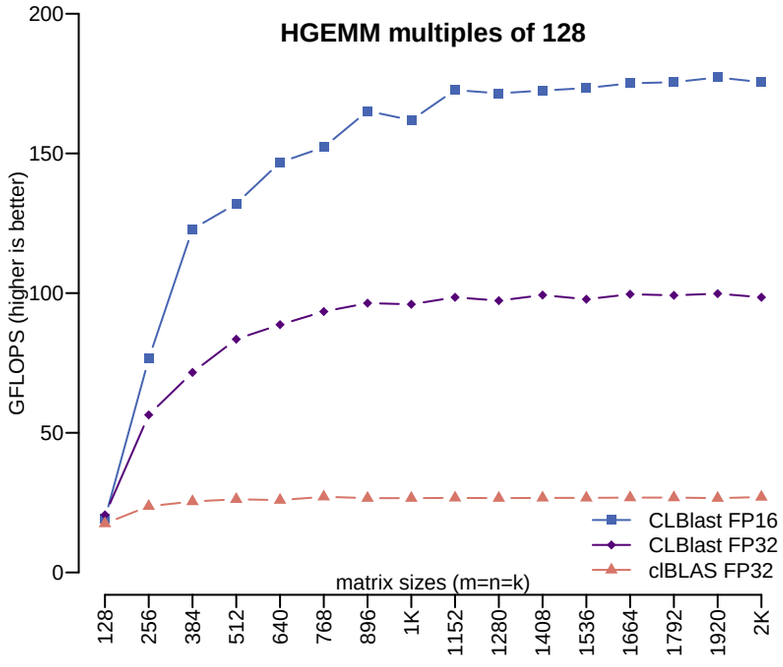
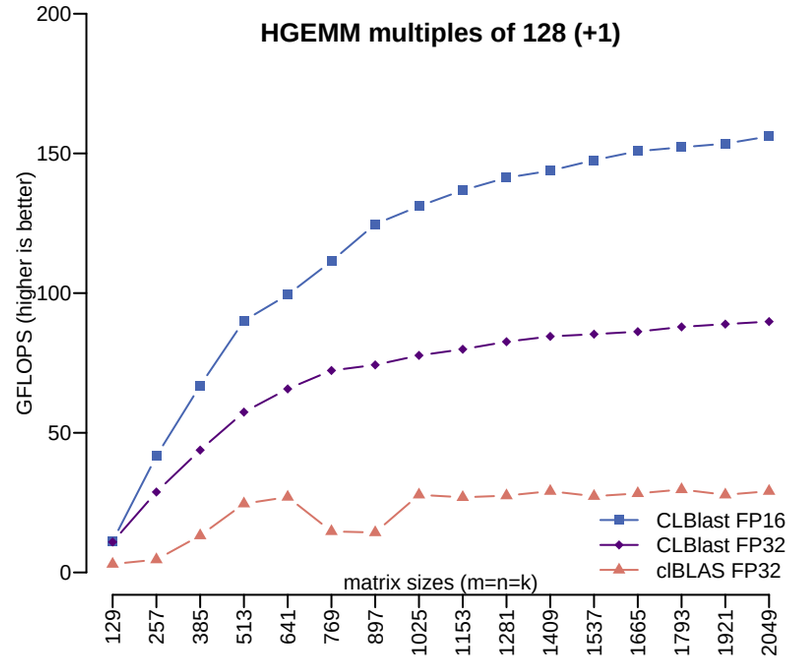
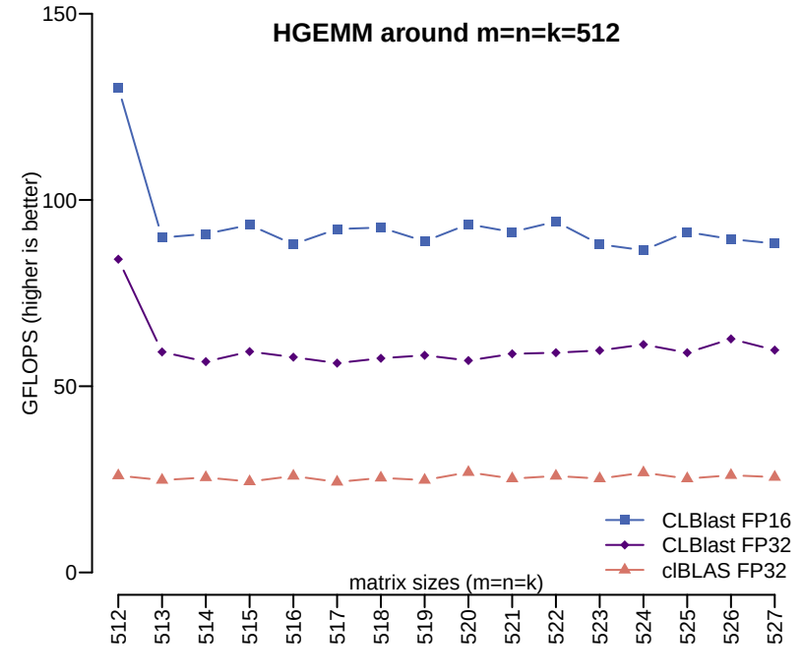
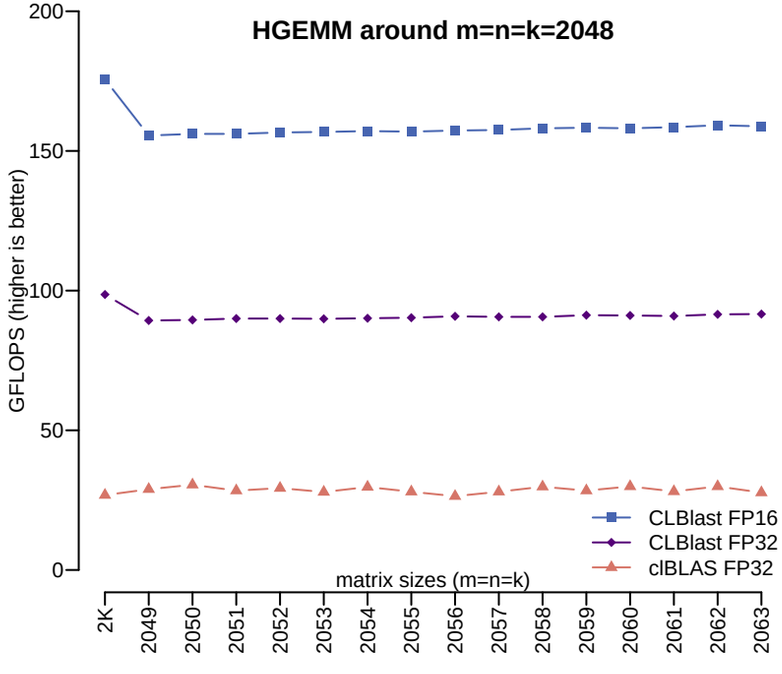
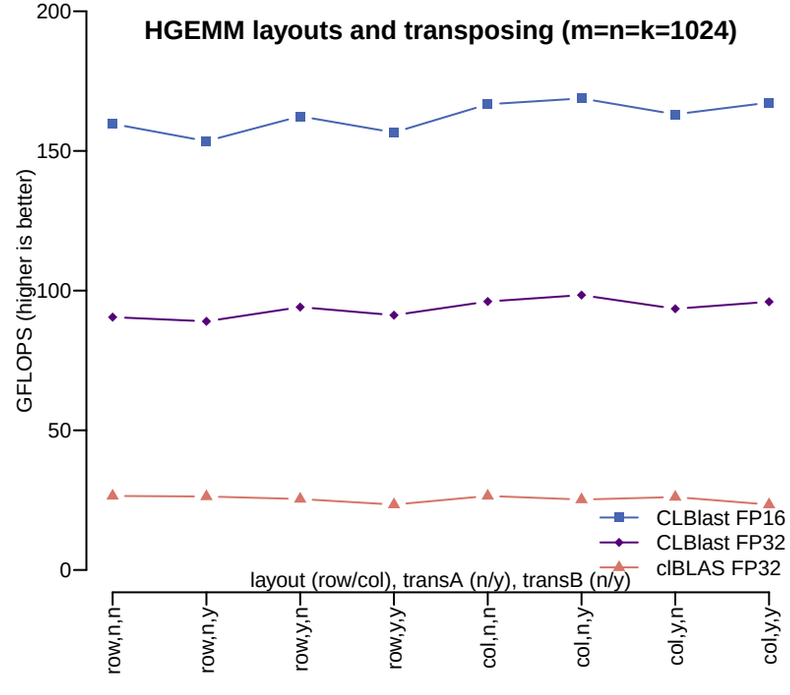
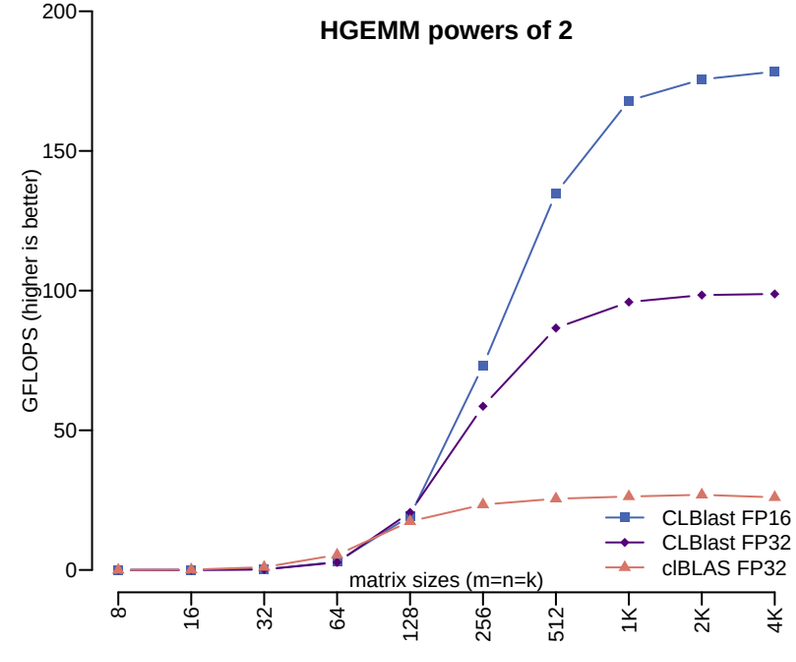

HGEMM multiples of 128
GFLOPS (higher is better)
0
50
100
150
200
matrix sizes (m=n=k)
128
256
384
512
640
768
896
1K
1152
1280
1408
1536
1664
1792
1920
2K
CLBlast FP16
CLBlast FP32
clBLAS FP32


HGEMM multiples of 128 (+1)
GFLOPS (higher is better)
0
50
100
150
200
matrix sizes (m=n=k)
129
257
385
513
641
769
897
1025
1153
1281
1409
1537
1665
1793
1921
2049
CLBlast FP16
CLBlast FP32
clBLAS FP32


HGEMM around m=n=k=512
GFLOPS (higher is better)
0
50
100
150
matrix sizes (m=n=k)
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
CLBlast FP16
CLBlast FP32
clBLAS FP32


HGEMM around m=n=k=2048
GFLOPS (higher is better)
0
50
100
150
200
matrix sizes (m=n=k)
2K
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
CLBlast FP16
CLBlast FP32
clBLAS FP32


HGEMM layouts and transposing (m=n=k=1024)
GFLOPS (higher is better)
0
50
100
150
200
layout (row/col), transA (n/y), transB (n/y)
row,n,n
row,n,y
row,y,n
row,y,y
col,n,n
col,n,y
col,y,n
col,y,y
CLBlast FP16
CLBlast FP32
clBLAS FP32


HGEMM powers of 2
GFLOPS (higher is better)
0
50
100
150
200
matrix sizes (m=n=k)
8
16
32
64
128
256
512
1K
2K
4K
CLBlast FP16
CLBlast FP32
clBLAS FP32