Evaluation of a Modular PET System Architecture with Synchronization over Data Links

Ramón J. Aliaga, Vicente Herrero-Bosch, Jose M. Monzo, Ana Ros, Rafael Gadea-Gironés, and Ricardo J. Colom

Abstract—A DAQ architecture for a PET system is presented that focuses on modularity, scalability and reusability. The system defines two basic building blocks: data acquisitors and concentrators, which can be replicated in order to build a complete DAQ of variable size. Acquisition modules contain a scintillating crystal and either a position-sensitive photomultiplier (PSPMT) or an array of silicon photomultipliers (SiPM). The detector signals are processed by AMIC, an integrated analog front-end that generates programmable analog outputs which contain the first few statistical moments of the light distribution in the scintillator. These signals are digitized at 156.25 Msamples/s with free-running ADCs and sent to an FPGA which detects single gamma events, extracts position and time information online using digital algorithms, and submits these data to a concentrator module. Concentrator modules collect single events from acquisition modules and perform coincidence detection and data aggregation. A synchronization scheme over data links is implemented that calibrates each link's latency independently, ensuring that there are no limitations on module mobility, and that the architecture is arbitrarily scalable. Prototype boards with both acquisition and concentration functionality have been built for evaluation purposes. The performance of a small PET system with two detectors based on continuous scintillators is presented. A synchronization error below 50 ps rms is measured, and energy resolutions of 19% and 24% and timing resolutions of 2.0 ns and 4.7 ns FWHM are obtained for PMT and SiPM photodetectors, respectively.

Index Terms—AMIC, clock distribution, data acquisition (DAQ), modular electronics, positron emission tomography (PET), self-calibration, serial links, silicon photomultiplier (SiPM), synchronization.

I. INTRODUCTION

TIME coincidence resolution is one of the most critical aspects of PET systems. Traditionally, finer resolutions have allowed a tightening of the coincidence window for event acceptance, yielding a better noise equivalent count rate (NEC) for cleaner reconstructed images. In the last decade, advances in scintillators have sparked a renewed interest in time-of-flight (TOF) PET systems [1]–[3], where time difference is used to estimate the radiotracer position along the line of response. This imposes a much more stringent limitation on coincidence resolution: a figure of 600 ps FWHM is estimated as the bare minimum for a modern TOF detector [4], and 500 to 600 ps are realized by the current generation of commercial full-body TOF PET scanners [1], [5]. Better resolutions are currently

achieved in finely tuned experimental setups [6] that can lead to even greater gains [5].

Coincidence is given by the difference between timestamps assigned to single gamma events on different acquisition boards in the system. Synchronization errors between timestamping boards need to be well below the coincidence resolution, since any inconsistency between reference times on acquisition nodes is directly reflected on the measured difference. Hence, it is mandatory to establish an accurate system-wide synchronization scheme that can guarantee timing mismatches on the order of 100 ps.

The typical synchronization method for PET and other high energy physics readout systems is based on a clock tree, using zero-delay clock buffers and distributors to span the whole system. Clock trees work best in a cableless environment, where all timestamping electronics are placed within the same crate or otherwise physically constrained, with clock distribution being implemented through controlled backplane connections [7]–[10]. However, this condition severely limits DAQ scalability and the mobility of the detectors, often hardwiring the maximum amount of supported detectors and forcing a hardware redesign for system expansions or whenever the detector topology is changed. In particular, cableless systems are not adequate in cases such as PET detectors with variable geometry [11], [12] and dual PET/MR systems where electronics separation is mandatory [13].

Systems where cables or fibers are necessary may use them for two different purposes: either as a means of transport for detector signals from the front-end to a centralized digitizing DAQ or as part of the clock distribution subsystem. In the former case, using long cables to transport sensitive sensor outputs degrades signal quality and timing [14], [15] and may require channelwise time alignment depending on implementation [16]. The latter case, on the other hand, requires transmission lines to be matched in length in order to obtain a fully balanced clock tree [17]–[19], which can be difficult when the number of acquisition nodes to be synchronized is large [20], forcing additional global timing calibrations [21]. Further problems arise when there is a large distance between system nodes, as fluctuating operating conditions such as temperature cause a variation of the delay of long lines [22].

In order to overcome these issues, a synchronization scheme over data links was proposed in [23] that was able to achieve state-of-the-art synchronization resolution. The idea of individually self-calibrating, bidirectional links has been applied recently following the same principle [24]–[26] or variations thereof [22], [27], [28] in the generic setting of large scale, distributed readout systems for physics experiments, with the
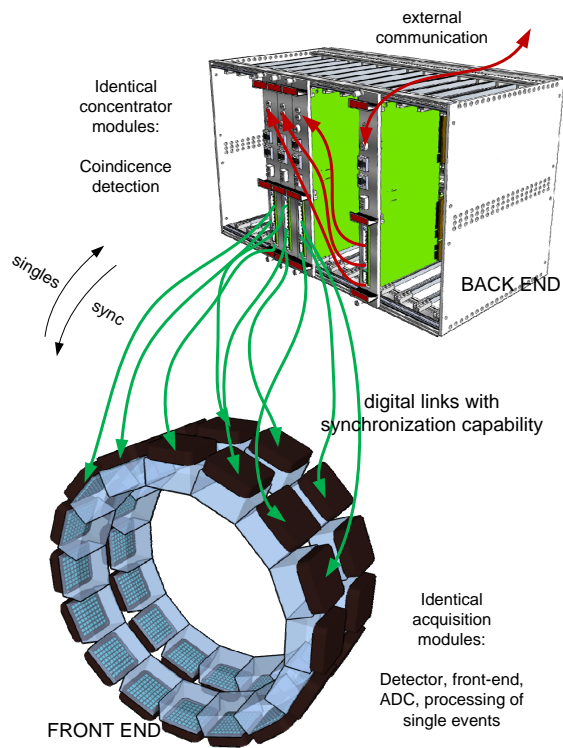
Fig. 1. Logical architecture of the DAQ system. Arrows represent digital links: cables between back-end and front-end, and either cables or backplane connections between back-end modules. Frequency and synchronization are propagated in the direction of the arrows.

main goal of compensating for drifts in the propagation delay of long transmission lines. However, this clocking method carries additional benefits for smaller systems like PET, namely: vastly reduced calibration needs, deterministic latency, and detector mobility with simplified cabling. A modular, scalable DAQ architecture for PET was thus proposed in [23] taking advantage of all these features. In this paper, the first working prototypes are described and their performance is evaluated.

II. HARDWARE ARCHITECTURE

The proposed DAQ arquitecture is outlined in Fig. 1. The system is logically divided into a front-end section and a back-end section, and each of them is formed by an arbitrary number of respectively identical modules with no constraints on physical location with respect to each other, i.e. the only placement restriction is given by the particle detectors at the front-end. The front-end section consists of acquisition modules, which contain the photodetectors and perform analog conditioning of detector signals, digitization, single event detection and position and timestamp extraction. The back-end section contains concentrator modules, which collect data from several acquisition modules and perform time coincidence detection. Data from several concentrators can themselves be collected by a higher-level concentrator with identical hardware. All modules are connected forming a hierarchic tree, with the top node being responsible for the transmission of all aggregated data to the external processor that handles the image reconstruction.

All connections between modules are purely digital, self-calibrating, full-duplex data links with an embedded clock that can be implemented either in a backplane or using cables, thus placing no restrictions on mobility. Each link's ends are regarded as master or slave according to the global system hierarchy. The links serve three different purposes:

- *Data transmission:* Single event and coincidence data are transmitted upward, and configuration commands are sent downward.
- *System frequency propagation:* The slave recovers the clock frequency from the data link and uses it for its own downlink transmissions and/or digitizing circuitry. Thus, the whole system is frequency-locked with the master oscillator which is located in the top concentrator module.
- *Synchronization:* Each link is capable of synchronizing the time reference for both nodes independently of all other module connections.

This architecture is arbitrarily scalable: no hard limit is imposed on the number of detectors or modules, not just at the design stage but also after the modules have been actually built. Soft limits are given by the degradation of synchronization resolution between acquisition nodes as the height of the hierarchy increases, and by the data link bandwidth, which has to be large enough to support the transmission of all coincidence data at the highest levels. Since all modules are identical copies of one of two different hardware designs and physically independent of each other, all boards are fully reusable in the case of system expansion (increase in the number of detector modules) or any other topology change.

III. IMPLEMENTATION

The preceding section concerns the generic DAQ architecture. In this section, a particular hardware implementation is described, focusing on acquisition modules. Let us remark that no specific concentrator module hardware has been designed yet; only their projected structure and role is described here. Instead of concentrators, acquisition boards with modified firmware have been used for validation purposes.

A. Acquisition Module

Figure 2 depicts a simplified diagram of the contents of a single acquisition module. Each module contains one gamma sensor consisting of a scintillating crystal coupled to a photodetector unit, which can be either a position-sensitive photomultiplier (PSPMT) or an array of silicon photomultiplier devices (SiPM). The high voltage source for the photodetector can be programmed in steps of 400 mV and 25 mV, respectively.

Photodetector outputs are sent to AMIC [29], an integrated analog front-end that converts 64 detector signals into 8 analog outputs, each of which is a weighted sum of the 64 inputs with digitally programmable coefficients. AMIC can thus be used as a replacement for a resistor network used as a charge division circuit for Anger-like logic [30], [31], benefiting from a higher bandwidth due to integrated preamplifiers, and the capability for automatic correction of photodetector channel gain mismatch by fine adjustment of weighted sum coefficients [32].
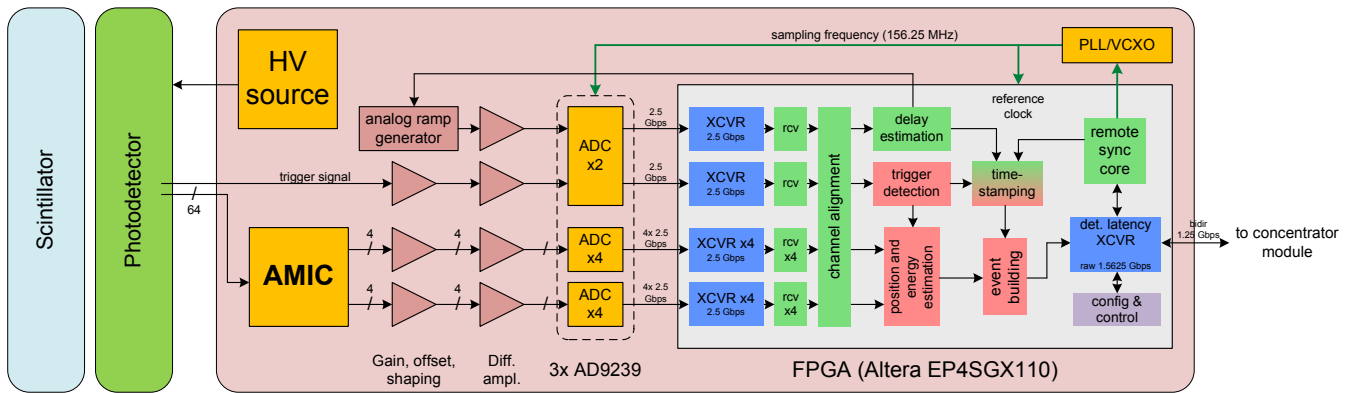
Fig. 2.   Simplified block diagram of a single acquisition module. Digital control lines are not shown.

Additionally, it can be used to obtain the first few statistical moments of the light distribution in a continuous scintillator [33], from which 3D event position can be extracted. In particular, the second moment contains information on depth of interaction within the crystal [34]. The newest version of AMIC is compatible with both PMT and SiPM-based detectors [35], so they can be used interchangeably by defining a common connector. Additionally, the AMIC architecture is fully expandable and allows the readout of detectors with more than 64 outputs by using several instances of AMIC and adding their corresponding current outputs together [36].

The resulting analog signals from AMIC go through a shaping and anti-aliasing stage using a second order, 30 MHz filter before being digitized by free-running ADCs. Shaped pulses have an approximate length of 150 ns. Signals are DC coupled, and a digitally controlled offset voltage is added to each channel in order to push the signal baseline as close as possible to the edge of the ADC input range, so as to maximize the dynamic range of detectable pulses. AD9239 12-bit converters by Analog Devices [37] are used with a 156.25 Msamples/s sampling rate. These are quad-channel ADCs with serial outputs, which help reduce the number of board components and digital traces, simplifying board layout and reducing signal integrity issues. A total of 10 ADC channels are used: eight for AMIC outputs, one for an additional fast trigger output from the detector, e.g. the last dynode signal from PSPMTs, and one for ADC delay calibration.

ADC outputs are read by a Stratix IV EP4SGX110 FPGA [38]; each data stream includes the actual samples and some framing overhead such as error-correcting codes. These serial signals have a data rate of 2.5 Gbps, so they must be read by embedded gigabit-speed transceivers. A channel alignment procedure is necessary after ADC frames are received and decoded, because the latency of the transceiver and frame decoding logic for each channel may be different. The alignment method is illustrated in Fig. 3. ADCs are configured to output fixed values, and all channels are simultaneously forced to transition between two known values. Data channels inside the FPGA are then selectively delayed so that the transitions are detected at the same clock cycle after the alignment logic.

From the sampled signals, single events are detected and time and position information is extracted online using purely
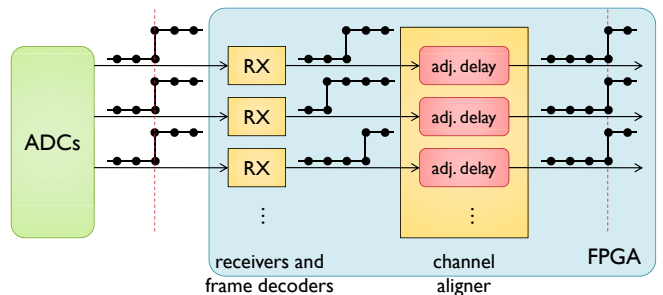

Fig. 3.   Channel alignment procedure inside the FPGA. ADCs are programmed to force a transition between two specific values, and that transition is time-aligned by channel-dependent adjustable delays.

digital algorithms. Each ADC channel can be used in either amplitude or charge mode, i.e. computations can be performed on the raw samples or on their accumulated value since the start of the current event, indicated by the trigger signal crossing a threshold value. The maximum value (i.e. pulse amplitude or charge) is recorded for each position channel. For timing, the Digital Constant Fraction Discriminator (DCFD) method [39] is applied to the trigger signal $t[n]$: the bipolar signal

$$b[n] = t[n] - A \cdot t[n-k] \tag{1}$$

is generated for programmable values of amplitude A and time shift k, and its zero-crossing point is estimated by linear interpolation on the clock interval where $b[n]$ changes its sign. By using either amplitude or charge signals, digital implementations of the CFD [40] and ARC (Amplitude and Rise Compensated) [41] methods are obtained, respectively; a graphical representation of both methods can be found in Fig. 4.

All data from a detected event is collected into a 160-bit frame containing a 48-bit timestamp with 1.5625 ps step size and a 12-bit value for each analog channel, and sent upstream to a concentrator module using an 8B/10B-encoded digital link with a net data rate of 1.25 Gbps.

B. ADC and FPGA Delay Compensation

Trigger signal delay from the digitization point to the timestamping algorithm block may be different for each acqui-
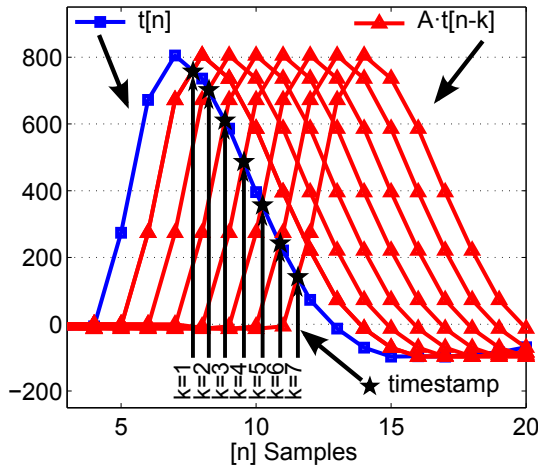
Fig. 4. Graphical depiction of the digital time pickoff methods for different time delays k. On the left, digital CFD for $A = 1$. On the right, digital ARC for $A = 2$ with the charge signal $q[n] = \sum_{j \leq n} t[j]$. The crosspoint between two linearly interpolated waveforms is used as the pulse timestamp.

sition module and hence must be taken into account in order to avoid timing errors for coincidence detection. Moreover, this delay may contain non-deterministic components such as ADC delay and FPGA transceiver latency, unless a specific deterministic latency mode is selected for the transceiver. In order to compensate for this effect, an analog linear ramp generator with digitally controlled charge and discharge signals is implemented on board and sampled by an ADC channel, with the goal of estimating the delay.

The FPGA continuously computes the linear regression coefficients corresponding to the last M samples from the ramp signal $y[n]$. If M is a power of two, this can be accomplished in an efficient way by iteratively computing the first two moments Y_0, Y_1 of the sample interval as

$$
\begin{aligned}
Y_0[n] &= Y_0[n-1] + y[n] - y[n-M] \\
Y_1[n] &= Y_1[n-1] + My[n] - Y_0[n]
\end{aligned}
\tag{2}
$$

and using them to obtain the instantaneous linear coefficients a_0, a_1 as

$$
\begin{aligned}
\frac{M(M+1)}{2}a_0[n] &= (2M-1)Y_0[n] - 3Y_1[n] \\
\frac{M(M^2-1)}{6}a_1[n] &= -(M-1)Y_0[n] + 2Y_1[n] .
\end{aligned}
\tag{3}
$$

Each delay estimation is obtained as follows: after fully discharging the ramp circuit, the charge signal is issued and the current signal baseline value $B = Y_0[n]/M$ is stored. Ramp start is detected when the linear coefficient a_1 crosses a certain threshold, and the number T of elapsed clock cycles is stored. At this point, the logic waits for M cycles until a linear fit $y[n] \approx a_1 n + a_0$ of the ramp waveform is obtained. The measured delay is then computed as

$$
D = T - \frac{a_0 - B}{a_1}
\tag{4}
$$

i.e. the time when the fitted ramp attains value B, using the charge signal trigger as the time origin. This measurement is repeated continuously, and a moving average of the last delay measures is used as the valid delay estimation. This delay
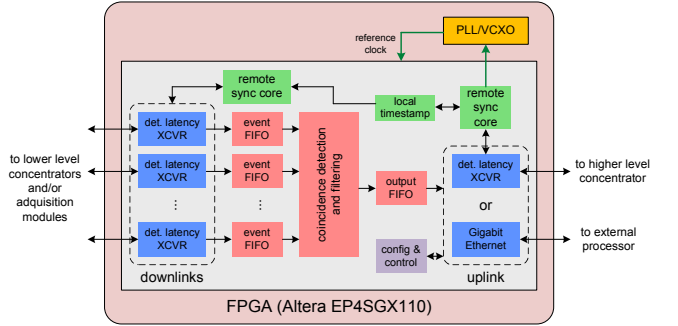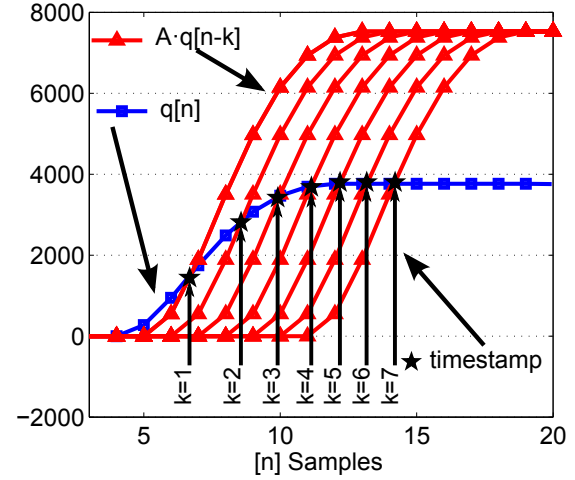

Fig. 5. Simplified block diagram of a concentrator module.

value is subtracted from the timestamp for all detected single events.

Notice that the delay value D contains not just the considered delay from ADC to timestamping logic, but also the propagation delay from the FPGA's delay estimation logic to the analog ramp circuit and to the ADC input; however, these additional components are considered to be equal in identical acquisition modules, so they are canceled when computing timestamp differences.

C. Concentrator Module

The main purpose of concentrator modules is to detect coincident gamma events and to keep its child nodes synchronized, i.e. acquisitors and lower level concentrators. A simplified scheme is shown in Fig. 5. Each concentrator hosts a number of links to lower level nodes where single event frames are received in chronological order and stored in FIFOs. At the top level concentrator, the coincidence detection engine continually compares the timestamps of the first available event from each detector, and checks whether the timestamp difference between the two oldest visible events is within the selected time coincidence window. If so, both events are registered as a coincidence; if not, the oldest one is discarded as a random event.

Lower level concentrators may simply aggregate all singles from different sources and transmit them upwards. The links described in section III-A can handle up to 7.8M singles per second, and a 4-fold increase in link data rate is possible without modifying the hardware. Further increases are possible by reducing the amount of information sent per single event. Hence, the architecture supports single rates corresponding to large PET systems even at the top hierarchy level [42]. Another possibility which is not discussed here is to offload part of the coincidence detection effort to lower level concentrators.

The top concentrator module implements a Gigabit Ethernet connection for communication with external processors. This is used for the transmission of detected coincidences and for configuration and control commands. The FPGA in each module contains a system on chip (SoC) with an embedded Nios II processor which handles these commands and relays them to lower level modules as needed.

IV. SYNCHRONIZATION OVER DATA LINKS

The key component behind the proposed architecture is the synchronization of all acquisition modules directly over data links with sub-nanosecond resolution. The general theory behind this synchronization scheme was presented in [23], as well as implementation details for Xilinx FPGAs. In this section, the operating principle and the most basic aspects are briefly described, and the differences in implementation when using Altera FPGAs are highlighted.

A. Frequency Propagation

Gigabit-rate data transmission between FPGAs is usually implemented using embedded high-speed transceivers and self-synchronous signaling, where the transmitter data clock is embedded in the data signal. The clock is recovered at the receiver using a PLL-based Clock Recovery Unit (CRU) and then used to sample and decode the incoming data stream. The transceiver is seeded by an external clock which is used both for transmission and as a reference for the CRU.

In the proposed architecture, it is mandatory to use the exact frequency of the recovered clock for the local logic at the slave node as well as for data transmission in the opposite direction, i.e. slave to master. However, the transmission clock is physically the same as the reference clock which is needed to recover the desired clock frequency in the first place. Because of this, a special clocking circuit is required in order to be able to use the same transceiver for both half-links.

The circuit from Fig. 6 is implemented in all system modules, using a National Semiconductor LMK02000 PLL and clock distributor [43] and an external voltage-controlled crystal oscillator (VCXO) with 156.25 MHz nominal frequency. The PLL device allows the charge pump connection to the loop filter to be switched open (tristate output) or closed on demand. At the top node, the switch is kept open so that the VCXO control input stays at a constant bias value and the circuit works as a fixed oscillator and clock distributor, feeding the transceiver's reference clock input. At slave nodes, the PLL is initially configured in the same way; however, once the recovered clock is stable, the loop switch is closed and the
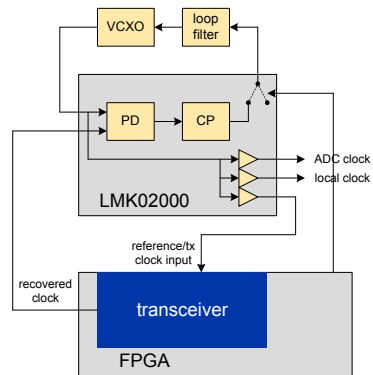

Fig. 6. Clock recovery circuit using an external PLL with VCXO. The PLL loop gets closed after the recovered clock from the transceiver becomes available.

VCXO output eventually converges to a jitter-filtered copy of the recovered clock. Transceiver operation, including clock recovery, is not affected during the PLL transient period because its reference clock suffers only very small variations while maintaining its nominal value. In particular, clock phase remains continuous. After PLL convergence, the half-link from slave to master is established, and the filtered recovered clock is used for local logic and for sampling in acquisition modules.

B. Timestamp Synchronization

After all links are established, the main clocks in all modules have exactly the same frequency $f_{\mathrm{clk}} = 1/T_{\mathrm{clk}}$, but different, fixed phases. Moreover, the phase relationship between clocks may be different each time the system is reset. An additional step is thus required in order to correct this mismatch. Precision clock distribution usually focuses on clock alignment by means of programmable phase shifters [26]–[28], [44]–[46]. A different approach is adopted here: to act directly on the timestamp counters instead of the clocks by adding a fractional part to the local timestamp, so as to account for these phase differences.

Every module runs its own timestamp counter, even concentration modules. A timestamp counter updates the integer part as usual, while the fractional part is managed exclusively by the synchronization algorithm. By adding the fractional part to the event timestamp values in the acquisition modules, the effect of varying ADC sampling clock phases is compensated when computing the time difference between events from separate modules.

Synchronization of timestamps is carried out independently on a link by link basis, following the same hierarchy as frequency propagation. In each master-slave data link, the master timestamp's fractional part remains fixed and the slave's is updated. The top module's phase is taken as a global reference and its timestamp's fractional part stays fixed at zero.

Timestamp synchronization over a single data link is based on the standard two-frame method, as used by the IEEE 1588 Precision Time Protocol (PTP) [47] among others. Two synchronization frames are sent, one from master to slave and later one from slave to master, with respective flight times t_{MS} and t_{SM}, and their local departure and arrival timestamps are
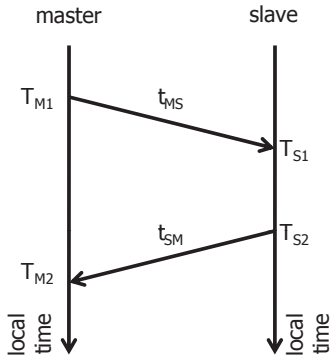
Fig. 7. Standard two-frame synchronization procedure. Local arrival and departure timestamp values are used to compute the round-trip time.

recorded, as shown in Fig. 7. Even if the timestamp counters from both nodes are not yet synchronized, the difference of timestamps from the same node is still a valid measure of a local time interval. Hence, the exact round-trip time is given by

$$t_{\mathrm{MS}} + t_{\mathrm{SM}} = (T_{\mathrm{M2}} - T_{\mathrm{M1}}) - (T_{\mathrm{S2}} - T_{\mathrm{S1}}). \qquad (5)$$

By measuring the latency t_{MS} (or t_{SM}), the correction offset for the slave timestamp counter can be obtained. Usual synchronization methods assume that $t_{\mathrm{MS}} \approx t_{\mathrm{SM}}$ and their value is estimated dividing (5) by two, but this introduces an error of $|t_{\mathrm{MS}} - t_{\mathrm{SM}}|/2$, which can be as high as T_{clk}.

In [23], a refinement was proposed that takes the skew between both half-links into account. Let us split each half-link's data path latency into components, as shown in Fig. 8. The full data path between the synchronization protocol's digital logic in each node has to be considered, because that is where the timestamps from (5) are assigned. The latency components are t_{TX} for the transmitter, t_{p} for external transmission lines like board traces and cables and t_{RX} for the receiver. An additional term $\Delta\varphi$ is needed to account for the phase changes in each node between the receiver clock domain (recovered clock) and the local clock domain used for protocol logic and transmission. To sum it up

$$
\begin{aligned}
t_{\mathrm{MS}} - t_{\mathrm{SM}} = {} & (t_{\mathrm{TX,M}} - t_{\mathrm{TX,S}}) + (t_{\mathrm{RX,S}} - t_{\mathrm{RX,M}}) \\
& + (t_{\mathrm{p,MS}} - t_{\mathrm{p,SM}}) + \Delta\varphi_{\mathrm{S}} - \Delta\varphi_{\mathrm{M}}. \quad (6)
\end{aligned}
$$

Knowing the values of (5) and (6), nodes can compute the exact half-link latencies and accurately correct the slave timestamp.

C. Measurement of Link Skew

Since (5) is always exact, synchronization resolution is governed by the accuracy of (6). Some considerations have to be made regarding the calculation of its addends.

- The term $t_{\mathrm{p,MS}} - t_{\mathrm{p,SM}}$ can be neglected with adequate board design and connection cable choice, e.g. by using composite cabling.
- The requirement that the differences in t_{TX} and t_{RX} be known exactly forces the use of transceivers that can be configured in specific deterministic latency modes, i.e. where these latencies are fixed after each reset and

their value, or at least their difference, can be obtained at runtime. Availability of such transceivers restricts the choice of FPGA from the main vendors: for Xilinx, a Virtex-5 LXT or better is needed; for Altera, an Arria II/Stratix IV GX or better. The current prototypes include Stratix IV GX devices, whose transceiver latency in deterministic mode for 8B/10B-coded links satisfies

$$
\begin{aligned}
t_{\mathrm{TX}} &= \mathrm{constant} \\
t_{\mathrm{RX}} &= \mathrm{constant} + n \cdot T_{\mathrm{clk}}/10
\end{aligned} \qquad (7)
$$

where the value of n may vary across resets but is available to the user logic. The constant terms in (7) are constant across transceiver resets. They are also constant across different FPGA instances, so they get canceled by subtraction in (6). If they were not, then a one-time calibration would be needed to compensate for them.

- A method for the measurement of fixed phase differences between same-frequency clocks in the same FPGA is required in order to obtain $\Delta\varphi_{\mathrm{M}}$ and $\Delta\varphi_{\mathrm{S}}$. In [23], it was proposed to combine measurements from a Digital Dual-Mixer Time Difference (DDMTD) system with a two-peak clustering algorithm that enhances resolution and filters invalid values. This method can be implemented entirely inside the FPGA using embedded PLLs and programmable logic.

The main contribution to synchronization inaccuracy comes from the resolution of $\Delta\varphi$ measurements. One key difference between the proposed architecture using Stratix IV FPGAs and the tests presented in [23] using Virtex-5 devices is that the Virtex transceivers can be configured in a special mode that allows the local and recovered clocks to be phase aligned on one of the two link nodes, guaranteeing $\Delta\varphi_{\mathrm{S}} \equiv 0$ by design and reducing the number of error contributing terms in (6). This is not possible with Altera devices, so a worse synchronization resolution is to be expected.

V. SETUP DESCRIPTION

Prototype circuit boards were implemented in order to evaluate the performance of the proposed DAQ architecture for a small PET system with two detectors. The family of boards used in the tests is pictured in Fig. 9. Each acquisition module prototype is formed by two boards: an acquisition board with 9 analog inputs (8 general ones and 1 for a fast trigger signal), and an analog front-end board with AMIC devices. Two different front-end boards were designed, with one and four AMICs, that can be used with photodetectors with 64 and 256 output channels, respectively. The acquisition board has two inter-module links, so it can be used as a small concentrator module with two downlinks, as well as a mixed acquisition/concentration module. A total of three acquisition boards were built for testing.

The setup used for evaluation consisted of two acquisition modules, one of them working as a concentrator as well and acting as the master in the module hierarchy. A continuous slab of 10 mm deep scintillating crystal covered by black epoxy was placed in each module, with a 49×49 mm^2 area coupled to a photodetector using optical grease. Two different types of photodetector unit were evaluated:
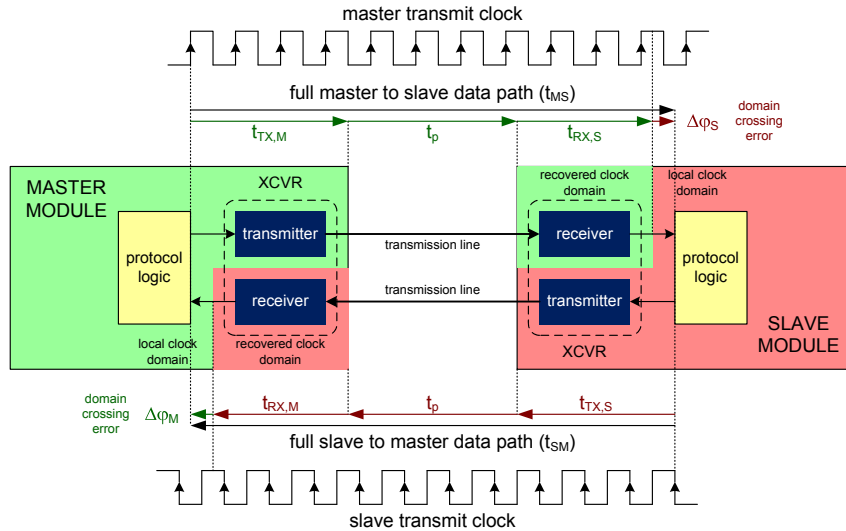
Fig. 8. Diagram of a data link, highlighting individual delay components and separate clock domains.
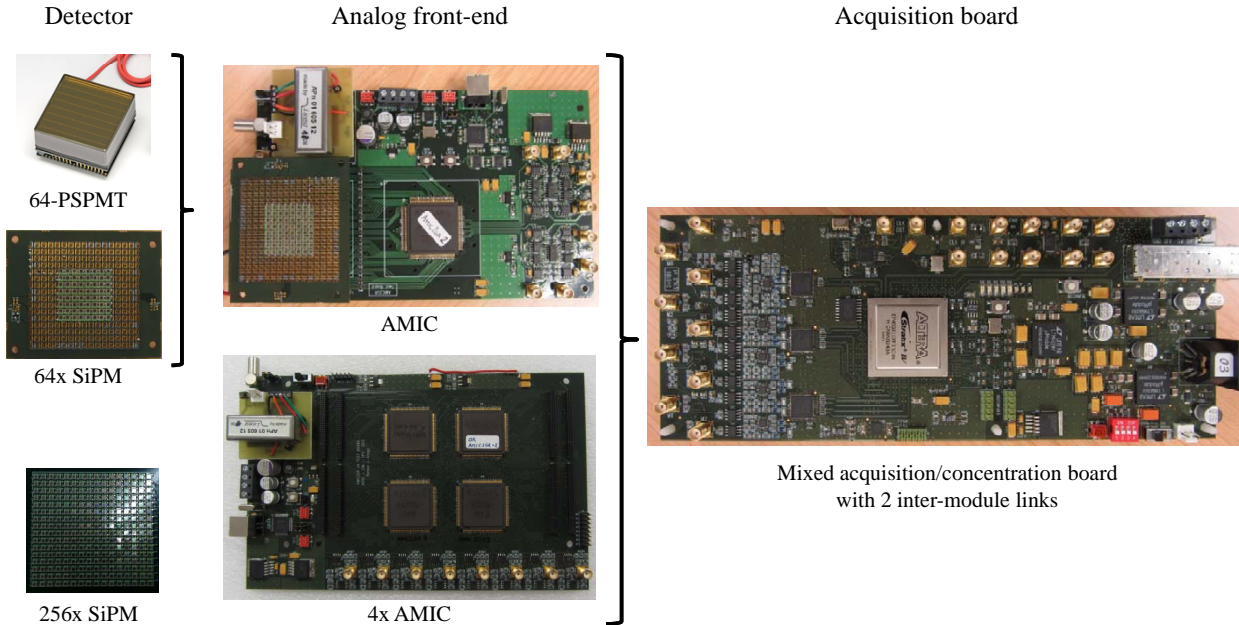

Fig. 9. Set of boards used for the module prototypes. Acquisition modules consist of a photodetector (with crystal), a front-end board, and an acquisition board. A stand-alone acquisition board can be used as a small-scale concentrator with two downlinks.

- A Hamamatsu H8500 position sensitive PMT. This detector has 64 outputs and its effective area matches that of the crystal. Parallelepiped LSO crystals were used with the PSPMTs.
- An array of 16×16 Hamamatsu S10362-11-50P MPPCs. These SiPM devices have an active area of 1×1 mm^2 but were soldered on a rectangular grid with 3.00 mm $\times$ 3.05 mm separation, so the effective scintillation area is only 10% of the total. A pyramidal frustum LYSO crystal was used in this case.

The test setup is shown in Fig. 10. A ^{22}Na point gamma-ray source was placed between both detectors, at 5 mm and 630 mm distance, respectively. A PMT detector was placed on the far side, used primarily for electronic colimation:

coincidences where the event on the far detector fell outside of the center region were filtered away. On the close side, PMT and SiPM detectors were tested. The close detector was mounted on a translation table in order to have the gamma source imping on different, known positions on the crystal.

VI. RESULTS

A. ADC and FPGA Delay Compensation

The ADC delay compensation method was evaluated first. A linear regression window of $M = 64$ samples was implemented, and the moving average of 256 measures was used as a valid delay estimation. A large number of delay estimations were captured for different power cycles and acquisition boards. The average delay exhibited a large variation
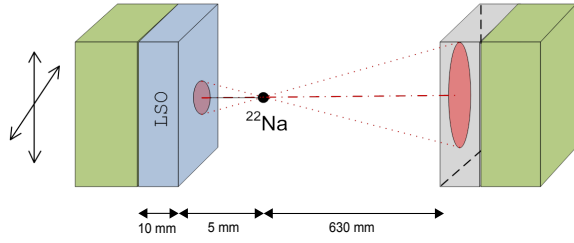
Fig. 10. Detector setup for coincidence measurements.

of up to 40 ns between different cases, showing that delay correction is indeed necessary. The variation seemed to take place mainly as integer multiples of T_{clk}, presumably due to frame decoding and channel alignment logic, while smaller variations depended on the phase difference between clock distribution nets, which was usually fixed. For individual cases, the worst-case variation of delay estimation was around $\sigma_{\text{comp}} = 17$ ps over 5 minutes.

B. Module Synchronization

The frequency replication and data link synchronization method was tested next. Phase measurements with DDMTD and two-peak clustering were implemented using the same parameters as in [23] for comparison purposes. Local phase differences took about 10 seconds to stabilize after a reset, and then yielded a resolution of $\sigma_{\Delta\varphi} = 49$ ps each. Clock jitter below 1.2 ps rms was measured on both boards using a Tektronix DSA70404C signal analyzer.

The timestamp adjustment procedure between modules was repeated every 100 ms, using 48-bit timestamps with 12 fractional bits. The synchronization algorithm converged in a single iteration, and then showed a timestamp variation of $\sigma_{\text{TS}} = 51$ ps over periods of 5 minutes. The algorithm result was found to be correct by comparison with the phase difference between master and slave clocks as measured with an oscilloscope.

The validity of the data path latency model was also tested. Using the model from Fig. 8 and (7), link round trip time (5) becomes

$$\text{RTT} = \text{constant} + (n_{\text{M}} + n_{\text{S}}) \frac{T_{\text{clk}}}{10} + \Delta\varphi_{\text{M}} + \Delta\varphi_{\text{S}} \quad (8)$$

where the constant term includes the constants from (7) and t_{p}, i.e. cable and board trace delay. Eq. (5) implies that RTT must be an integer multiple of T_{clk}; therefore, one way to confirm that the latency model is correct is to check whether the non-constant part of (8) remains constant mod T_{clk} between resets for all possible link states. For a given test setup i.e. choice of boards, the value was stable with $\sigma = 14$ ps and a 51 ps peak-to-peak variation. A variation up to 86 ps was observed between different sets of boards using the same cables.

C. Time Coincidence

In order to test time coincidence, a common pulse source was distributed to two different acquisition channels, either on the same board or different boards. An Agilent 33250A

arbitrary waveform generator was used as a source of fixed-waveform, randomly separated pulses. The pulses were independently detected and timestamped, and the time difference distribution for at least 10^5 pulses was obtained. For each setup, measurements were repeated interchanging the cable connections to both channels in order to eliminate the possible fixed time bias from cable length mismatch. The DCFD method was used for pulse timestamping, using $A = 1$ and $k = 4$ in (1) and working in amplitude mode.

The timestamp difference should ideally be zero in all cases. Hence, the mean value μ of the time difference distribution represents the systematic error that needs to be calibrated away. Different values ranging from 0 to 150 ps were measured for different setups, i.e. choices of boards and channels within each board, although μ was always fixed for a given setup. Moreover, μ seemed to be independent of the choice between using two channels on the same board or on different boards. The conclusion is that these systematic errors are caused by fixed delay differences in the analog stages in each channel, so they can be corrected with a one-time calibration of individual boards. In particular, ADC delay compensation and timestamp synchronization do not seem to add systematic error.

For the evaluation of synchronization resolution, ADC delay compensation was disabled in order to eliminate the effect of σ_{comp} on the measurements; a systematic error is hereby introduced, but it does not affect variance. The time difference distribution was measured first for two channels on the same acquisition board, in order to get an estimation of the impact of the timing algorithm on these measurements, and a resolution $\sigma_1 = 85$ ps was obtained. Using two channels on different acquisition boards yielded a resolution $\sigma_2 = 98$ ps. This value includes only the variation due to the timing algorithm and the timestamp synchronization method, and they can be assumed to be independent of each other; hence, an estimate of synchronization resolution can be obtained as

$$\sigma_{\text{sync}} \approx \sqrt{\sigma_2^2 - \sigma_1^2}. \quad (9)$$

This formula yields $\sigma_{\text{sync}} = 49$ ps, or a FWHM resolution of 115 ps assuming a gaussian distribution. Notice that this measure is very similar to σ_{TS}, that can be obtained simply by monitoring timestamp changes in the synchronization algorithm.

D. Photodetector Comparison

For each detector type, a large number of coincidences were captured with the gamma source at different positions with respect to the close detector. Five positions were used, forming an X shape centered on the center of the crystal, with a 4 mm × 4 mm separation between them. AMICs were programmed to emulate an ideal 2D charge division circuit with 4 outputs for Center of Gravity (CoG) positioning and to generate a trigger signal proportional to the sum of all detector outputs. The AMIC coefficients were not calibrated, i.e. detector channel gain spread was not compensated for. DCFD in amplitude mode with $A = 2$ and $k = 1$ was used on the specified trigger signal for event timestamping. A wide time coincidence window of ±50 ns was used in order to

observe the random coincidence background. Events where any ADC channel reached its full scale value were considered saturated and filtered away.

For position resolution, coincident events were energy filtered around the photopeaks, and electronic colimation was applied by using only events that were detected less than 15 mm away from the field-of-view (FOV) center on the far side. An appropriate time coincidence window was applied so as to remove random coincidences. Figure 11 shows the 2D histogram of detected event position for both detectors. For PMT, the five positions are clearly separated, and a spatial resolution of 2.7 mm and 2.6 mm in different axes is obtained at the center. For SiPM, the image is noisier and the points are blurred but still distinguishable; the measured resolutions at the center are 4.4 mm and 3.9 mm.

Energy and time resolution were measured only at the center point. For each event, energy was estimated as the maximum detected amplitude value of the sampled trigger signal. Energy resolution was measured by histogramming energy measures and fitting a gaussian curve around the photopeak. FWHM resolutions of 19% for PMT and 24% for SiPM were obtained, as depicted in Fig. 12. Similarly, time coincidence resolution was measured by histogramming timestamp differences and fitting a gaussian curve around the peak. The result is shown in Fig. 13, where FWHM resolutions of 2.0 ns for PMT and 4.7 ns for SiPM are obtained.

VII. SUMMARY AND DISCUSSION

A fully modular and scalable system architecture for PET has been proposed that is based on synchronization over data links. A particular implementation of said architecture has been described where the same circuit boards accept both PMT and SiPM based photodetectors. Prototype boards have been designed and the architecture has been successfully tested for a small-scale PET system with two detectors. The design and validation of a full-scale coincidence detection system using several concentrator modules remains pending.

The synchronization method has been evaluated in a realistic setting, and its resolution has been shown to be within TOF PET specifications, as it is completely negligible compared to the 600 ps figure. However, the results appear to be inferior to those obtained in [23] using Virtex FPGAs. One reason is proposed for this: the lack of a phase alignment mode in Stratix transceivers that would allow the elimination of one $\Delta\varphi$ addend in (6). This conclusion is tentative, however, as the testing conditions for both implementations were not identical.

The implemented modules have been shown to work with both PMT and SiPM based photodetectors; in particular, the ability to work with arrays of 256 SiPM has been proved. The performance of both photodetectors can be compared but the conditions were not the same, either: for instance, the scintillator area coverage for the SiPM detector was only 10% of that of the PMT, so worse results are to be expected. Using only the simplest configuration for the front-end and digital algorithms (AMIC as a CoG network; lack of gain calibration; event detection by fixed threshold crossing; basic, fixed, non-interpolated DCFD for timing), decent results have

been obtained for all measured specifications. For comparison purposes, a time resolution of 2.0 ns has been measured with PMTs, while 3.4 ns were obtained in [39] under similar conditions but with a different DAQ [48]. No reference has been found for comparison of measured timing resolution with SiPM under the same conditions, but previous work has shown that the employed detector configuration imposes a limit close to 2.0 ns [49]. In any case, the measured resolutions are expected to improve by optimizing the digital algorithm parameters and by introducing online waveform interpolation techniques, as suggested in [39].

REFERENCES

[1] J. S. Karp *et al.*, "Benefit of time-of-flight in PET: experimental and clinical results," *J. Nucl. Med.*, vol. 49, no. 3, pp. 462–470, Mar. 2008.

[2] M. Conti, "State of the art and challenges of time-of-flight PET," *Phys. Medica*, vol. 25, pp. 1–11, 2009.

[3] V. Bettinardi *et al.*, "Physical performance of the new hybrid PET/CT Discovery-690," *Med. Phys.*, vol. 38, no. 10, pp. 5394–5411, Oct. 2011.

[4] W. W. Moses, "Recent advances and future advances in time-of-flight PET," *Nucl. Instr. and Meth. A*, vol. 580, no. 2, pp. 919–924, Oct. 2007.

[5] M. Conti, L. Eriksson, and V. Westerwoudt, "Estimating image quality for future generations of TOF PET scanners," *IEEE Trans. Nucl. Sci.*, vol. 60, no. 1, pp. 87–94, Feb. 2013.

[6] W. W. Moses *et al.*, "Optimization of a LSO-based detector module for time-of-flight PET," *IEEE Trans. Nucl. Sci.*, vol. 57, no. 3, pp. 1570–1576, Jun. 2010.

[7] L. Njejimana *et al.*, "Design of a real-time FPGA-based data acquisition architecture for the LabPET II: an APD-based scanner dedicated to small animal PET imagin," *IEEE Trans. Nucl. Sci.*, vol. 60, no. 5, pp. 3633–3638, Oct. 2013.

[8] B. E. Atkins *et al.*, "A data acquisition, event processing and coincidence determination module for a distributed parallel processing architecture for PET and SPECT imaging," in *Proc. NSS-MIC Conf.*, San Diego, CA, Oct. 2006, pp. 2439–2442.

[9] M. Budassi *et al.*, "First results from the BNL/Penn PET-MRI system for whole body rodent imaging at 9.4T," in *Proc. NSS-MIC Conf.*, Anaheim, CA, Oct. 2012, pp. 2753–2755.

[10] W. W. Moses *et al.*, "OpenPET: A flexible electronics system for radiotracer imaging," *IEEE Trans. Nucl. Sci.*, vol. 57, no. 5, pp. 2532–2537, Oct. 2010.

[11] S. Lee *et al.*, "Ethernet-based flash ADC for a plant PET detector system," in *Proc. NSS-MIC Conf.*, Anaheim, CA, Oct. 2012, pp. 1320–1322.

[12] I. N. Weinberg *et al.*, "Flexible geometries for hand-held PET and SPECT cameras," in *Proc. NSS-MIC Conf.*, San Diego, CA, Nov. 2001, pp. 1133–1136.

[13] J. E. Mackewn *et al.*, "Performance evaluation of an MRI-compatible pre-clinical PET system using long optical fibers," *IEEE Trans. Nucl. Sci.*, vol. 57, no. 3, pp. 1052–1062, Jun. 2010.

[14] G. F. Knoll, *Radiation detection and measurement*, 3rd ed. John Wiley & Sons, 2000, pp. 578–583.

[15] M. F. Bieniosek, P. D. Olcott, and C. S. Levin, "Time resolution performance of an electro-optical-coupled PET detector for time-of-flight PET/MRI," in *Proc. NSS-MIC Conf.*, Valencia, Spain, Oct. 2011, pp. 2531–2533.

[16] G. Sportelli *et al.*, "Reprogrammable acquisition architecture for dedicated positron emission tomography," *IEEE Trans. Nucl. Sci.*, vol. 58, no. 3, pp. 695–702, Jun. 2011.

[17] J. Imrek *et al.*, "Development of an FPGA-based data acquisition module for small animal PET," *IEEE Trans. Nucl. Sci.*, vol. 53, no. 5, pp. 2698–2703, Oct. 2006.

[18] M. Streun *et al.*, "The data acquisition system of ClearPET Neuro - a small animal PET scanner," *IEEE Trans. Nucl. Sci.*, vol. 53, no. 3, pp. 700–703, Jun. 2006.
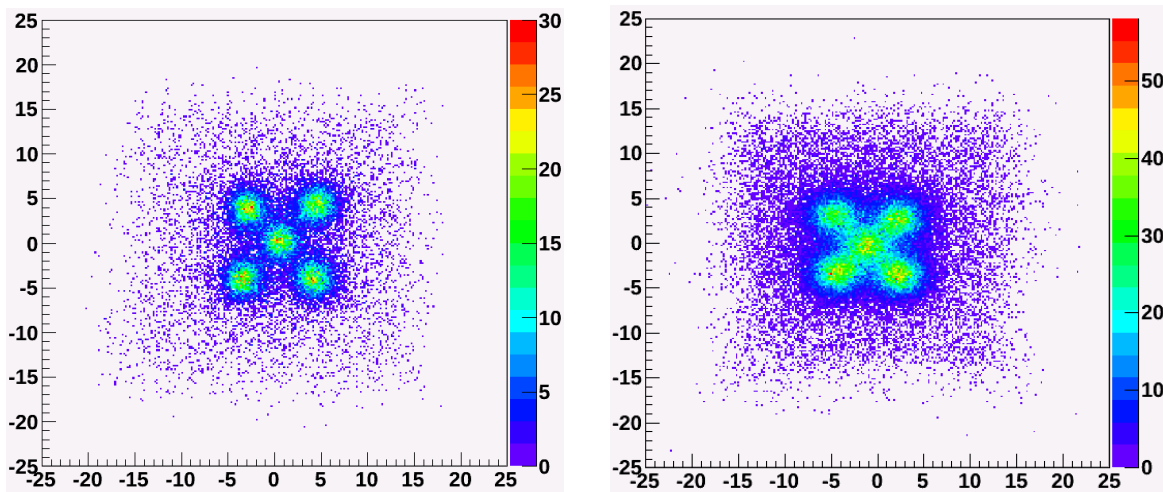
Fig. 11.   Histogram of measured positions for PMT (left) and 256 SiPM (right), using a centered X-shaped grid of source positions with 4 mm separation along each axis.
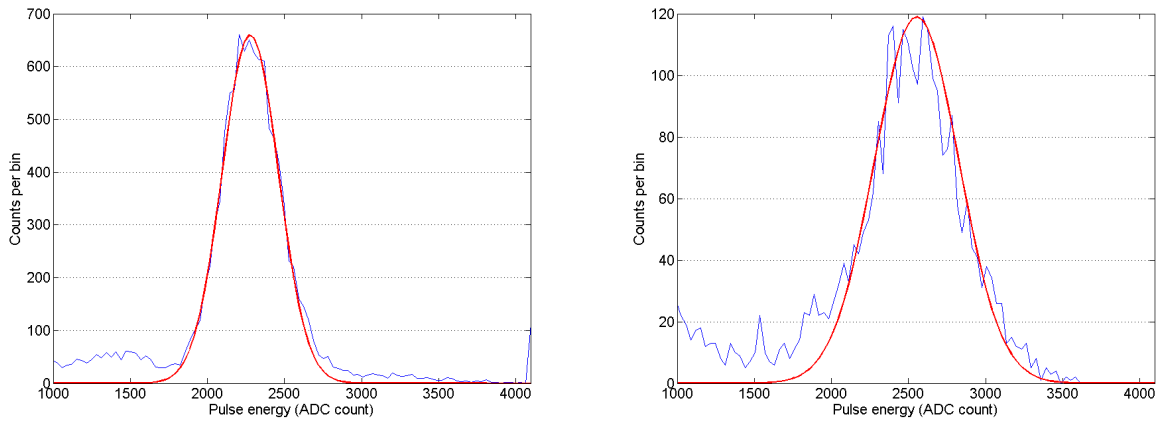

Fig. 12.   Histogram of measured pulse energy and gaussian fit of the photopeak for PMT (left) and 256 SiPM (right) at the center of the detector.
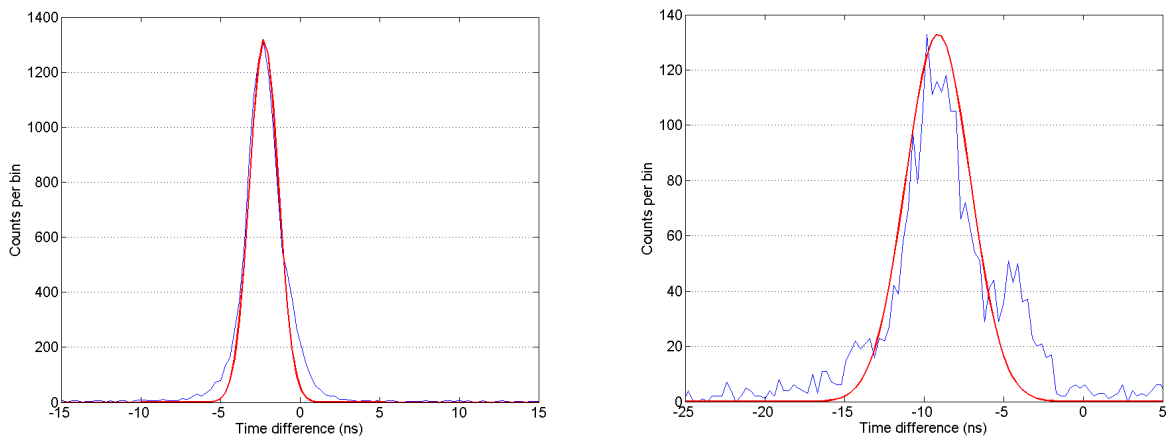

Fig. 13.   Histogram of measured time difference and gaussian fit for coincident events for PMT (left) and 256 SiPM (right) at the center of the detector. In the former, the mean value matches the distance difference of 625 mm between the gamma source and the detectors. In the latter, the mean value is different due to PMT and SiPM front-end boards having different analog delays.

[19] A. Mann *et al.*, "A sampling ADC data acquisition system for positron emission tomography," *IEEE Trans. Nucl. Sci.*, vol. 53, no. 1, pp. 297–303, Feb. 2006.

[20] A. Navarro-Tobar, C. Fernández-Bedoya, and I. Redondo, "Low-cost, high-precision propagation delay measurement of 12-fibre MPO cables for the CMS DT electronics upgrade," *J. Instrum.*, vol. 8, no. 2, p. C02001, Feb. 2012.

[21] W. W. Moses and C. J. Thompson, "Timing calibration in PET using a time alignment probe," *IEEE Trans. Nucl. Sci.*, vol. 53, no. 5, pp. 2660–2665, Oct. 2006.

[22] A. Hidvégi *et al.*, "Timing and triggering system prototype for the XFEL project," *IEEE Trans. Nucl. Sci.*, vol. 58, no. 4, pp. 1852–1856, Aug. 2011.

[23] R. J. Aliaga *et al.*, "PET system synchronization and timing resolution using high-speed data links," *IEEE Trans. Nucl. Sci.*, vol. 58, no. 4, pp. 1596–1605, Aug. 2011.

[24] M. Lipiński *et al.*, "White Rabbit: a PTP application for robust sub-nanosecond synchronization," in *Proc. IEEE Int. Symp. on Precision Clock Synchronization (ISPCS)*, Munich, Germany, Sep. 2011, pp. 25–30.

[25] H. Le Provost *et al.*, "A readout system-on-chip for a cubic kilometer submarine neutrino telescope," *J. Instrum.*, vol. 6, no. 12, p. C12044, Dec. 2011.

[26] L. Shang *et al.*, "A prototype clock system for LHAASO WCDA," *IEEE Trans. Nucl. Sci.*, vol. 60, no. 5, pp. 3537–3543, Oct. 2013.

[27] I. Papakonstantinou *et al.*, "A fully bidirectional optical network with latency monitoring capability for the distribution of timing, trigger and control signals in high-energy physics experiments," *IEEE Trans. Nucl. Sci.*, vol. 58, no. 4, pp. 1628–1640, Aug. 2011.

[28] A. Rohlev *et al.*, "Sub-nanosecond machine timing and frequency distribution via serial data links," in *Proc. Topical Workshop on Electronics for Particle Physics (TWEPP)*, Naxos, Greece, Sep. 2008, pp. 411–413.

[29] V. Herrero *et al.*, "AMIC: An expandable front-end for gamma-ray detectors with light distribution analysis capabilities," *IEEE Trans. Nucl. Sci.*, vol. 58, no. 4, pp. 1641–1646, Aug. 2011.

[30] S. Siegel *et al.*, "Simple charge division readouts for imaging scintillator arrays using a multi-channel PMT," *IEEE Trans. Nucl. Sci.*, vol. 43, no. 3, pp. 1634–1641, Jun. 1996.

[31] P. D. Olcott *et al.*, "Compact readout electronics for position sensitive photomultiplier tubes," *IEEE Trans. Nucl. Sci.*, vol. 52, no. 1, pp. 21–27, Feb. 2005.

[32] N. Ferrando *et al.*, "Cellular automaton-based position sensitive detector equalization," *Nucl. Instr. and Meth. A*, vol. 604, no. 1, pp. 211–214, Jun. 2009.

[33] C. W. Lerche *et al.*, "Fast circuit topology for spatial signal distribution analysis," in *Proc. 17th IEEE-NPSS Real Time Conf.*, Lisbon, Portugal, May 2010.

[34] ——, "Depth of interaction detection for gamma-ray imaging," *Nucl. Instr. and Meth. A*, vol. 600, no. 3, pp. 624–634, Mar. 2009.

[35] V. Herrero-Bosch *et al.*, "Programmable integrated front-end for SiPM/PMT PET detectors with continuous scintillator crystal," *J. Instrum.*, vol. 7, no. 12, p. C12021, Dec. 2012.

[36] A. Ros *et al.*, "Expandable programmable integrated front-end for scintillator based photodetectors," in *Proc. NSS-MIC Conf.*, Anaheim, CA, Oct. 2012, pp. 3196–3200.

[37] *Quad, 12-Bit, 170 MSPS/210MSPS/250 MSPS Serial Output 1.8V ADC*, AD9239, Analog Devices, 2010. [Online]. Available: http://www.analog.com/static/imported-files/data_sheets/AD9239.pdf

[38] *Stratix IV GX Device Handbook, Volume 1*, SIV5V1-4.1, Altera, 2010. [Online]. Available: http://www.altera.com/literature/hb/stratix-iv/stratix4_handbook.pdf

[39] J. M. Monzó *et al.*, "Digital signal processing techniques to improve time resolution in positron emission tomography," *IEEE Trans. Nucl. Sci.*, vol. 58, no. 4, pp. 1613–1620, Aug. 2011.

[40] J.-D. Leroux *et al.*, "Time determination of BGO-APD detectors by digital signal processing for positron emission tomography," *IEEE Trans. Nucl. Sci.*, vol. 56, no. 5, pp. 2600–2606, Oct. 2009.

[41] M. Nakhostin *et al.*, "Development of a digital front-end electronics for the CdTe PET systems," *Nucl. Instr. and Meth. A*, vol. 614, no. 2, pp. 308–312, Mar. 2010.

[42] E. Kim *et al.*, "Trends of data path topologies for data acquisition systems in positron emission tomography," *IEEE Trans. Nucl. Sci.*, vol. 60, no. 5, pp. 3746–3757, Oct. 2013.

[43] *LMK02000 Precision Clock Conditioner with Integrated PLL*, SNAS390D, National Semiconductor, 2007. [Online]. Available: http://www.ti.com/lit/ds/symlink/lmk02000.pdf

[44] B. Taylor, "Timing distribution at the LHC," in *8th Workshop on Electronics for LHC Experiments*, Colmar, France, Sep. 2002.

[45] P. Moreira *et al.*, "The GBT project," in *Proc. Topical Workshop on Electronics for Particle Physics (TWEPP)*, Paris, France, Sep. 2009, pp. 342–346.

[46] M. Nakao, "Timing distribution for the Belle II data acquisition system," *J. Instrum.*, vol. 7, no. 1, p. C01028, Jan. 2012.

[47] *Standard for a precision clock synchronization protocol for networked measurement and control systems*, IEEE Std. 1588-2002, 2002.

[48] R. Esteve *et al.*, "A high performance data acquisition system for a 16-head PET scanner," presented at the 11th Int. Workshop on Radiation Imaging Detectors (iWoRID), Prague, Czech Republic, Jun. 2009.

[49] J. M. Monzo *et al.*, "Evaluation of a timing integrated circuit architecture for continuous crystal and SiPM based PET systems," *J. Instrum.*, vol. 8, no. 3, p. C03017, Mar. 2013.