Balanced Ternary CMOS Logic for Scientific Computing: Complete Implementation Guide

Balanced ternary logic circuits offer compelling advantages for scientific computing: **symmetric signed representation eliminates sign-bit overhead**, simplified arithmetic reduces carry propagation, and natural alignment with neural network weight quantization enables 16× model compression with minimal accuracy loss. This guide provides a complete workflow from circuit design through simulation to FPGA and ASIC synthesis, specifically optimized for molecular dynamics, force field modeling, neural computing, and complex number arithmetic.

Circuit simulation tools differ dramatically in ternary support

The choice of simulation tool fundamentally impacts the viability of ternary circuit development. Commercial tools provide the most mature device model support, while open-source alternatives have reached production quality for research applications.

Cadence Spectre leads for multi-threshold voltage MOSFET simulation, supporting BSIM3, BSIM4, BSIM-CMG (FinFET), and BSIM-IMG models (Cadence) through its Compiled Model Interface. Research demonstrates successful ternary logic design in Virtuoso using 180nm and 90nm CMOS, with transient simulations validating full custom ternary multiplexers and half-adders. The tool handles the multi-voltage level simulation essential for ternary logic (0V, VDD/2, VDD) without modification.

Synopsys HSPICE provides comparable capability with Level 54 (BSIM4), Level 72 (BSIM-CMG), and Level 77 (BSIM6/BSIM-Bulk) models. Multi-threshold configuration uses distinct (.MODEL) statements with varying VTH0 parameters for LVT, SVT, and HVT devices—the foundation of ternary CMOS design.

ngspice (v45), the most capable open-source option, supports BSIM3 (Level 49), BSIM4 (Level 54), and BSIMSOI (Level 10) with OpenMP multicore acceleration providing approximately 2× speedup. The OSDI/OpenVAF interface enables Verilog-A model loading for custom devices including TFETs. When foundry models are unavailable, Arizona State University's Predictive Technology Models (PTM) provide research-grade alternatives. (Ngspice)

Xyce (v7.10) from Sandia National Labs offers MPI-based parallel computing scalable to hundreds of processors, (sandia) making it ideal for large ternary circuit verification. Its DAE formulation and Trilinos solver library (including KLU direct solver) (sandia) handle complex ternary topologies that cause convergence issues in traditional SPICE.

For **T-CMOS (Tunneling FET) devices**, Verilog-A compact models provide SPICE compatibility. Physics-based implementations combine the BSIM4 framework with band-to-band tunneling current models, capturing unique TFET behaviors including steep subthreshold swing and unidirectional conduction. Cadence Spectre and HSPICE support these through Verilog-A integration, while Xyce uses the ADMS/Xyce back-end compiler.

Simulating ternary gates requires specific configurations

The three fundamental ternary inverters—Standard (STI), Positive (PTI), and Negative (NTI)—each require different multi-threshold transistor arrangements. For a 1.8V VDD process, standard Vth of ~0.5V (VDD/4) and high Vth of ~1.35V (3×VDD/4) provide optimal switching characteristics. Simulation setup should use piecewise-linear inputs covering all three logic levels with explicit load capacitance (typically 1-50fF) and timesteps ≤1/10 of minimum signal transition time.

Convergence issues are common in ternary circuits due to intermediate voltage nodes. Setting $\boxed{\text{gmin=1e-12}}$, enabling $\boxed{\text{method=gear}}$ for stiff circuits, and using $\boxed{\text{.ic}}$ or $\boxed{\text{.nodeset}}$ for initial conditions at VDD/2 nodes resolves most convergence failures. Monte Carlo analysis across process corners (tt, ff, ss, fs, sf) is essential for validating threshold voltage mismatch tolerance.

FPGA implementation requires binary encoding with optimized primitives

Since no FPGAs support native ternary logic, implementation requires encoding schemes that map three-valued signals to binary fabric. The encoding choice significantly impacts area, speed, and power consumption.

Two-bit encoding (00→-1, 01→0, 10→+1) is most practical, wasting only 20.75% of information capacity versus the theoretical Shannon limit of $\log_2(3) \approx 1.585$ bits per trit. (hal) The unused state (11) enables error detection. **Compressed encoding** saves additional storage: 3-trit/5-bit compression provides 16% memory savings, while 5-trit/8-bit compression achieves 20% savings (hal) at the cost of decompression logic overhead.

FPGA family selection impacts achievable performance

AMD/Xilinx Versal ACAP with AI Engines provides the highest raw performance at **133 INT8 TOPS** (VC1902), scalable to 405 INT4 TOPS. (Xilinx) The Network-on-Chip interconnect enables high-bandwidth communication between compute engines (Medium) essential for ternary neural network inference. DSP58 blocks with 58-bit capabilities support efficient ternary arithmetic packing.

Xilinx UltraScale+ (VU9P, VU13P) offers the best balance for research applications. The 6-input LUTs enable optimized ternary adders: (HAL) a 3-input ternary adder fits in 2 LUTs (versus 3 from synthesis), a 4-input adder in 3 LUTs (versus 6), and a 7-input adder in 6 LUTs (versus 13)—achieving **51-57% LUT savings** in adder trees. (hal) Published results show **18.7 TOP/s** ternary neural network performance with peak efficiency of 1.62 TOP/s/W. (hal)

Intel Agilex 7 features AI Tensor Blocks with 20× INT8 multiplications per block, providing 5× density improvement over previous generations. (Cytech) The architecture supports INT8 down to INT2, with claimed **92 INT8 TOPS** (EEJournal) on M-Series devices.

HDL encoding follows established patterns

VHDL enumerated types with explicit encoding attributes provide clean abstraction:

```
vhdl
```

```
type ternary_t is (NEG_ONE, ZERO, POS_ONE);
attribute enum_encoding : string;
attribute enum_encoding of ternary_t : type is "10 00 01";
```

SystemVerilog enables more compact synthesis:

```systemverilog
typedef enum logic [1:0] {
  NEG_ONE = 2'b10, ZERO = 2'b00, POS_ONE = 2'b01
} ternary_t;
```

The ternary multiply-accumulate—the core operation for neural networks—eliminates actual multiplication through conditional selection: weight +1 adds the activation, weight -1 subtracts it, and weight 0 contributes nothing.

High-level synthesis tools accelerate development

hls4ml provides the most mature open-source HLS path for ternary neural networks, (GitHub) supporting AMD Vitis HLS, Intel Quartus HLS, and Siemens Catapult backends. (GitHub) Binary "-1" is encoded as "0" for efficient XNOR operations, (arXiv) with automatic optimization passes including ternary threshold extraction. The QKeras integration enables quantization-aware training directly targeting ternary weights.

FINN Framework from AMD Research generates dataflow-style architectures specifically for binarized and quantized networks, (GitHub) achieving sub-microsecond latency (KRS Documentation) and up to **50 TOP/s** on appropriate hardware. (ResearchGate) The workflow uses Brevitas (PyTorch) for quantization-aware training (KRS Documentation) and QONNX intermediate representation.

ASIC synthesis enables native multi-threshold ternary implementation

Unlike FPGA emulation, ASIC implementation can exploit true multi-threshold transistors for native ternary voltage levels, eliminating binary encoding overhead and achieving **33% reduction in interconnect** versus binary encoding.

Open-source PDKs provide viable fabrication paths

SKY130 (SkyWater 130nm) offers the most mature multi-Vth support (GitHub) with standard ((sky130_fd_pr__nfet_01v8)), low-Vt ((sky130_fd_pr__nfet_01v8_lvt)), and native devices ((sky130_fd_pr__nfet_03v3_nvt)). Research demonstrates CMOS180 technology achieving ternary logic with three threshold levels: LVT_NMOS ~0.292V, MVT_NMOS ~0.42V, HVT_NMOS ~0.76V. (University of Rochester) Access is free through Google/Efabless collaboration. (Skywater Technology)

IHP SG13G2 (130nm SiGe BiCMOS) represents the most advanced open PDK for analog/RF applications, featuring SiGe:C npn-HBT bipolar devices with fT up to 350 GHz and dual gate oxides (1.2V digital, 3.3V

analog). (GitHub) The bipolar+CMOS combination enables precise threshold control essential for ternary circuits. PSP 103.6 compact models include advanced effects modeling. (Skillsurf)

GF180MCU (GlobalFoundries 180nm) provides 3.3V and 6V supply options (GitHub) with comprehensive documentation, though less characterized for multi-Vth ternary applications than SKY130.

Standard cell library creation follows established methodology

Creating ternary standard cells (STI, PTI, NTI, TMIN, TMAX) requires transistor-level design using multi-Vth devices, followed by characterization for Liberty (.lib), LEF, and GDS generation. For ternary cells, Liberty files must characterize **six transition types** (0→1, 1→2, 2→1, 1→0, 0→2, 2→0) rather than the binary two, with modified measurement thresholds appropriate for three-level operation. (GitHub)

LibreCell (lctime) provides open-source characterization using ngspice simulation, generating Liberty format output. The tool supports combinational and sequential cells with configurable output loads and slew times. (PyPI) Alternative approaches include vsdStdCellCharacterizer_sky130 for SKY130-specific characterization with bisection-based setup/hold extraction.

OpenLane adaptation requires workarounds for binary synthesis

Yosys, the synthesis engine in OpenLane, fundamentally operates on binary logic. (Yosyshq) Three practical workarounds exist:

1. **Binary encoding approach**: Encode ternary signals as 2-bit pairs, synthesize with standard binary logic, then map to ternary cells post-synthesis

2. **Direct instantiation**: Instantiate ternary cells as modules in RTL, bypassing synthesis for ternary portions

3. **MRCS Tool (Mixed Radix Circuit Synthesizer)**: Browser-based EDA specifically for ternary, generating HSPICE and Verilog netlists with MVL synthesis algorithms—**four MRCS designs have successfully taped out using OpenLane** (Usn)

Custom cell integration requires adding LEF and Liberty files to the OpenLane configuration (GitHub) and potentially modifying power distribution for three-rail systems (VDD, VMID, VSS).

Fabrication options range from free shuttles to commercial MPW

Efabless Open MPW Shuttle provides free fabrication for fully open-source designs using SKY130A, with the Caravel carrier chip offering 10mm² user area. (Efabless) (Skywater Technology) **ChipIgnite** ($10,000-30,000) enables private shuttle runs for proprietary designs. (Efabless) **Tiny Tapeout** offers ultra-low-cost entry ($50-150) suitable for educational projects— (Skywater Technology) a **REBEL-2 Balanced Ternary ALU** has already been submitted through this pathway.

Balanced ternary arithmetic simplifies scientific computing operations

Balanced ternary uses digits {-1, 0, +1}, enabling symmetric signed representation without separate sign bits.

Negation requires only flipping signs, subtraction converts trivially to addition, and **truncation equals rounding**—a unique property noted by Donald Knuth that eliminates rounding mode complexity in floating-point implementations.

Complex number representation leverages Eisenstein integers

The cube roots of unity provide natural ternary encoding for complex numbers. **Eisenstein integers** of form $z = a + b\omega$ (where $\omega = e^{(2\pi i/3)}$) form a triangular lattice in the complex plane with unique factorization properties. Signal constellations using Eisenstein integers have cardinality $M = 3^m$, with addition corresponding to **addition with carry over ternary number fields**. This property enables efficient physical-layer network coding and MIMO transmission with 4.77 dB SNR gain per level ($10 \times \log_{10}(3)$) in set partitioning applications.

Ternary neural networks demonstrate practical scientific computing benefits

Ternary Weight Networks (TWN) constrain weights to $\{-1, 0, +1\}$, eliminating multiplication entirely—only additions and subtractions remain. **Trained Ternary Quantization (TTQ)** learns both ternary values and assignments, achieving accuracy improvements over full precision on ResNet-32/44/56 (CIFAR-10) and reducing AlexNet ImageNet error from 46.1% (DoReFa-Net) to 42.5%.

Recent FPGA accelerators demonstrate the approach at scale: **TerEffic (2025)** achieves 16,300 tokens/sec for fully on-chip 370M parameter models with **192× throughput versus Jetson Orin Nano** and 19× power efficiency improvement. The **xTern RISC-V ISA extension** provides 67% higher throughput than 2-bit equivalents with only 5.2% power increase, yielding **57.1% energy efficiency improvement**.

For **molecular dynamics**, balanced ternary's symmetric representation naturally handles positive/negative charge interactions and attractive/repulsive force terms. However, direct ternary implementations for force calculations remain limited in literature—the primary pathway is through neural network-based surrogate models for force field prediction, where ternary quantization enables efficient inference of machine-learned potentials.

Design methodology adapts traditional approaches to three-valued logic

Ternary combinational design extends Karnaugh map minimization to 3-valued logic. A 2-input ternary truth table has 9 cells (3^2), with groups covering powers of 3 rather than powers of 2. The **27 monadic functions** (3^3) include STI (standard inverter), PTI/NTI (threshold inverters), increment/decrement (cyclic rotation), and decoder functions. Critically, MIN, MAX, and negation alone do **not** form a complete basis—additional decoding functions or increment operations are required.

Sequential logic uses tri-flops with three stable states

Ternary flip-flops—termed "flip-flap-flops"—maintain three stable states. The **PZN Tri-Flop** provides Set Positive, Set Zero, and Set Negative inputs. **D-Type Tri-Flops** and master-slave **T-Type Tri-Flops** enable

standard sequential design patterns. (Google Patents) US Patent 4,107,549 documents comprehensive CMOS ternary sequential circuits implementable with CD4007AE and CD4016AE chips.

Binary-ternary interfaces require explicit conversion circuits

Double Pass-Transistor Logic (DPL) converters implement 4-stage binary-to-ternary conversion with delay equalization, validated in TSMC 0.18μm. The reverse direction uses threshold detection circuits, with recent MTCMOS-based converters achieving 8.61% delay reduction and 28.72% power reduction versus earlier designs. (IEEE Xplore)

Memory interfaces use either 2-bit encoding (simplest) or packed encoding (5 trits in 8 bits, since $3^5=243 < 256=2^8$) for efficient storage. Error detection leverages the unused fourth state in 2-bit encoding.

Verification extends binary methodologies

Testbench design adapts VHDL's std_logic_1164 package by mapping Logic 0 → '0', Logic 1 → 'Z', Logic 2 → '1'. Coverage metrics extend to **trit coverage** (each signal achieves all three values), **9-transition coverage** per signal, and **3^n exhaustive coverage** for n-input functions.

Fault models add **Stuck-at-2 (SA2)** to traditional SA0/SA1, with total single stuck-at faults = 3n for n signal lines. ATPG algorithms (D-algorithm, PODEM, FAN) extend using D0, D1, D2 fault values, though the larger state space (3^n vs 2^n) increases test generation complexity.

Performance comparison reveals context-dependent tradeoffs

The FPGA-ASIC gap narrows significantly with hard block utilization. From Kuon & Rose's comprehensive 90nm study:

Metric	Logic Only	With DSP + Memory
Area ratio	40×	18-21×
Speed ratio	3.2×	2.1×
Dynamic power	12×	9.0×

For ternary implementations, these gaps have additional implications: ASIC enables native multi-Vth ternary (eliminating 2-bit encoding overhead), while FPGA's programmable routing represents unavoidable overhead regardless of encoding.

Power analysis tools span commercial and open-source options

Synopsys PrimePower provides golden signoff accuracy with cycle-accurate peak power analysis, accepting VCD/SAIF/FSDB switching activity files. **Intel PowerPlay** and **Xilinx Power Estimator** enable early-stage FPGA estimation before synthesis completion.

OpenSTA in OpenROAD provides research-grade power analysis using SAIF files and Liberty cell characterization, with standard format compatibility enabling cross-tool validation. Power breakdown across logic, routing, I/O, memory, and clock networks enables targeted optimization.

Accuracy requirements vary by application

Ternary floating-point using 27 trits (Ternary27 format) provides approximately **8 decimal digits precision**—better than 32-bit binary float (~5 digits), less than 64-bit double. The symmetric rounding property (truncation equals rounding) eliminates guard digit requirements and reduces catastrophic cancellation severity.

For **molecular dynamics**, most production codes require double precision for energy conservation, though single precision suffices for some force calculations. Machine-learned potentials can achieve CCSD(T) quantum chemistry accuracy while benefiting from ternary network inference efficiency.

Neural network inference with ternary weights shows minimal accuracy degradation: ResNet-32 on CIFAR-10 achieves 93.7% versus 93.9% full precision (0.2% loss), while ResNet-34 on ImageNet reaches 71.6% versus 74.0% (2.4% loss). Medical imaging applications demonstrate statistical equivalence between ternary and full precision implementations.

Complete workflow synthesis

The recommended development path proceeds through four phases:

Phase 1 (Simulation): Design ternary cells in ngspice or Cadence Spectre using multi-Vth models from SKY130 PDK or PTM. Validate STI, PTI, NTI inverters and TMIN/TMAX gates through transient simulation with Monte Carlo process variation analysis.

Phase 2 (FPGA Prototyping): Implement algorithms using 2-bit binary encoding on UltraScale+ or Versal hardware. Use hls4ml for neural network applications or hand-coded HDL with optimized ternary adder primitives for arithmetic circuits. Verify functional correctness and baseline performance.

Phase 3 (ASIC Cell Library): Create ternary standard cells in SKY130, characterize using LibreCell, generate Liberty/LEF views. Integrate with OpenLane flow using direct instantiation or MRCS synthesis. Validate through post-layout simulation.

Phase 4 (Fabrication): Submit through Efabless Open MPW (free, open-source requirement) or ChipIgnite (proprietary allowed). Validate silicon performance against simulation predictions.

This pathway enables ternary circuit development from concept through fabrication using entirely open-source tools and free fabrication programs, while commercial tools and foundries remain available for production applications requiring advanced nodes or higher performance guarantees.