Conclusion:

There are 3 varying factors: input size, cutoff and degree of parallelism

Let’s keep 2 constant and change one to find its relationship with time.

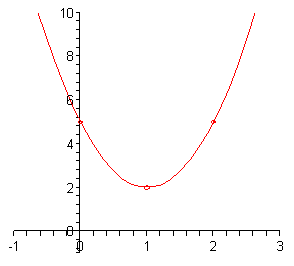
1. Change input size, keep cutoff and DoP constant (64)

Increase input size saw almost a linear increase in time

1. Change cutoff for constant input size and DoP (16)

Cutoff shouldn’t be too small neither too large

I guess the relationship here is hyperbolic. Something like,



y = ax^2+bx+c where b^2-4ac<0

1. Change DoP, keep cutoff and input size constant

Same case for changing DoP, you don’t want too many parallel threads running to create a bottleneck situation. Speed up plateaus after a certain DoP.

We might pin down the DoP for the lowest time however.

My best guess is for a given value of cutoff, lowest time is for

DoP = 2^(inputsize/cutoff -1)

1. As input increases we would increase DoP at bigger rate than cut off because cutoff brings steeper change in time