Resource-Efficient FPGA Pseudorandom Number Generation

Hüsrev Cılasun*, Ivy Peng†, Maya Gokhale†
†Lawrence Livermore National Laboratory
*University of Minnesota, Twin Cities

Introduction

- Probability distributions play a critical role in diverse application domains.
  - In simulations, modeling physical properties of materials, of processes, or of behaviors.
  - For instance, molecular dynamics codes often utilize the Maxwell-Boltzmann distribution for modeling temperature.
- We introduce a resource-efficient hardware pseudo-random number generator (RNG) and two optimizations:
  - **Alias table partitioning:** Separates a target distribution into multiple sub-ranges and facilitates local optimizations in each sub-range to improve overall resource utilization
  - **Adaptive threshold resolution:** Adjusts bitsize for representing threshold values to the precision of underlying partition
- Our main contributions:
  - Analytic study driven by dual considerations of improving accuracy and hardware mapping optimization
  - Automated HDL generation of both simulation and synthesis scripts
  - Diverse use cases: emulating Gaussian delay profile in FPGA-based LiME memory system emulator [1]; random number server for HPC applications

Methodology

- Walker's Alias Method [2] is an efficient algorithm for FPGA hardware implementation. It generates arbitrary discrete distributions from uniformly generated random numbers. For a target distribution $E(\cdot)$, this method generates and uses a table of real threshold values $F(\cdot)$ and alternative index values $A(\cdot)$, where $F(\cdot)$, $A(\cdot)$, and $E(\cdot)$ are of the same length. Each output sample Y is generated as

$$Y = \begin{cases} X & U \leq F(X) \\ A(X) & U > F(X), \end{cases}$$

where U is a real uniform random number and X is a uniform random integer. The output quality is a function of the precision of U, i.e., increasing the bit size or representing U as a floating-point number [3] improves the quality.
- We target following Maxwell-Boltzmann distribution (Eq.1) which has its PDF as a function of temperature T and the Planck distribution (Eq.2) which is parameterized by the factor a.

$$f(x) = \frac{2hc^2}{x^5} \exp -\frac{hc}{xkT} \quad (1)$$

$$f(x) = \sqrt{\frac{2}{\pi}} x^2 \frac{\exp \frac{-x^2}{2a^2}}{a^3} \quad (2)$$

Integration with MATLAB

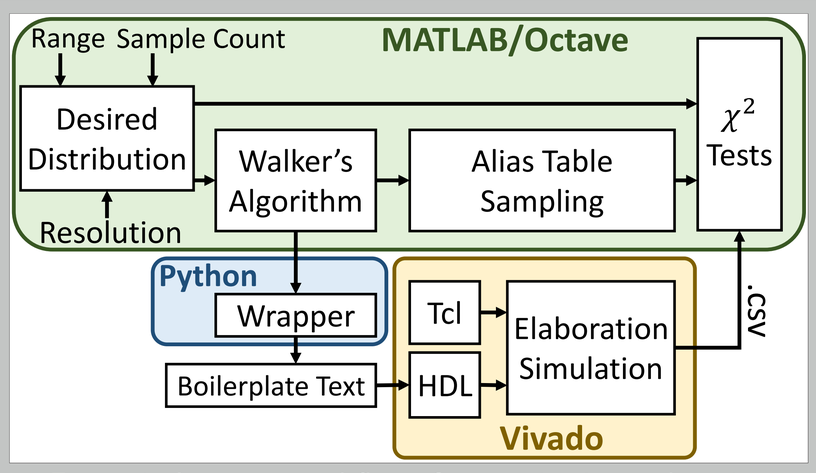


Figure 1: An automated flow of customization and testing

PwCLT Architecture

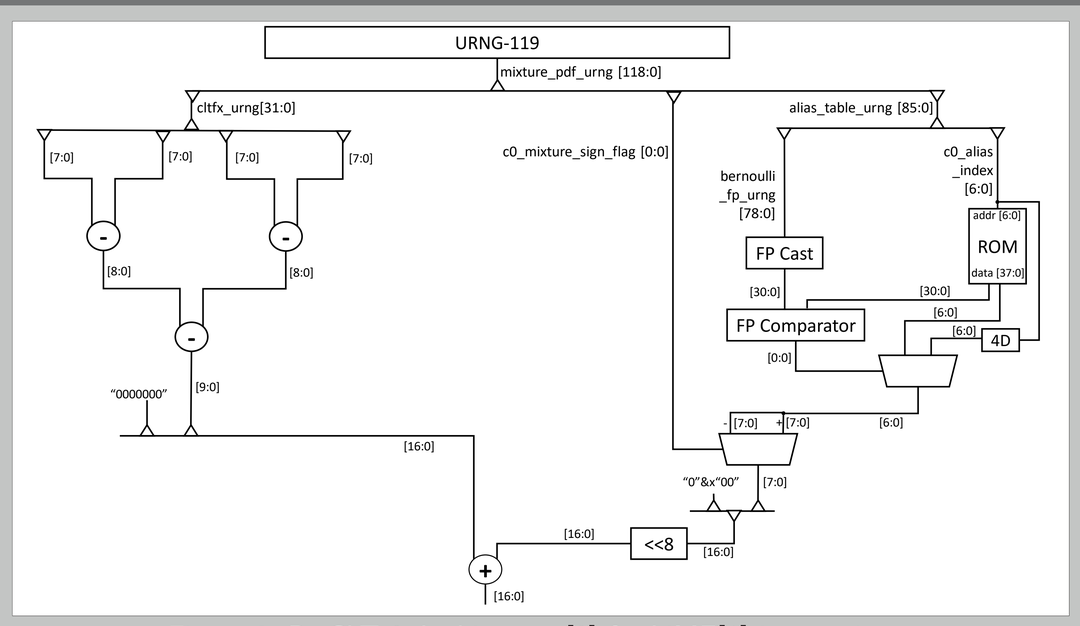


Figure 2: PwCLT-8 Architecture[3] for LiME[1] integration.

Alias Table Partitioning

- We improve the resource utilization for alias tables by **separating the target distribution into multiple subranges** (four subranges are exemplified in Fig. 3).
  - In each subrange, the standard alias table implementation is performed.
  - This separation allows each table to be optimized locally, i.e., alias tables whose target distribution is smoother can be configured to have fewer threshold bits in $F(\cdot)$ table per entry.
  - Consequently, the alias tables can be selected based on their relative probability range and lifted accordingly.
- We propose **adaptive threshold resolution** to adjust the threshold bitsize while maintaining statistical accuracy.
  - The quality of the generated samples is determined by the threshold resolution.
  - When alias table partitioning is employed, partitions with higher variance yield larger bitsize while smaller bitsize is required for those partitions with lower variance.

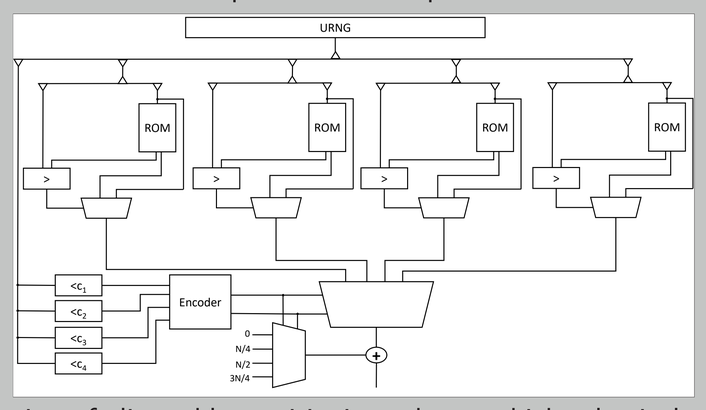
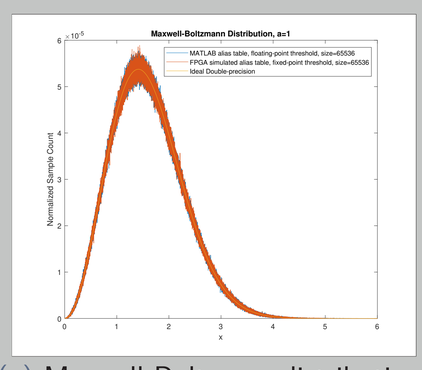
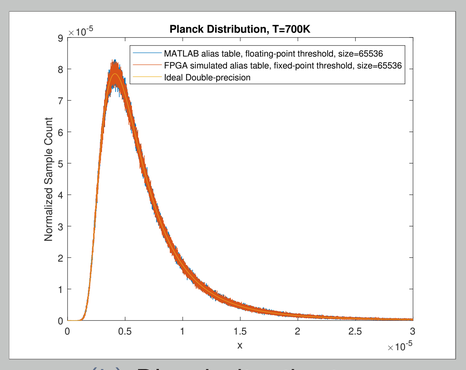


Figure 3: An illustration of alias table partitioning scheme which selectively combines sub distributions by comparing a uniform random variable with CDF values of each distribution in partition boundaries.

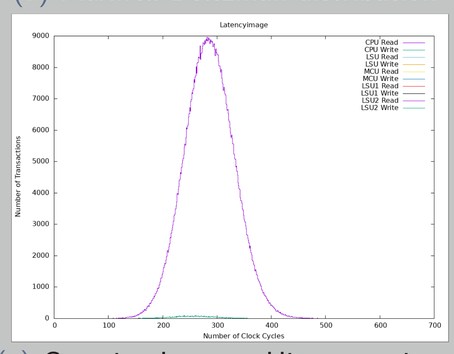
Validation and Evaluation



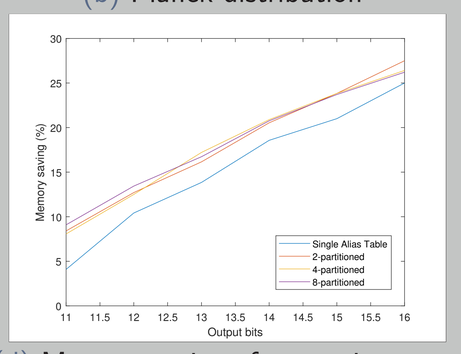
(a) Maxwell-Boltzman distribution



(b) Planck distribution



(c) Gaussian Latency Histogram in LiME



(d) Memory savings from various partitioning schemes.

Conclusion

- We introduced a resource-efficient hardware RNG whose accuracy is validated by χ^2 test.
- We proposed an alias table partitioning technique for optimizing resource utilization.
- Our RNG is evaluated in three use cases for memory emulations and scientific simulations.

References

[1] A. K. Jain, S. Lloyd, and M. Gokhale.
Microscope on memory: Mpsoc-enabled computer memory system assessments.
In *2018 IEEE 26th Annual International Symposium on Field-Programmable Custom Computing Machines (FCCM)*, pages 173–180, 2018.

[2] Alastair J Walker.
An efficient method for generating discrete random variables with general distributions.
ACM Transactions on Mathematical Software (TOMS), 3(3):253–256, 1977.

[3] D. B. Thomas.
FPGA gaussian random number generators with guaranteed statistical accuracy.
In *2014 IEEE 22nd Annual International Symposium on Field-Programmable Custom Computing Machines*, pages 149–156, 2014.

Acknowledgments

This work was supported by LLNL LDRD 19-ERD-004. LLNL-ABS-813772.