any recycling measure applied after the fact. That is the strongest environmental argument of the project, and it only appears if one reasons over the complete cycle.

As for the deliverable I produced myself, its end of life is a software one: a generation tool that is not maintained becomes unusable as soon as the environment changes. That is why I documented its scope and versioned its code rather than delivering an isolated script.

Adapting the solution to the various affected parties. A technical proposal never addresses a single actor, and I had to take account of at least four parties with different interests. The foundry protects its model by encryption: my transcapacitance measurement method was designed to work without ever accessing the internals of the model, which makes verification possible without undermining its intellectual property. The company needs a distributable product and a cycle short enough to be practised, hence the automation of the path from extraction to kit. The designer who will use the model needs it to fit into their existing flow like any foundry device, a constraint that dictated the form of the deliverable far more than its ease of implementation. The researcher, finally, needs to know the limits of validity, which justifies the report devoting as much space to the reservations as to the results. These four needs are not spontaneously aligned, and several of my design choices are arbitrations between them.

Proposals formulated during the second-year internship. In that context I had formulated several concrete avenues: recycling development samples to recover gold and titanium, optimising cleanroom protocols and the packaging of the final devices to reduce the energy cost, and extending the pooling of equipment already begun through a shared platform. Those proposals remained at the level of incremental improvement, but they constitute my first practical application of the approach, and I had then self-assessed at four out of five on this competency.

Limits I acknowledge. I will not present cryogenic electronics as a frugal technology. A cryostat consumes power permanently, fabrication in an advanced technology is resource intensive, and fault-tolerant quantum computing remains a technological bet whose societal return is not demonstrated today. I also note that my own results are simulation results without post-layout extraction, and that only a physical implementation would turn my relative conclusions into absolute ones.

My position is that the role of the engineer consists in making these costs and these reservations explicit, and in reducing what is reducible, rather than turning them into a commercial argument or ignoring them. That is what I tried to do by treating the cooling budget as the first-rank design constraint, and by stating the limits of validity of my own results in the same place as those results.

D.6.2.2 Resources drawn on

- **Knowledge**: thermal constraints of cryogenic systems, life cycle and environmental cost of microelectronics, architecture and scaling limits of quantum processors.
- **Skills**: structured impact analysis, translation of a technical gain into a system consequence, formulation of argued and verified proposals.
- **Attitudes**: critical thinking about one's own discipline, refusal of promotional discourse, respect for third-party intellectual property, ability to state a balanced assessment rather than a verdict.

D.6.2.3 Evidence

Thesis: chapter 1 for the analysis of the existing situation, sections 4.1.9 to 4.1.13 for the optimisation and its verifications, chapter 5 for the statement of the limits of validity. Appendix D.8.2 for the second-year impact analysis.

D.6.3 Progression radars

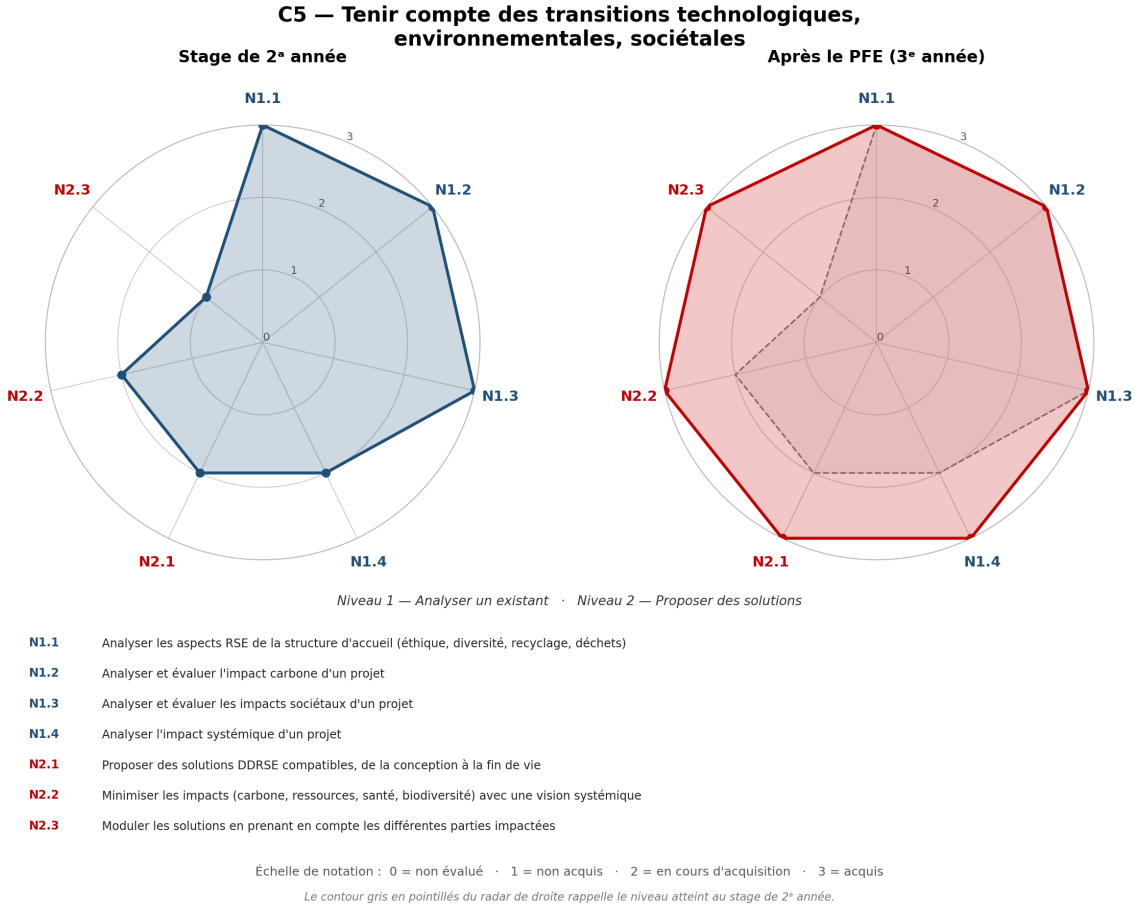


Figure D.5: Competency C5: radar of the second year and radar updated after the final-year project.

D.6.4 Final assessment

Level 1 is achieved on the four intended learning outcomes: analysis of the societal and environmental aspects of the host organisation, carbon footprint, societal impacts and systemic impact. Level 2, the third-year target, is achieved: the proposal did not remain declarative, it translated into a measurable design result, it reasons through to end of life by acting on the number of objects manufactured rather than on their final treatment, and it was arbitrated between the needs of four distinct stakeholders.

D.7 Summary and overall assessment

Three elements seem to me to structure this whole path.

The first is that nothing was easy, and that is precisely what I am most proud of. I did not arrive at Phelma with a natural facility for science, I did an extra year of preparatory classes after a second year that had not opened the schools I was aiming at, I entered a track that did not lead to the

Competency	3A target	Self-assessment	Status
C1 Design or implement	Level 2	Level 2	Achieved
C2 Research approach	Level 2	Level 2	Achieved
C3 Collaborate in a team or project	Level 2	Level 2 through project leadership, without supervising people	Partially achieved
C4 International	Parts A and B	Parts A and B, English certified at B2	Achieved
C5 Technological and societal transitions	Level 2	Level 2	Achieved

Table D.1: Summary of the self-assessment against the third-year targets.

programme I initially wanted, and I obtained Nanotech by refusing to coast on what I already knew. The same pattern repeated itself in Turin and then in Lausanne, where the level of demand each time exceeded what I believed I could sustain. No step was given, and I finish my education having validated every one of them.

The second is that autonomy has become my main professional resource, and that it was built out of necessity rather than by choice. The very small structure of the final-year project, where supervision is intermittently available and where the expectation is concrete and rapid results, forced me to decide alone, to document my decisions in order to be able to defend them, and to distinguish at all times what is demonstrated from what is assumed. That last habit is the one I consider the most solid of my technical acquisitions, and it is the one found in the very structure of my thesis, where every result comes with its domain of validity.

The third is the point on which I place my clearest margin for progress: supervising people. I have led a project, facilitated a subgroup, contributed to several collectives, but I have never had a colleague under my responsibility. That is the step I identify as the next one in my professional development, and I prefer to name it clearly rather than present a progression with no blind spot.

D.8 Evidence and appendices

D.8.1 Master's thesis

The thesis, to which this document is appended, is the main evidence for competencies C1, C2, C4 and C5. The cross-references used in the body of this report are the following.

- Chapter 1: analysis of the existing situation, cabling bottleneck, cooling budget of the 4 K stage, breakdown of the design loop at low temperature.
- Chapter 2 and figure 2.1: `gen_skillfinit.py` tool, motivation, structure of the generator and scope.
- Section 3.1 and appendix A: static validation of the short-channel model, circuit-level comparison and associated Monte-Carlo analysis.
- Section 3.2: three-transistor testbench for the extraction of the intrinsic transcapacitances.
- Section 3.3: gate resistance model, calibration to 2.2 % and simplification decision based on a sensitivity study.
- Section 4.1.3: addition of the channel noise model with a bias-dependent Fano factor.
- Sections 4.1.4 and 4.1.5: verification at transistor level and on the published circuit.
- Section 4.1.7 and appendix C: analytical chain-matrix model of the amplifier.
- Sections 4.1.8 to 4.1.13: recalibration of the passive network, bias optimisation, stability, time-domain verification and linearity.
- Section 4.1.14 and chapter 5: statement of the relative result, attribution of contributions and prioritised recommendations.

D.8.2 Impact analysis document from the second-year internship

Document entitled “Preuve de compétence: analysis of the environmental and societal impacts of the internship”, produced at the end of the second-year internship. It documents the carbon footprint of the project, the reduction initiatives, the ethical and societal aspects, the usefulness of the project, the balance of positive and negative impacts and the proposals for sustainable solutions. *Attach the complete document.*

D.8.3 Planned and actual schedule of the final-year project

Gantt chart comparing the planned schedule with the actual course of the three phases of the project. This document also appears in the appendix of the thesis.

D.8.4 Nanogenerator group project

Second-year project on an energy-harvesting module based on nanogenerators for bridge monitoring, carried out in a group of seven students. *Attach the project report, the final presentation or the assessment received if these documents are still available.*

D.8.5 Transcripts and language certification

Attach the transcripts from Politecnico di Torino and EPFL, together with the B2 English certification.

D.8.6 Competency radars

The eight radars, two per competency for C1, C2, C3 and C5, are inserted in the corresponding sections of the body of this report. They are grouped into four image files, `radar_C1.png`, `radar_C2.png`, `radar_C3.png` and `radar_C5.png`, each containing the second-year radar and the updated one, placed in `images/competences/`.