

Implementation of multi-GHz digital shaper for high-rate nuclear spectroscopy

Introduction

The trapezoidal filter is considered the optimum filter for measuring the energy of particles in nuclear spectroscopy applications due to its superior performance in extracting energy information from detected signals while minimizing noise contributions. The trapezoidal filter achieves this by leveraging two primary characteristics: pulse shape and time-domain filtering:

Pulse shape: The trapezoidal filter, has a trapezoidal pulse shape, which is characterized by a flat top and linear rise and fall times. The flat top of the trapezoid provides a stable and consistent region for energy measurement, which reduces the influence of trigger jitter on the energy estimation.

Time-domain filtering: The trapezoidal filter operates in the time domain and offers a trade-off between energy resolution and processing time. By adjusting the filter's shaping time constants, it is possible to find an optimal balance between noise reduction and throughput. Longer shaping times result in improved energy resolution, as they allow for better noise suppression, particularly for low-frequency noise components such as 1/f noise and baseline drifts. However, longer shaping times can lead to pulse pile-up.

When the particle rate is very high, and the detector produces short signals, it is crucial to have a high-speed filter for several reasons:

Accurate energy measurement: High-speed filters can more accurately capture the short signals generated by the detector, ensuring that energy measurements are precise and reliable. If the filter operates at a slower speed, it may not be able to effectively process the fast signals because the peak of the short exponential signal will be missed, leading to potential inaccuracies or loss of information.

Reduced pulse pile-up: In high-rate environments, closely spaced pulses can overlap or "pile up," making it difficult to separate individual pulse events and extract accurate energy information from each pulse. A high-speed filter on the trigger path can better resolve closely spaced pulses, reducing the probability of pulse pile-up and enabling more accurate energy measurements. On the energy path a short filter with a relevant number of samples allows to easy recover piled up event increasing the throughput without a significant reduction in resolution.

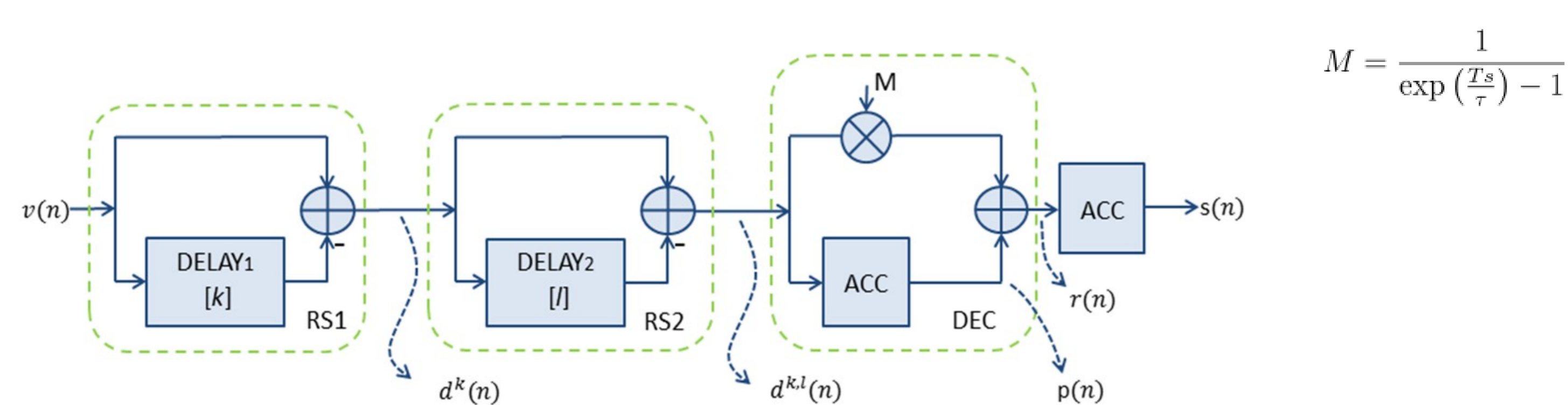
Enhanced noise suppression: at equal shaping time, an high-speed filters exhibit better noise suppression capabilities (there are more points to average the noise), particularly for high-frequency noise components, leading to more accurate energy measurements and better trigger performance on very low energy.

Parallel implementation of trapezoidal filter

The implementation of the trapezoidal filter in FPGA devices has been known in the literature for years [Unfolding-synthesis technique for digital pulse processing. Valentin T. Jordanov].

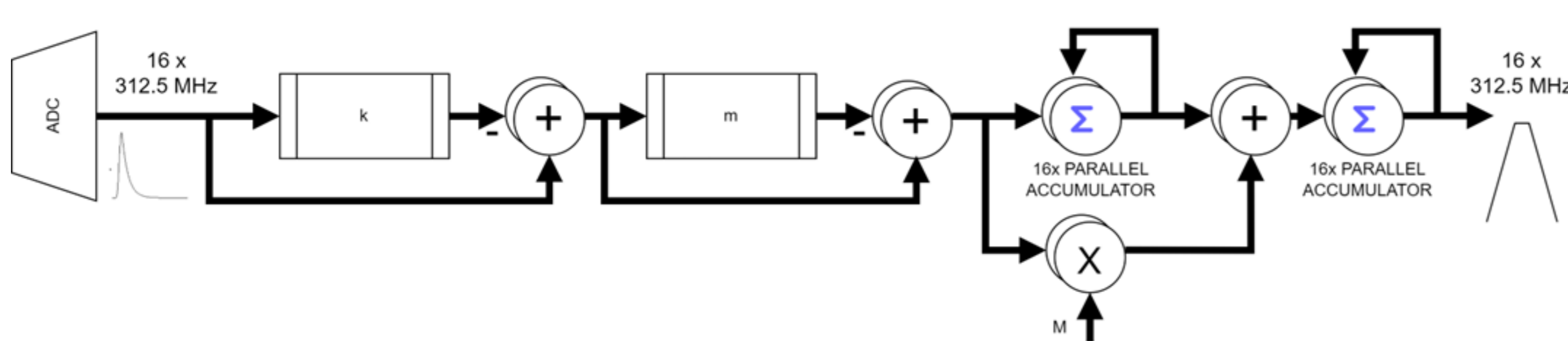
The recursive implementation allows for real-time processing of the input data to the ADC, sample by sample, using a single filter to perform both the deconvolution of the exponential signal and the convolution with the filter with a trapezoidal impulse response.

The following image depicts the processing schematic of the filter. The delay elements define the shaping time (k) and the flat top (l-k). The multiplication constant M defines the deconvolution constant, given by $M = \frac{1}{\exp(\frac{Ts}{\tau}) - 1}$, where Ts is the sampling period and τ is the time constant of the exponential signal.



Although it may appear as an FIR filter, within the filter are two accumulators.

In order to work with detectors whose pre-amplified signals are very fast exponentials (<20 ns decay constant), there arose the need to develop a new series of digital signal processors whose ADCs operate at frequencies exceeding 1 Gsps, even reaching up to 10 Gsps. Often, signal processing is done offline, with the digitizer simply saving the waveforms and then processing them on a PC. The technique proposed is an improvement on the classic implementation of the trapezoidal filter, allowing for parallel processing of data in order to achieve multi-GHz throughput. Obviously, it is unfeasible to think of using a 5 GHz processing clock in the FPGA if the sampler operates at 5 Gsps. Therefore, a parallel development of the various elements of the trapezoidal filter was chosen in order to keep the power dissipation in check and, even with a complex firmware architecture, to meet the timing requirements of the FPGA device.



Implementation of a trapezoidal filter exploiting a parallel architecture in the FPGA device. Each element in the pipeline as been replicated by a factor of 16 in order to parallel process the data produced by the ADC converter. The diagram in figure implement both the deconvolution of the exponential signal to a delta and the trapezoid.

The underlying idea of the implementation shown in the figure is that the signal from the fast ADC (i.e., 5GHz) is provided over multiple parallel buses N (for example, 16) at a word rate of F_s/N (5GHz/16 = 312.5 MHz). At this frequency, it becomes possible in an FPGA to operate on the data, performing even complex operations, in parallel. The graph clearly shows how all the elements are replicated for N. There are N delays k, N delays m, N adders, and N multipliers. In fact, the FIR part of the trapezoidal filter is easily parallelized by multiplying the computational elements by N. However, the accumulators are much more complex to parallelize.

Pipelined implementation of accumulator

In a trapezoidal filter, the use of accumulators with a large number of bits is essential for several reasons. Accumulators often manage small incremental values, and having more bits ensures these minute values aren't lost due to truncation, preserving the precision of the accumulated signal. As accumulators sum up values over time, a large bit-count prevents overflow, especially if the processed signal has a vast dynamic range or during long integration intervals. Lastly, a higher number of bits provides a greater resolution, which is crucial when the differences between signal values are minimal or in contexts like high resolution spectroscopy. Accumulators, especially those with a large number of bits, can introduce timing challenges when implemented in FPGA architectures. Accumulators, by nature, involve addition operations. As the number of bits increases, the adder's complexity, required for the accumulator, increases. Pipelining accumulators can be tricky. An accumulator inherently relies on feedback, as the accumulated value from the previous operation (or clock cycle) is used along with the new value for the next addition.

The proposed (novel) implementation calculate partial sum, using a pipelined architecture. We consider for simplicity a parallelization factor equal to 4 (allowing to operate at 1.25 GHz with 312.5 clock). The architecture produce 4 sums per clock cycle.

$$\begin{aligned} Y_0(n) &= Y_0(n-1) + X_0(n); & Y_1(n) &= Y_1(n-1) + s_0(n); \\ Y_2(n) &= Y_2(n-1) + s_1(n); & Y_3(n) &= Y_3(n-1) + s_4(n); \\ s_1(n) &= X_0(n) + X_1(n); & s_2(n) &= s_1(n) + X_2(n); \\ s_3(n) &= X_2(n) + X_3(n); & s_4 &= s_1(n) + s_3(n) \end{aligned}$$

The partial sum s1, s2, s3, and s4 can be calculated using pipelined adders. Because they are in the "FIR" part of the accumulator, their pipelined can be optimized to close the FPGA timing. It's important to balance the respective pipeline delays of each node of the accumulator in order to maintain true the mathematical expressions.

Sampling Jitter in random distributed pulses

The exponential pulses produced by the detector are not typically in any way correlated with the sampling comb of the acquisition card. This results in the exponential signal being effectively sampled with a probability evenly distributed between $-T_s/2$ and $T_s/2$. The error is minimal when the peak of the exponential coincides with a sampling point, and it is maximal when the peak occurs shortly after sampling. Naturally, this introduces into the digital samples an error proportional to T_s and inversely proportional to the decay time (τ) of the exponential. When working with detectors whose preamplifier produces long tails, the effect is generally negligible. In this publication, we consider working with fast detectors, where typically the τ of the preamplifier is 10-20ns.

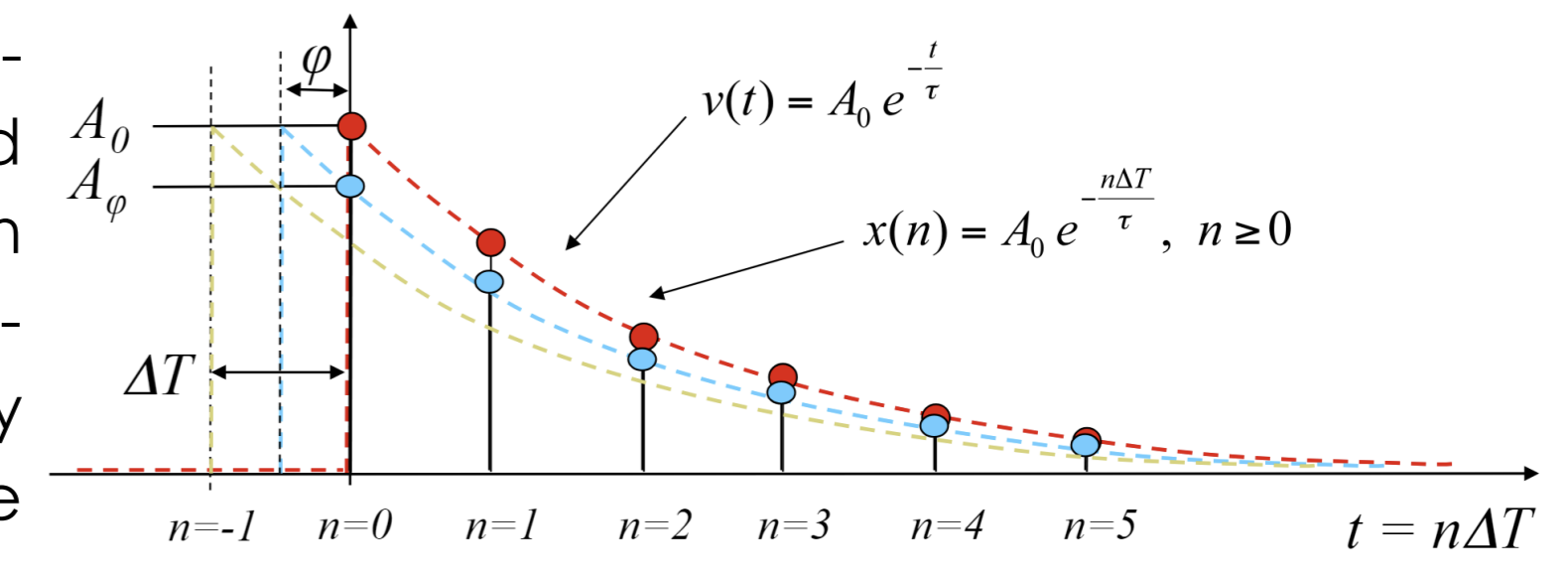
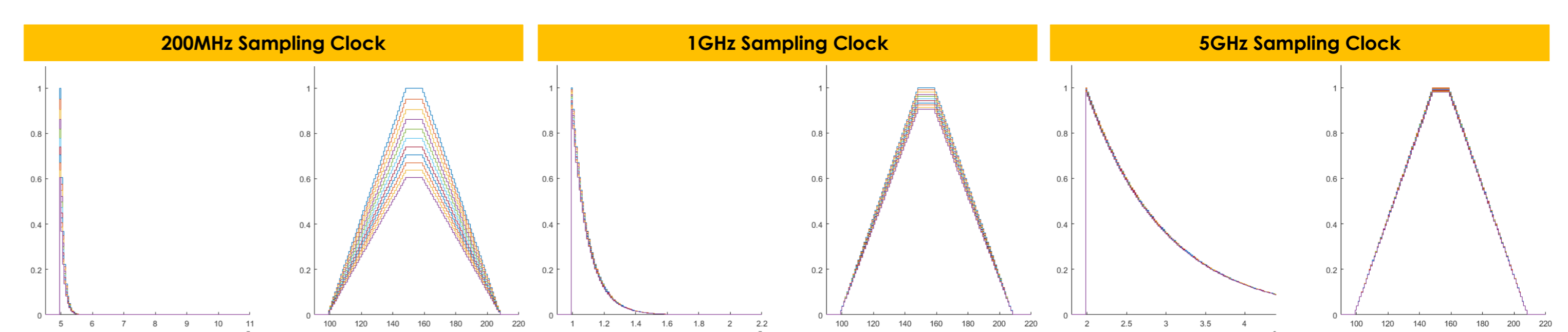


Image credit: lectures of Prof. Valentin T. Jordanov.
Effect of random pulse sampled with fixed sampling period. Exponential amplitude is affected by random phase

$$\begin{aligned} A(\phi) &= A_0 \cdot \exp\left(-\frac{\phi}{\tau}\right); & e(\phi) &= A_0 - A(\phi); \\ e(\phi) &\approx A_0 \cdot \frac{\phi}{\tau} \\ \text{mean: } m_\phi &= -\frac{A_0 \cdot T_s}{2\tau} \\ \text{variance: } \sigma_\phi^2 &= \int \left(\mu - \frac{A_0 \cdot T_s}{2\tau}\right)^2 p(u) du = \frac{A_0^2 \cdot T_s^2}{\tau^2} \\ \sigma(\phi) &\approx 0.29 \frac{A_0 \cdot T_s}{\tau} \end{aligned}$$

where $p(u)$ is the white distributed probability of sampling A between A_0 and $A_0 \cdot T_s/\tau$



	10ns			25ns		
	A (% pp)	E (% pp)	$E_n \sigma^2$	A (% pp)	E (% pp)	$E_n \sigma^2$
200 M	49	49	4,20E-04	19	19	1,50E-04
1 G	10	10	3,03E-04	4	4	4,7,80E-06
2.5 G	4	4	9,60E-06	1,6	1,6	2,50E-06
5 G	2	2	3,20E-06	0,8	0,8	1,50E-06

The table on the left, and the plots above show the error on the estimation of the peak A (%pp), the error on the energy E (%pp). These measurements are calculate peak-to-peak because the probability to measure the maximum or the minimum is white and not gaussian. We consider two different decay time (10ns and 25ns), sampled with 200 Msps, 1Gsps, 2.5Gsps and 5 Gsps. We fixed the shaping time of the trapezoidal to 30ns and we measured the sigma of the noise at different sampling rate.

Applications—Fast Diamond Detectors

Diamond, with its exceptional physical properties, is increasingly recognized as a robust radiation detector material. This unique crystalline structure is particularly well-suited for challenging radiation environments and elevated temperatures. Its sensitivity extends across a diverse spectrum of radiation, including charged particles, neutrons, and photons. This versatility has enabled diamond detectors to play pivotal roles in particle accelerators as beam loss monitors, in Synchrotron Light Sources for advanced photon detection, and notably in thermal neutron fields for precise neutron diagnostics. Yet, one of the most promising and vital applications lies in its ability to monitor the fusion processes, specifically Deuterium-Deuterium (D-D) and Deuterium-Tritium (D-T) fusion reactions.

The short signal duration or pulse width observed in diamond detectors is largely attributed to the rapid collection of charge carriers generated when radiation interacts with the diamond lattice. This quick response minimizes the risk of pulse pile-up, where multiple radiation interactions could produce overlapping signals in detectors with longer response times. When dealing with high particle fluxes, the distinction between individual radiation events becomes paramount to obtain accurate readings and understand the underlying processes.

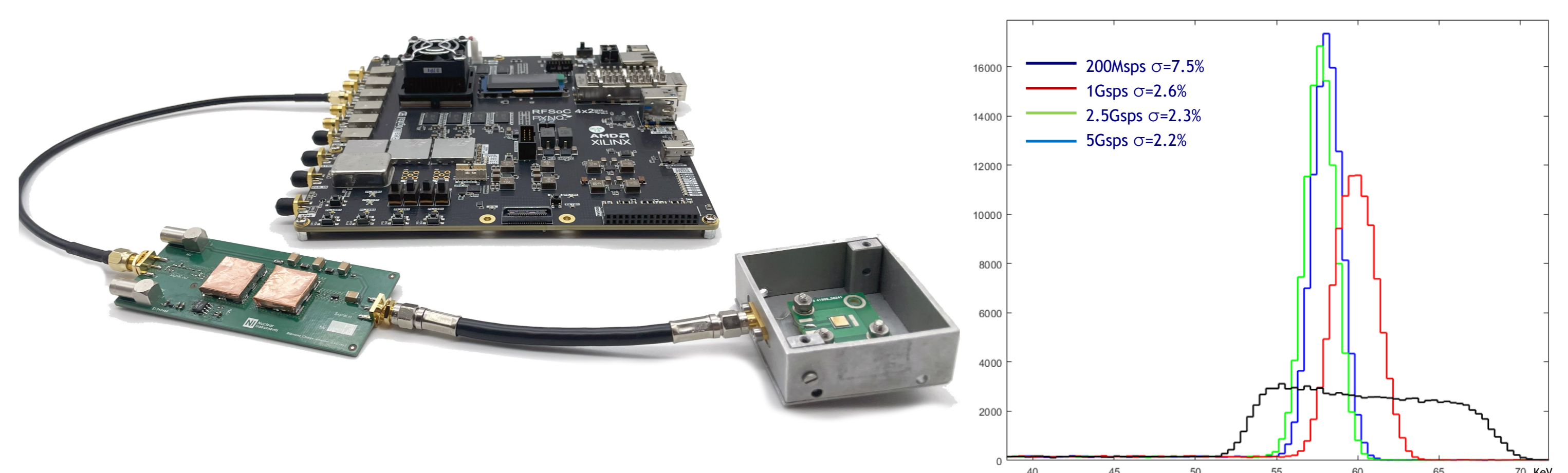
Furthermore, the high mobility of charge carriers in diamond, combined with its minimal charge carrier trapping, ensures that the detector can reset swiftly between detections. This rapid reset capability enhances the detector's ability to measure high fluxes by reducing the time window for noise introduction, thereby maintaining an excellent signal-to-noise ratio even under high rate environments.

In the realm of fusion reactions, where particle fluxes are exceedingly high, the short signal response of diamond detectors offers a critical advantage. Not only does it provide clear and distinct signals for each fusion event, but it also paves the way for real-time monitoring, enabling instantaneous adjustments and optimizations in fusion reactors.

In this study, the high-speed trapezoidal filter algorithm was developed using high-level synthesis (HLS) to enable efficient hardware implementation on FPGA devices. HLS allows for the design of hardware accelerators using high-level programming languages such as C++.

it was implemented on a Zynq RFSoc development board. The Zynq RFSoc is a versatile and powerful platform that combines the processing capabilities of a high-performance ARM-based processing system with the flexibility of programmable logic, and 4 channels 5 GSps 14 bit ADC, making it an ideal choice for implementing the high-speed trapezoidal filter algorithm. The development board used in this study is the PYNQ-ZU RFSoc board.

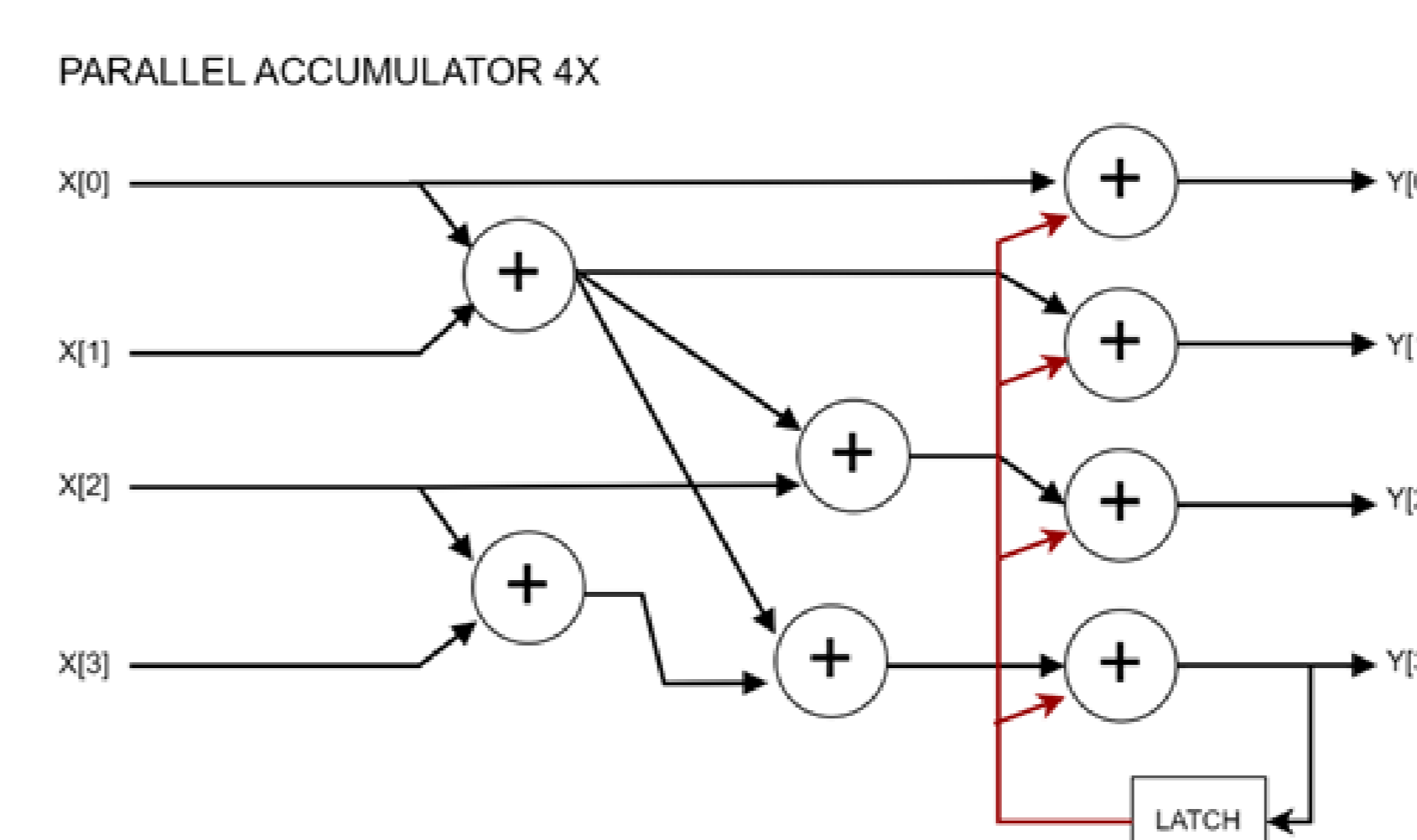
We tested the processing system with a diamond detector and a custom designed ultra fast charge pre-amplifier (with 20ns decay time) designed to operate in fusion reactor.



Zynq Pynq RFSOC connected to our pre-amplifier and diamond detector. Sample date has been decimated to 5G, 2.5G, 1G and 200Msps in order to evaluate the improvement in increasing the sample rate. It appear that a trade of between resolution ad FPGA resources can be 2.5 Gsps where on the 59KeV the degradation on the sigma is about the 0.1%.

The only critical part of the pipelined accumulator is the last adders. They need to produce the result in a single clock cycle.

On modern FPGA the single adder core on 64 bits can operate at 350 MHz or more. It's important to note that the implementation of the pipelined accumulator produces data with a pipeline delay > 1 (unlike the classic accumulator). The implementation with N=4, for instance, introduces a pipeline delay of 3. It's essential to be aware of this delay within the algorithms that use the accumulator.



Implementation of the parallel pipelined accumulator. The accumulator is the most critical part in the filter design because it need the result at t-1 to calculate the sample at t. Except the last adder, all the circuit can be pipelined with arbitrary depth. In the design care should just been taken to compensate pipeline delay. The last sum on all parallel path must be calculated in a single clock cycles. With Zynq Ultrascale+ the circuit can operate up to 330MHz with 64 bit accumulators

Different version of the design has been synthesized on Zynq Ultrascale+ RFSoc XCZU48DR at different level of parallelization to operate from 200Msps up to 5 Gsps. Also the resource usage may seems important, we should consider that this kind of devices has a very big reconfigurable area and the overall usage is below the 1% per channel

	Trapezoidal Filter resource usage			
	LUT	FF	DSP	BRAM
200M	1602	1405	5	6.5
2.5G	3581	5620	24	16
5G	9111	17540	48	32
	Parallel Accumulator			
	LUT	FF	DSP	BRAM
2.5G	620	876	0	0
5G	1802	2330	0	0