Project 3: DFT

Author: Jason Yuan

Student ID: A69042479

Github Repo: https://github.com/jas0xf/pp4fpgas-project3

Throughput (MSPS) = 1000 / (Estimated clock ns) / (Achieved II cycles)

Q1 DFT32 Baseline Implementation

(a)CORDIC accuracy vs. resources/performance

Accuracy knobs:

- **Iterations (rotations)** ↑ ⇒ accuracy ↑
- **Fixed-point total bits** ↑ ⇒ accuracy ↑

From Assignment 2 Results:

- **More rotations:** latency ↑; throughput ↓; **area ~unchanged** (e.g., ~+2 FF, LUT ~same)
- **More bits: FF/LUT increase noticeably** (clear growth with wider word-lengths); latency ~unchanged; throughput shows no clear trend

Q2 DFT32 Table Lookup

(a) Table lookup vs. baseline

DESIGN	EST. CLOCK (NS)	LATENCY (CYCLES)	INTERVAL (CYCLES)	THROUGHPUT (MSPS)	BRAM_18K	DSP	FF	LUT
Q1 Baseline (cos/sin)	7.256	55,908	55,909	0.002465	2	53	5,275	9,111
Q2 Table lookup (2D ROM)	7.256	257	258	0.5342	66	640	59,208	91,730

Q3 DFT32 Interface Change

(a) Why & Impact (from the end of 4.7 of the textbook)

- **Why:** Separate I/O avoids **in-place** read-after-write hazards and the final copy loop; each top-level array maps to its **own memory port**, reducing the **weakest-link** memory bottleneck.
- **Impact on optimizations:** Enables higher **memory bandwidth** for pipelining/unrolling (easier to hit low **II**), cleaner **array_partition/banking**, and pairs well with **loop interchange** and later dataflow/streaming.

(b) Before vs. after

DESIGN	EST. CLOCK (NS)	LATENCY (CYCLES)	INTERVAL (CYCLES)	THROUGHPUT (MSPS)	BRAM_18K	DSP	FF	LUT
Q2 DFT32 – Table lookup (in-place)	7.256	257	258	0.5342	66	640	59,208	91,730
Q3 DFT32 – Table lookup (separate I/O)	7.256	221	222	0.6208	64	640	59,189	91,521

(c) What changed
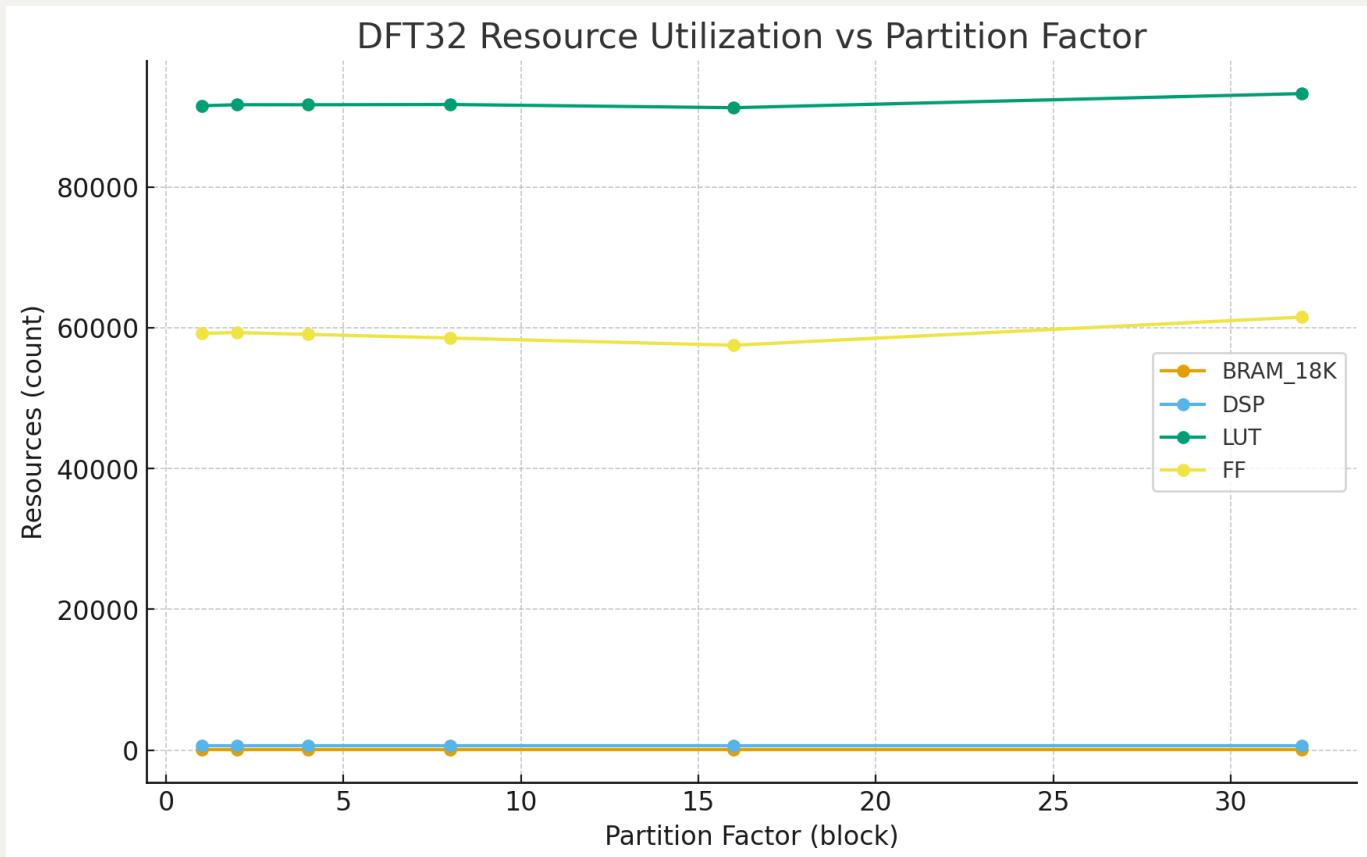
- **Throughput:** ↑ **0.5342 → 0.6208 MSPS** (~**+16.2%**), via smaller **Interval** (258 → 222).
- **Latency:** ↓ **257 → 221** cycles by eliminating the end-of-kernel copy.
- **Resources:** slight **BRAM** drop (66 → 64) and minor **FF/LUT** reductions; **DSP** unchanged.
- **Takeaway:** Separating I/O removes the in-place bottleneck and unlocks higher sustained parallelism for subsequent partitioning & unrolling.

Q4 DFT32 Array Partitioning

(a) Results table
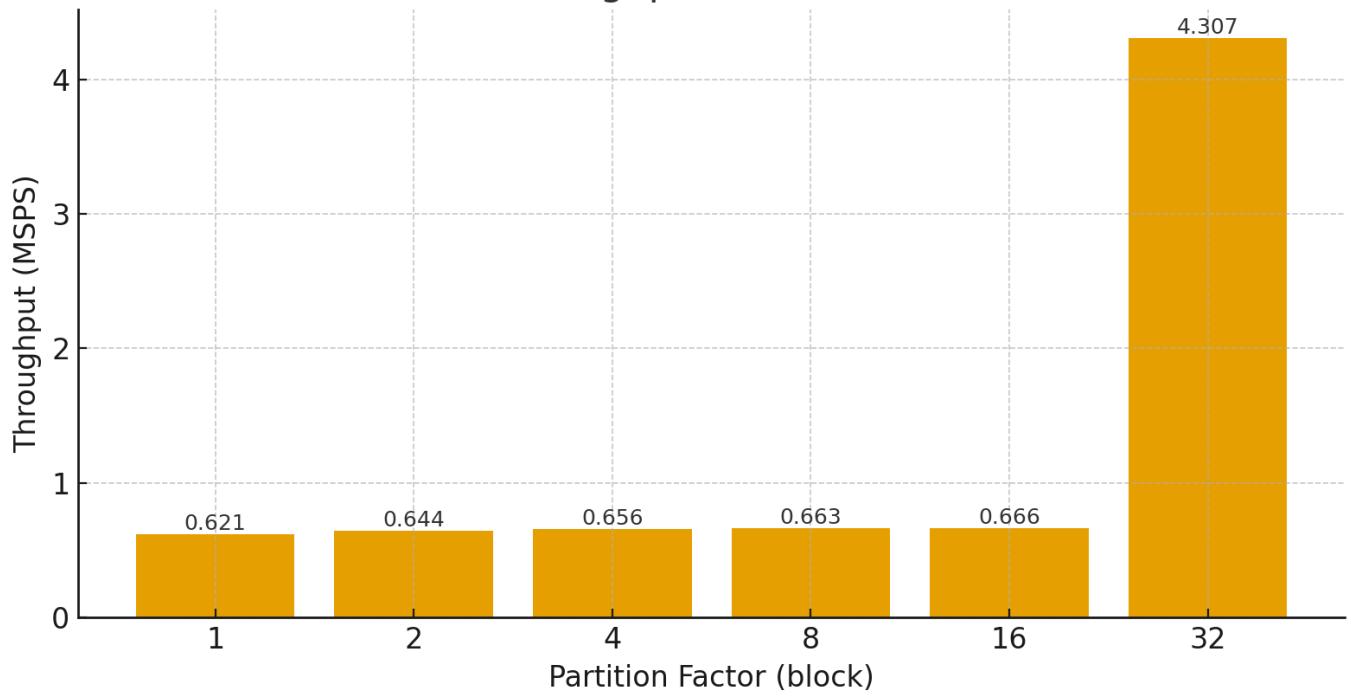
PARTITION FACTOR	EST. CLOCK (NS)	LATENCY (CYCLES)	ACHIEVED II (CYCLES)	THROUGHPUT (MSPS)	BRAM_18K	DSP	FF	LUT
1	7.256	221	222	0.6208	64	640	59,189	91,521
2	7.256	213	214	0.6440	64	640	59,314	91,675
4	7.256	209	210	0.6563	64	640	59,054	91,669
8	7.256	207	208	0.6626	64	640	58,540	91,705
16	7.256	206	207	0.6658	64	640	57,515	91,252
32	7.256	203	32	4.3068	64	640	61,509	93,250

(b) Resource utilization vs partition factor
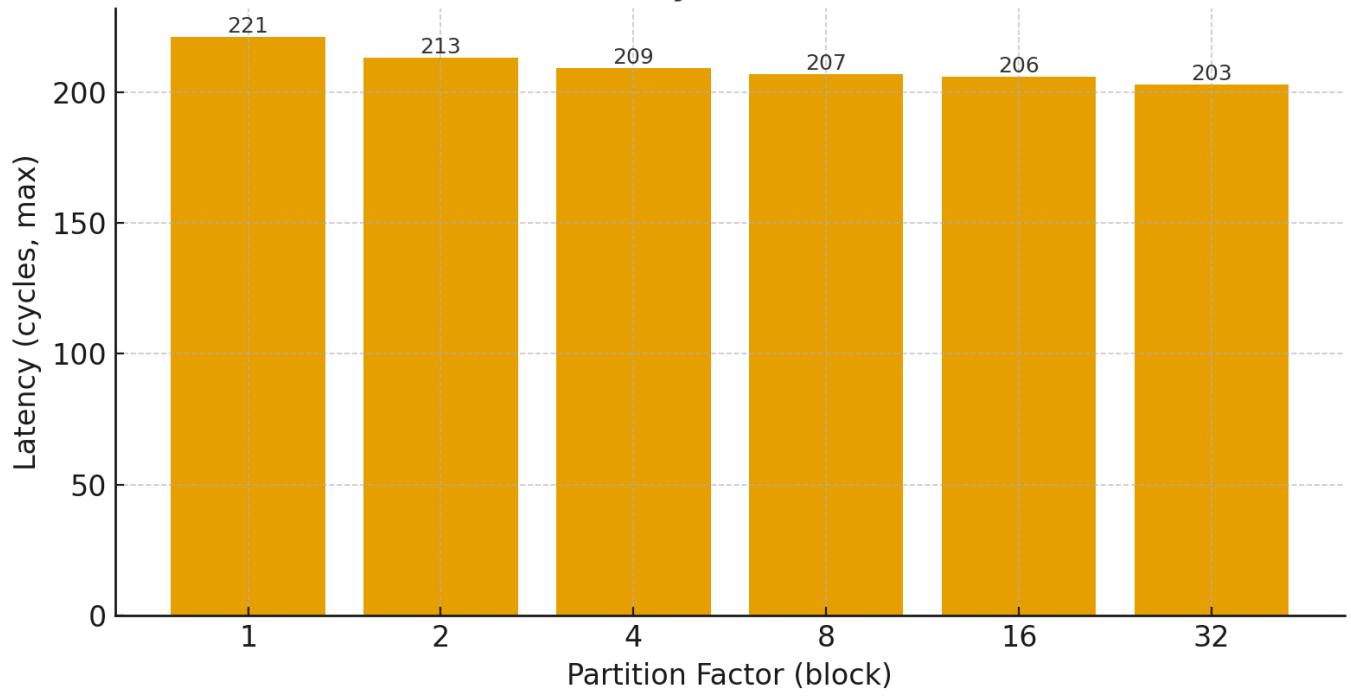
(c) Throughput & Latency vs partition factor

DFT32 Throughput vs Partition Factor


DFT32 Latency vs Partition Factor
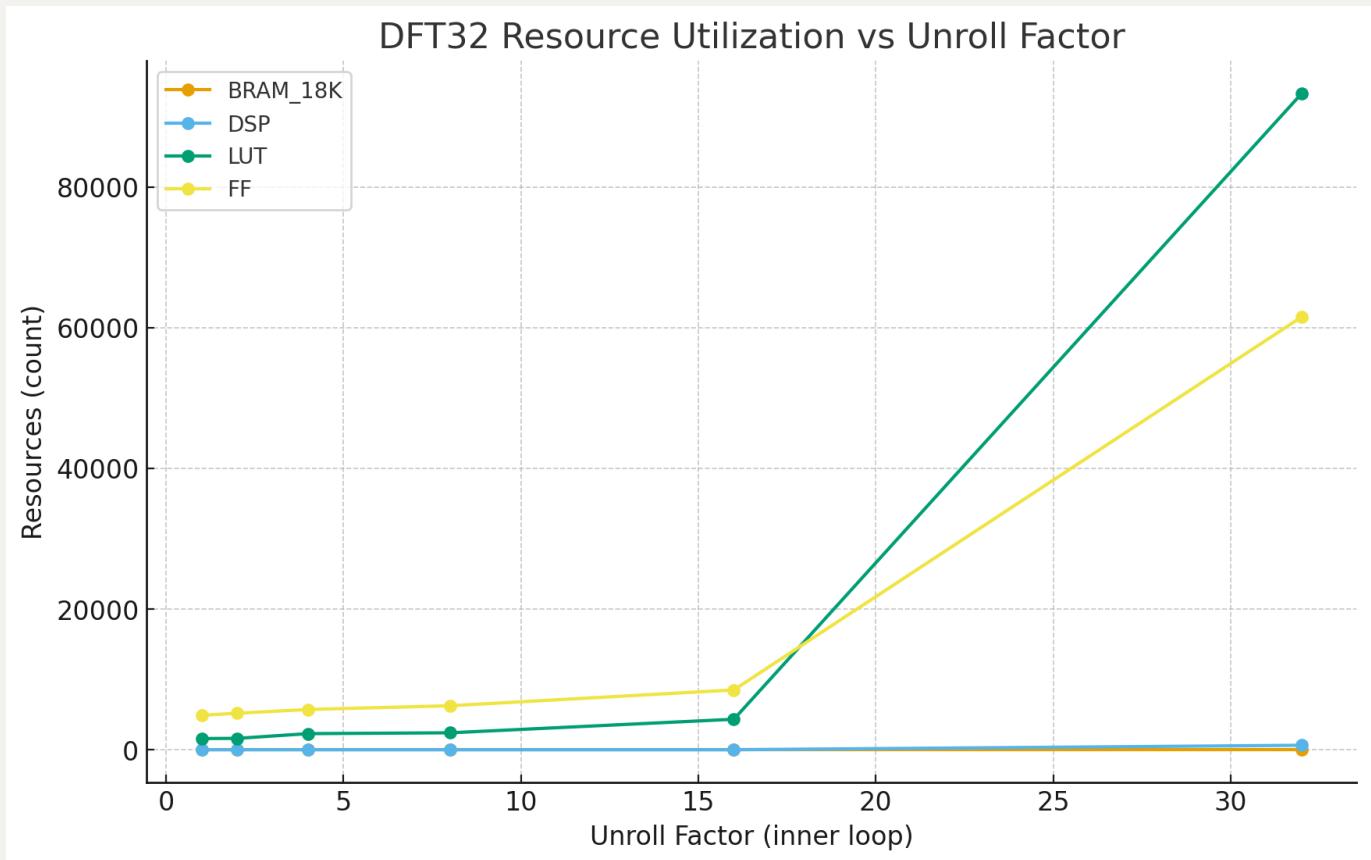
Q5 DFT32 Loop Unrolling

(a) Loop unrolling results
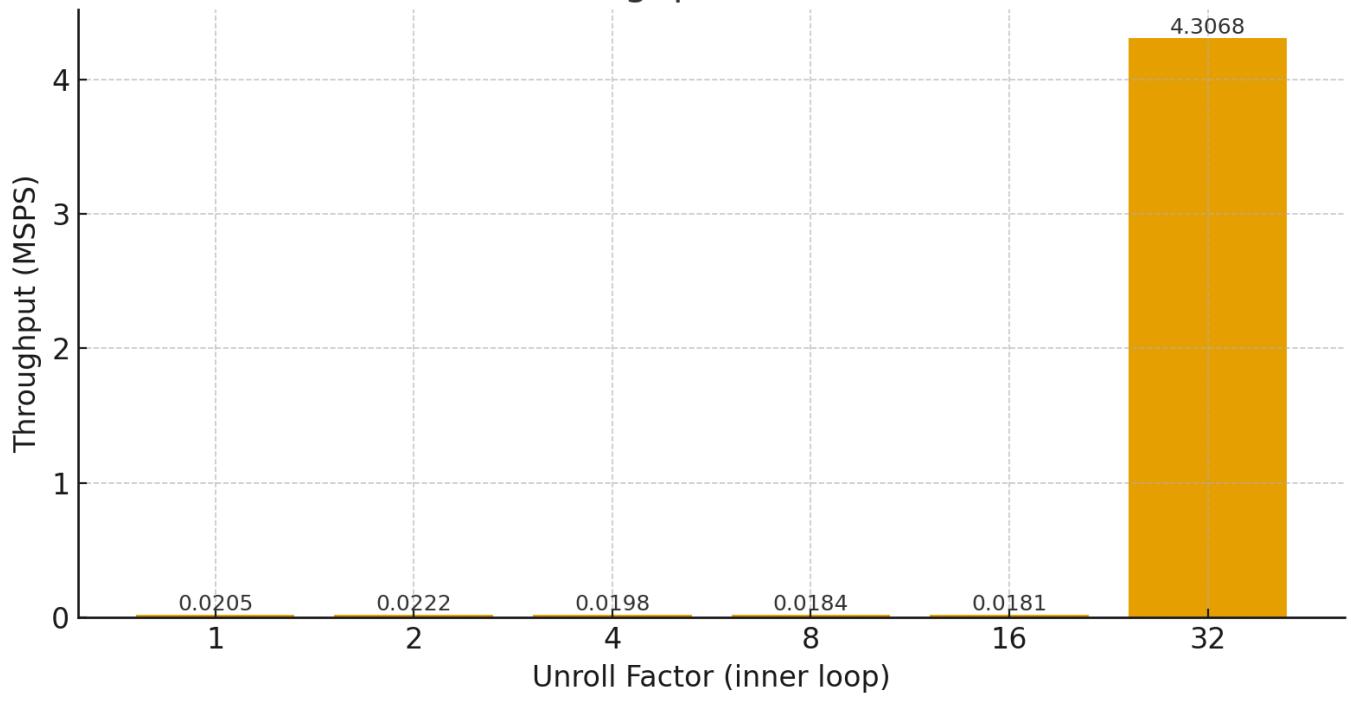
UNROLL FACTOR	EST. CLOCK (NS)	LATENCY (CYCLES, MAX)	ACHIEVED II (CYCLES)	THROUGHPUT (MSPS)	BRAM_18K	DSP	FF	LUT
1	7.256	6721	6722	0.02050	64	5	4,904	1,596
2	7.256	6209	6210	0.02219	64	5	5,195	1,621
4	8.495	5953	5954	0.01977	64	7	5,727	2,295
8	9.305	5825	5826	0.01845	64	7	6,259	2,413
16	9.591	5761	5762	0.01809	64	7	8,507	4,339
32	7.256	203	32	4.30678	64	640	61,509	93,250

## (b) Resource utilization
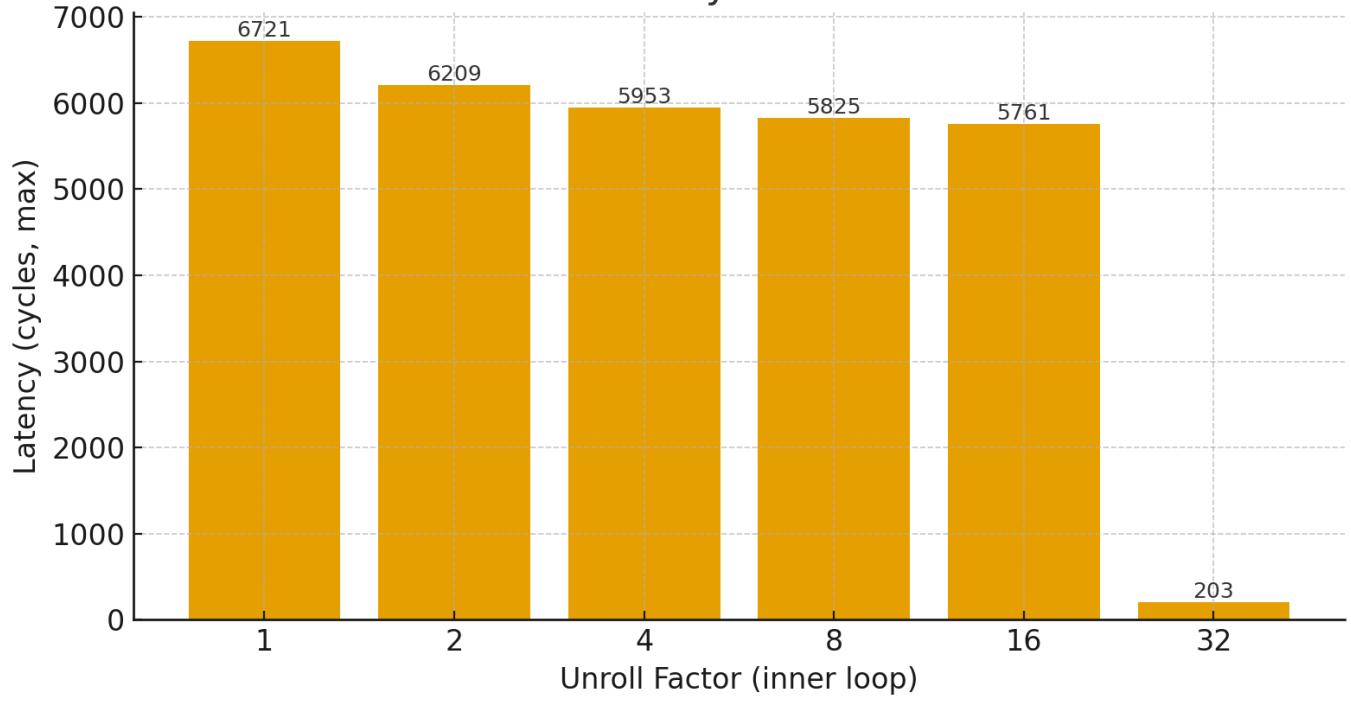
(c) Throughput & latency

DFT32 Throughput vs Unroll Factor

DFT32 Latency vs Unroll Factor

Q6 DFT1024 Baseline

(a) sin/cos

EST. CLOCK (NS)	LATENCY (CYCLES)	INTERVAL (CYCLES)	THROUGHPUT (MSPS)	BRAM_18K	DSP	FF	LUT
7.297	102,774,785	102,774,786	1.333e-6	16	220	14,764	17,815

Q6(b) — DFT-1024 with 1-D LUT

EST. CLOCK (NS)	LATENCY (CYCLES)	INTERVAL (CYCLES)	THROUGHPUT (MSPS)	BRAM_18K	DSP	FF	LUT
7.297	8,388,631	8,388,637	0.000016	4	14	2688	2517

Q7 DFT1024 Loop Interchange

EST. CLOCK (NS)	LATENCY (CYCLES)	INTERVAL (CYCLES)	THROUGHPUT (MSPS)	BRAM_18K	DSP	FF	LUT
7.297	1,048,603	1,048,602	0.000131	4	56	5397	7380

Q8 DFT1024 Best Design

Q8(a) — Methodology

- **Loop interchange**: outer `j`, inner `i` to eliminate inner-loop write hazards and enable a clean inner pipeline.
- **1-D LUT twiddles**: `k = (i*j) & (SIZE-1)` → use `cos[k]` and `sin[k]` (forward

DFT sign). Avoids 1M sin/cos calls.

- **Fixed-point arithmetic**: `DTYPE = ap_fixed<45,20>` to cut DSP usage vs float/double while keeping precision.

- **Moderate memory parallelism**: `#pragma HLS array_partition ... cyclic factor=5` on inputs, outputs, and LUTs to feed a 5-wide datapath.

- **Balanced inner parallelism**: `#pragma HLS pipeline II=1` + `#pragma HLS unroll factor=5` **on the inner loop** for ~1 result update/cycle at 5 lanes, without blowing the DSP budget.

- **Power-of-two fast modulo**: `k = (i*j) & (SIZE-1)` (SIZE=1024) for free index wrap.

Rationale: unroll=5 + cyclic factor=5 gives enough ports for II=1 per lane while keeping **DSP < 220** (PYNQ-Z2 limit), and fixed-point keeps resources comfortable.

Q8(b) — Best design results (fits PYNQ-Z2)

Throughput (MSPS) = 1000 / (Estimated clock ns) / (Interval cycles)

EST. CLOCK (NS)	LATENCY (CYCLES)	INTERVAL (CYCLES)	THROUGHPUT (MSPS)	BRAM_18K	DSP	FF	LUT	NOTES
6.562	209,950	209,949	0.0007259	60	188	18,271	13,071	All < 100% utilization

- **Fits PYNQ-Z2** (availability: BRAM_18K=280, DSP=220, FF=106,400, LUT=53,200):
    - BRAM ≈ **21%**, DSP ≈ **85%**, FF ≈ **17%**, LUT ≈ **25%**.
- Throughput in **Hz**: 0.0007259 MSPS × 1e6 ≈ **726 samples/second**.

Q9 — DFT1024 Streaming

EST. CLOCK (NS)	LATENCY (CYCLES)	INTERVAL (CYCLES)	THROUGHPUT (MSPS)	BRAM_18K	DSP	FF	LUT
6.851	2,099,227	2,099,228	0.00006953	10	18	2693	2680