A front-end channel in 65 nm CMOS for pixel detectors at the HL-LHC experiment upgrades

L. Ratti, *SeniorMember*, *IEEE*, F. De Canio, L. Gaioni, M. Manghisoni, *SeniorMember*, *IEEE*, V. Re and G. Traversi, *SeniorMember*, *IEEE*

***Abstract*—A front-end channel prototype for pixel detectors has been designed for the upgrades of the HL-LHC experiments. The circuit is based on a Krummenacher feedback network to continuously reset the charge sensitive amplifier and on a fast threshold discriminator to implement a time-over-threshold (ToT) method and perform amplitude measurement. The front-end circuit was developed in a 65 nm CMOS technology and takes an overall area not exceeding 1250 μm^2, i.e., half of the overall pixel area. The current consumption per channel is around 4 μA at $VDD = 1.2$ V. A very small charge sensitivity dispersion was detected in the set of characterized samples. An equivalent noise charge of 120 e^- was found for a detector capacitance of 100 fF. The response of the channel is compatible with the speed requirements of the foreseen application in the innermost layers of the CMS pixel detector.**

INTRODUCTION

IN the upgraded version of the Large Hadron Collider (LHC), the so-called High-Luminosity (HL) LHC, the instantaneous luminosity will be pushed to the unprecedented level of 5×10^{34} $cm^{-2}s^{-1}$, i.e., 5 times the present nominal one. This will result in a set of extremely severe requirements for the tracking pixel detectors in the ATLAS and CMS experiments upgrades, especially for the layers closer to the interaction points. Improved accuracy in momentum measurement will be achieved by increasing the granularity of the detector, down to a pixel size of 50 $\mu m\times50$ μm (or 25 $\mu m\times100$ μm) in the inner layers. In order to minimize the material budget and the complexity of the cooling system, the power dissipation will have to be kept around 0.5 W/cm^2 or below at hit rates in the order of 2 GHz/cm^2. Such a high rate will impact on the bandwidth requirements for the readout circuits, boosting the design efforts to include as much intelligence as possible in the elementary pixel front-end channel and reduce the amount of data to be transferred off chip. The foreseen growth in luminosity will also lead to a remarkable increase of the radiation levels. The front-end circuits will have to ensure reliable operation in a harsh radiation environment, with a predicted total ionising dose of 10 MGy and a 1 MeV neutron equivalent fluence of 2×10^{16} cm^{-2} accumulated during their lifetime. With respect to the present pixel detector systems, where 300 μm is the typical thickness of the sensor substrate, thinner detectors are being proposed for the so-called phase II experiment upgrades to reduce multiple scattering and improve radiation tolerance. This will lead to smaller signals, which in turn will exacerbate the requirements on the noise performance of the analog front-end electronics. In order to preserve the detection efficiency, the readout channel needs to be operated at relatively small thresholds, around 1000 e^- or lower. This sets demanding requirements on the equivalent noise charge (ENC) and on the threshold dispersion.

The design of a new pixel readout chip complying with the above specifications is being tackled in the framework of the RD-53 Collaboration using a 65 nm CMOS technology [1]. The CHIPIX65 project, funded by the Italian Institute for Nuclear Physics (INFN), is contributing to the design effort led by the RD-53 consortium with the development of several building blocks, including a couple of options for the very front-end channel for the innermost layer of the CMS pixel detector. Section I of this paper will discuss the main design features of one of the proposed solutions, based on a Krummenacher network for the continuous reset of the charge sensitive amplifier (CSA) and on a fast discriminator with locally adjustable threshold voltage. This solution is also designated as asynchronous, as opposed to the other one developed in the frame of the same CHIPIX65 collaboration, based on the synchronous operation of a comparator with offset cancellation [2]. Section II will focus on the experimental characterization of the charge preamplifier and of the discriminator and will discuss their performance in view of their application to the HL-LHC experiments.

I. ASYNCHRONOUS READOUT CHANNEL DESCRIPTION

A block diagram of the asynchronous readout channel is shown in Fig. 1. The signal from the sensor (modeled by the current source I_D in parallel with the capacitor C_D) is converted to a voltage by means of a charge sensitive amplifier, continuously reset by means of a Krummenacher feedback network [3]. C_{in} is the parasitic capacitance at the preamplifier input, mainly contributed by the gate-to-source capacitance in the CSA input device. The signal at the

L. Ratti is with Università di Pavia, Dipartimento di Ingegneria Industriale e dell'Informazione, Via Ferrata 5, I-27100 Pavia, Italy and INFN, Sezione di Pavia, Via Bassi 6, I-27100 Pavia, Italy (phone: +39 0382 985222; fax: +39 0382 422583; email: lodovico.ratti@unipv.it).

F. De Canio is with Università di Pavia, Dipartimento di Ingegneria Industriale e dell'Informazione, Via Ferrata 5, I-27100 Pavia, Italy and INFN, Sezione di Pavia, Via Bassi 6, I-27100 Pavia, Italy

L. Gaioni, M. Manghisoni, V. Re and G. Traversi are with Università di Bergamo, Dipartimento di Ingegneria e Scienze Applicate, Viale Marconi 5, I-24044 Dalmine (BG), Italy and INFN, Sezione di Pavia, Via Bassi 6, I-27100 Pavia, Italy.

The activity leading to the results presented in this paper was carried out in the framework of the CHIPIX65 project, funded by the Italian Institute for Nuclear Physics (INFN), and was supported by the H2020 project AIDA-2020, GA no. 654168.

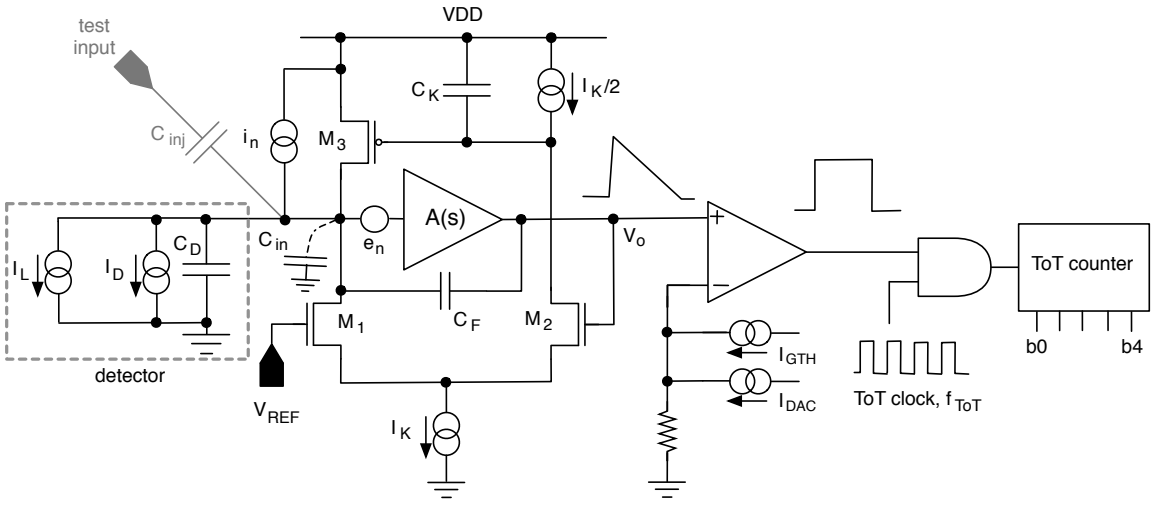

Fig. 1. Block diagram of the asynchronous front-end channel. A separate path, in parallel to the input pad (connecting to the detector , is used to inject charge through capacitor C_{inj} for test purposes.

preamplifier output is fed to a threshold discriminator, turning the signal amplitude into a time interval or ToT, time-over-threshold. The threshold discriminator is based on a low power transimpedance amplifier for fast switching operation. Given the triangular shape of the preamplifier response, featuring a very fast leading edge and a return to baseline with a constant slope, a linear relationship between amplitude (or input charge) and ToT is expected. The threshold discriminator output is used as a gate signal (through the AND gate) for the ToT clock, which is delivered to the 5 bit ToT counter for time-to-digital conversion. The current I_{GTH} is used for chip-wide threshold configuration and is set by circuits located in the chip periphery. Channel to channel dispersion of the threshold voltage is addressed by means of a local circuit for threshold adjustment, based on a 4-bit binary weighted DAC, which generates the current I_{DAC} adding to I_{GTH}. The block diagram in Fig. 1 also includes a capacitor C_{inj}, which is used to inject a signal at the channel input in the test chip.

A. Charge preamplifier and Krummenacher network

The gain stage of the charge preamplifier is based on a single-ended, NMOS input, folded cascode architecture, with a regulated cascode load serving the purpose of boosting the impedance seen at the amplifier output node [4]. This configuration is well known for allowing the designer both to maximize the DC gain and to optimize the output swing in a single stage amplifier. The transistor-level schematic is shown in Fig. 2. The input transistor, M_4, has W/L=5 µm/100 nm and a transconductance g_m=67 µA/V. The overall current consumption is slightly in excess of 3 µA, resulting in a power dissipation of about 3.6 µW at VDD=1.2 V. The DC gain and the -3dB cut-off frequency of the open loop response, as obtained from simulations, are A_0=76 dB and f_0=140 kHz respectively. Therefore, using a dominant pole approximation, the transfer function $A(s)$ of the gain stage is given by

$$\frac{V_{out}}{V_{in}} = A(s) = -\frac{A_0}{1+\frac{s}{2\pi f_0}}. \quad (1)$$

The gain bandwidth product GBP of the circuit can be defined as

$$GBP = A_0 f_0 = \frac{g_m}{2\pi C_Z} \quad (2)$$

where C_Z is the total capacitance at the output node.
The Krummenacher feedback network, shown in Fig. 1, consists of the transistors M_1 and M_2, forming a differential pair, and M_3, the capacitor C_K and the two current sources I_K and

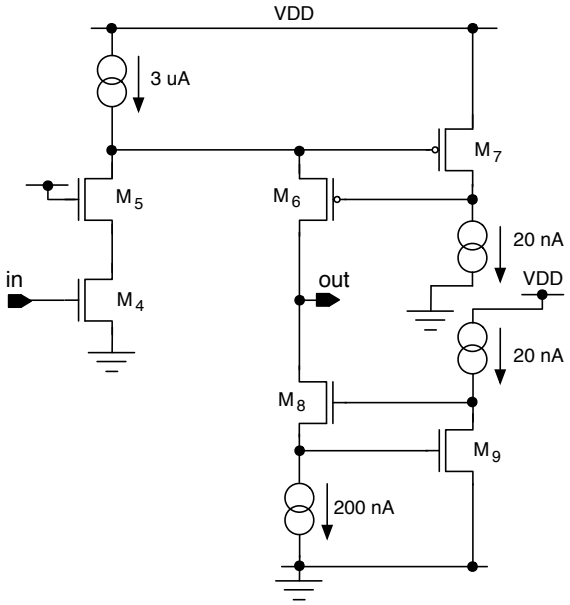

Fig. 2. Gain stage of the charge sensitive amplifier.

$I_K/2$. This architecture was specifically chosen for its capability to compensate for the detector leakage current, which is expected to increase significantly during the experiment as a result of the exposure to very high radiation doses. The negative feedback loop around the preamplifier sets the DC voltage at the output node, corresponding to the gate of M_2, to the same V_{REF} value as at the gate of M_1. The quiescent current flowing into the drain of M_1, $I_K/2$, is the same as the one flowing through M_2 and is provided by M_3. If a leakage current I_L is present, the low frequency loop, featuring a relatively large time constant depending on the resistance seen at the gate of M_3 and on the capacitance C_K (about 430 fF in this circuit), adjusts the voltage $V_{G,3}$ at the gate of M_3 so that it can supply the current required by the detector. Therefore, the drain current in M_3 will be $I_K/2+I_L$. Note that the current sourced by M_3 can be much larger than I_K. $V_{G,3}$ can be decreased, i.e., the drain current in M_3 can be increased, as long as M_2 and M_3 remain in saturation. In simulation, the circuit was found to be capable of compensating for leakage currents in excess of 100 nA, with $I_K \simeq$25 nA.

For small input charge, the response of the preamplifier can be studied by using small signal device models. If $C'_D = C_D + C_{in}$, the feedback capacitance C_F and A_0 are such that $\frac{C'_D+C_F}{A_0}$ is negligible with respect to C_F, then the transfer function can be written as

$$F(s) = \frac{V_o}{Q} \simeq \frac{2C_K}{g_{m_n} g_{m_p}} \frac{s}{(1+s\tau_1)(1+s\tau_2)(1+s\tau_3)}, \quad (3)$$

where Q is the charge delivered by the detector with a Dirac delta shaped pulse, V_o is the Laplace transform of the v_o signal at the preamplifier output, g_{m_n} is the transconductance of either M_1 or M_2, g_{m_p} is the transconductance of M_3, $\tau_1 = \frac{C_K}{g_{m_p}}$, $\tau_2 = \frac{2C_F}{g_{m_n}}$ and $\tau_3 = \frac{C'_D+C_F}{2\pi C_F GBP}$. Under the hypothesis of linearity, as it is the case for the system considered here, the gain of the charge preamplifier G_Q, or charge sensitivity, can be calculated as the peak value of the Laplace anti-transform of $F(s)$. Under the assumption that $\tau_1 >> \tau_2 >> \tau_3$,

$$G_Q \simeq \frac{1}{C_F}(1 - h_1 - h_2), \quad (4)$$

where

$$h_1 = \frac{g_{m_n}(C'_D + C_F)}{4\pi C_F^2 GBP} \left[1 + \ln \frac{4\pi C_F^2 GBP}{g_{m_n}(C'_D + C_F)}\right], \quad (5)$$

$$h_2 = 2\frac{g_{m_p} C_F}{g_{m_n} C_K}. \quad (6)$$

In the circuit under analysis, $h_2 << h_1 << 1$. Equations (4), (5) and (6) account for the dependence of the charge sensitivity on the circuit parameters and configuration, to be discussed later on in this paper. Note that, if $x = \frac{g_{m_n}(C'_D+C_F)}{4\pi C_F^2 GBP}$, (5) takes the form $h_1 = x\left[1 + \ln\left(\frac{1}{x}\right)\right]$, with $\frac{dh_1}{dx} > 0$ for $x < 1$, which is the case in the considered circuit. This means that, if for instance g_{m_n} is increased, h_1 will also increase.

In the ideal case, where A_0, f_0 and C_K all go to infinity, the transfer function assumes a simpler form

$$\frac{V_o}{Q} \simeq \frac{1}{C_F} \frac{1}{s + \frac{g_{m_n}}{2C_F}}, \quad (7)$$

$G_Q = C_F^{-1}$ and, in the time domain, the response is of the exponential kind with a single time constant,

$$v_o(t) = H(t) \cdot \frac{Q}{C_F} \exp\left(-\frac{g_{m_n}}{2C_F} t\right), \quad (8)$$

$H(t)$ being the step function.

For sufficiently large input signals, i.e., for $Q \gtrsim 3v_T C_F$ (where v_T is the thermal voltage), the differential pair in the Krummenacher network leaves the linear region. The above limit is obtained under the reasonable assumption that the transistors in the network are operated in the sub-threshold region, leading to a relationship of the hyperbolic tangent type between the output current and the input differential voltage [5]. Under the assumption of a very large gain-bandwidth product for the amplifier, as a response to a charge signal Q, the voltage signal at the CSA output goes promptly to about Q/C_F. M_1 is switched off and the current in the M_2 drain rises to I_K (i.e., the differential pair reaches a saturation condition). Half of this current flows through C_K. However, if C_K is large enough, the change in $V_{G,3}$ and in the current sourced by M_3 during signal processing, can be neglected. Therefore, while part of the drain current of M_3, I_L, is absorbed by the sensor, the remaining part, $I_K/2$, will discharge the feedback capacitance C_F in a linear fashion. The quiescent condition is restored as soon as v_o is back to the baseline. In the time domain, as anticipated, the response to a current pulse with area Q occurring at $t = 0$ will have a roughly triangular shape like in the following piecewise defined function,

$$v_o(t) \simeq \begin{cases} 0 & \text{t<0, t> } \frac{2Q}{I_K} + t_p, \\ \frac{Q}{C_F}\frac{t}{t_p} & 0 \le t < t_p, \\ \frac{Q}{C_F} - \frac{I_K}{2C_F}(t - t_p) & t_p \le t \le \frac{2Q}{I_K} + t_p. \end{cases} \quad (9)$$

In the above approximate equation, the leading edge of the response, which has a raising exponential shape in the actual circuit, is depicted with a simple straight line segment. Again in (9), t_p is the peaking time of the CSA response which, to a large extent, is independent of Q. Over most of the input charge range, t_p is negligible with respect to the time needed to return to the baseline. So, in these operating conditions, since the current discharging the capacitor is constant, the falling edge slope is also constant and the output signal duration is an approximately linear function of the input charge. Note that the Krummenacher stage is designed to enter the saturation state for relatively small input signals. This is done to extend the linear relationship between ToT and Q over the largest possible input charge interval.

B. ToT and choice of the front-end parameters

The time-over-threshold, as the time interval during which the signal at the preamplifier output exceeds the discriminator threshold, can be calculated with reference to Fig. 3. Here, the

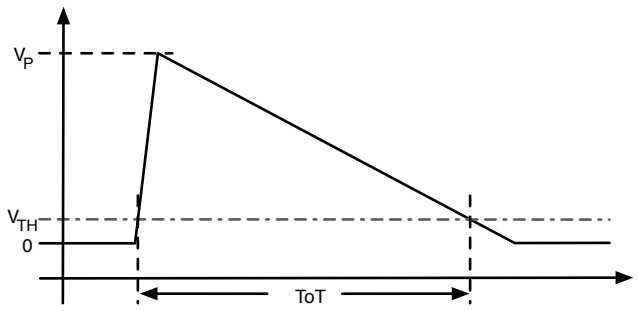

Fig. 3. Calculation of the time-over-threshold in the case of a front-end channel with triangular response.

response to a large input charge is represented with a triangular waveform with a steep rising edge and a slow falling one, as in (9). If $t_p \simeq 0$ is assumed, then

$$ToT \simeq \frac{V_P - V_{TH}}{\frac{I_K}{2C_F}} = 2\frac{Q - V_{TH}C_F}{I_K} \tag{10}$$

where $V_P = \frac{Q}{C_F}$ is the peak amplitude of the preamplifier response to an input charge Q and V_{TH} is the threshold voltage at the discriminator inverting input. The I_K current value can be set once the input dynamic range and the maximum allowed ToT are defined. Therefore, if the maximum ToT value, ToT_{MAX}, is the one obtained for the maximum input charge, Q_{MAX}, then

$$ToT_{MAX} \simeq 2\frac{Q_{MAX} - V_{TH}C_F}{I_K} \simeq 2\frac{Q_{MAX}}{I_K} \Rightarrow$$
$$\Rightarrow I_K \simeq 2\frac{Q_{MAX}}{ToT_{MAX}}, \tag{11}$$

where $Q_{MAX} \gg V_{TH}C_F$ is assumed. $V_{TH}C_F = Q_{TH}$ represents the discriminator threshold referred to the channel input, which will be no lower than 600 e^- in the real experiment. On the other hand, the maximum expected charge, set by the experiment and by the detector thickness, is 30 ke^-. In order to keep the hit loss from in-pixel pile-up well below 1%, the dead time of the channel, roughly corresponding to the time needed to restore the feedback capacitor, should not exceed 400 ns. This value represents an upper limit for the ToT, which, for large input signals, is very close to the CSA response duration. These specifications for Q_{MAX} and ToT_{MAX} yield $I_K \simeq 25$ nA. Also, given a CSA output dynamic range $\Delta V_{MAX} \simeq$450 mV, since $\frac{Q_{MAX}}{C_F}$ has to equal ΔV_{MAX} if full advantage is to be taken of the CSA linear operating region, then $C_F \simeq$10 fF.

The ToT will be measured in each channel by means of a counter clocked at a frequency f_{ToT}. A resolution equal to or larger than 4 bits is required by the experiments for the hit charge measurement. If a 5-bit, dual edge counter (i.e., sensitive to both the rising and the falling edge of the clock) is used, the maximum measurable ToT can be expressed as

$$ToT_{MAX} = \frac{2^5 \times T_{ToT}}{2}, \tag{12}$$

yielding a clock period T_{ToT}=25 ns for ToT_{MAX}=400 ns. Therefore, the whole ToT range can be covered by leveraging the LHC machine clock signal, which is only slightly different from 40 MHz.

C. Equivalent noise charge

The noise performance of the charge preamplifier is mainly determined by the contributions from the CSA input device and from the PMOS transistor M_3 part of the feedback network. In Fig. 1 they are represented by the voltage source e_n and the current source i_n respectively. Both sources feature a white thermal and a flicker noise term. However, the latter will be omitted in the following discussion, as simulations indicate that its contribution is negligible as compared to the contribution from thermal noise. Possible noise contributions in the detector leakage current will also be neglected here, as the purpose of this section is to provide a noise model for the preamplifier alone. Nevertheless, it is worth mentioning here that the noise from the detector, in the non-irradiated device, is extremely small with respect to the front-end electronics sources. On the other hand, due to radiation exposure, it may substantially increase during the actual experiment and might become comparable with the noise contribution from the charge preamplifier should the leakage current rise up to the worst case limit of 10 nA. The power spectral density of the noise in e_n and i_n can eventually be expressed as

$$\frac{\overline{de_n^2}}{df} = \frac{4k_BT\gamma_n}{g_m}, \tag{13}$$
$$\frac{\overline{di_n^2}}{df} = 4k_BT\gamma_p g_{m_p}, \tag{14}$$

where k_B is the Boltzmann's constant, T is the absolute temperature and γ_n and γ_p are the channel thermal noise coefficients for the CSA input device and M_3 respectively (depending on the operating conditions of the transistors). The mean square noise at the preamplifier output, $\overline{v_n^2}$, is found by definition as

$$\overline{v_n^2} = \int_0^{+\infty} |F(j\omega)|^2 \left[\omega^2(C_D' + C_F)^2\frac{\overline{de_n^2}}{df} + \frac{\overline{di_n^2}}{df}\right] df. \tag{15}$$

By using, for the transfer function $F(s)$, the approximated expression in (4), the following equation is obtained for the noise at the CSA output

$$\overline{v_n^2} \simeq \frac{2k_BT}{C_F} \cdot \frac{\pi\frac{\gamma_n}{g_m}(C_D' + C_F)\cdot GBP + \gamma_p\kappa_g}{1 + 2\kappa_g\frac{C_F}{C_K}} \tag{16}$$

In (16), $\kappa_g = g_{m_p}/g_{m_n}$. Note that the transistors M_1, M_2 and M_3 in the CSA feedback network work very close to weak inversion, where the transconductance is directly proportional to the drain-source current. Therefore, transconductance ratio κ_g is therefore close to unity and is subjected to very small changes when the current I_K is made to vary in a limited range. Again in (16), the second term in the denominator is much smaller than 1, as $C_K >> C_F$.

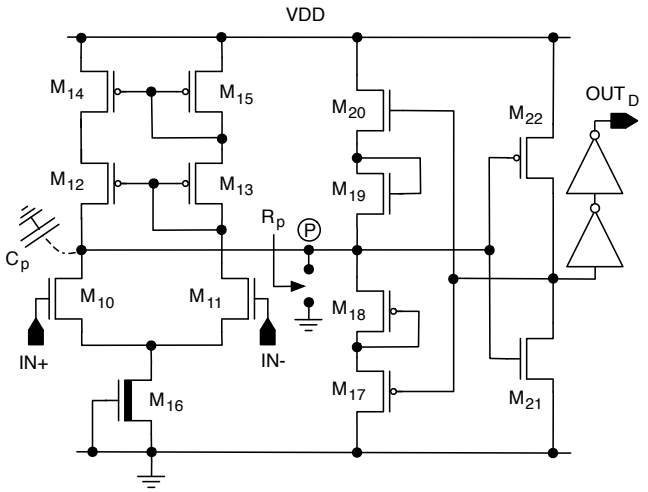

Fig. 4. Schematic diagram of the threshold discriminator.

The noise feature of a charge preamplifier is generally expressed in terms of equivalent noise charge, or ENC, which is defined as

$$ENC^2 = \frac{\overline{v_n^2}}{G_Q^2}. \tag{17}$$

If h_2 is neglected in (4), since $\frac{1}{(1-h_1)^2} \simeq 1 + 2h_1$, the square of the ENC approximates to

$$\begin{aligned} ENC^2 &\simeq \overline{v_n^2} C_F^2 (1 + 2h_1) = \\ &= ENC_0^2 \left[1 + 2h_1(g_{m_n}, C'_D, C_F)\right], \end{aligned} \tag{18}$$

where

$$ENC_0^2 = 2k_B T C_F \frac{\pi \frac{\gamma_n}{g_m}(C'_D + C_F) \cdot GBP + \gamma_p \kappa_g}{1 + 2\kappa_g \frac{C_F}{C_K}}. \tag{19}$$

Equation (18) emphasizes the dependence of h_1 on g_{m_n}, C'_D and C_F. In particular, it is worth recalling here that, under the conditions already specified in Section I-A, h_1 is an increasing function of all of the three variables.

D. Threshold discriminator

In the asynchronous front-end channel discussed in this work, the threshold discriminator serves a threefold purpose. As the signal from the charge preamplifier exceeds the preset threshold, the rising edge of the signal at the discriminator output indicates the occurence of a significant event. The same signal is used to freeze the content of a local time stamp register, set by a chip-wide distributed clock, so as to label the event with a time stamp. Finally, the discriminator signal enables a local counter to count the pulses from a clock and perform direct digitization of the time-over-threshold interval. The schematic diagram of the discriminator is shown in Fig. 4. The first stage consists of a transconductor, whose configuration is based on an NMOS differential pair (M_{10},M_{11}) with a cascoded PMOS mirror load (M_{12} to M_{15}). The tail current source, draining about 900 nA, is implemented through a native NMOSFET (M_{16}) in a self-biased configuration. The current at the transconductor output is fed to a transimpedance amplifier (TIA, transistors M_{17} to M_{22}), providing a low impedance path and a small time constant for fast switching. In the TIA, transistors M_{20} and M_{19} will be ON and transistors M_{18} and M_{17} OFF when the signal from the preamplifier (at the $IN+$ terminal) exceeds the threshold (set at the $IN-$ terminal). The opposite happens when the voltage at $IN+$ is below the threshold. Assuming that the transistors M_{17} to M_{20} are in the saturation region when ON, and that the output inverter is operated in the linear region (which is a reasonable approximation for the behavior of the stage during transition between the logic levels), the time constant τ_p of the signal at the TIA input is given by

$$\tau_p = C_p R_p = \frac{\tau_{p0}}{1 - G_{loop}}, \tag{20}$$

where

$$\tau_{p0} = \frac{C_p}{G_m}, \quad G_m = \begin{cases} \dfrac{g_{m,19} \cdot g_{m,20}}{g_{m,19} + g_{m,20}}, & 0 \rightarrow 1 \text{ transition} \\ \dfrac{g_{m,17} \cdot g_{m,18}}{g_{m,17} + g_{m,18}}, & 1 \rightarrow 0 \text{ transition} \end{cases} \tag{21}$$

$$G_{loop} = -\left(g_{m,21} + g_{m,22}\right)\left(r_{o,21} // r_{o,22}\right). \tag{22}$$

In the previous equations, C_p is the parasitic capacitance shunting node P, R_p is the resistance seen between node P and ground, τ_{p0} is the open loop time constant, G_{loop} is the loop gain of the TIA, $g_{m,i}$ and $r_{o,i}$ are, respectively, the transconductance and the output resistance of transistor M_i. The above result shows how the negative loop in the transimpedance amplifier has the effect of boosting the response speed of the discriminator.

As compared to a slightly simpler version discussed in the literature [6], the TIA in Fig. 4 includes two diode-configured MOSFETs (M_{18} and M_{19}), which were added to the input branch of the circuit to prevent the transistors in the output inverting stage from being both ON in stationary conditions, therefore minimizing crowbar current contributions [7]. Two cascaded inverters are used after the TIA to consolidate the output logic level and to minimize the time delay.

TABLE I
NOMINAL VALUES FOR CHARGE SENSITIVITY AND RECOVERY CURRENT IN THE DIFFERENT SETTINGS.

	Charge sensitivity	Recovery current
low	7.5 mV/ke$^-$ (C_F=20 fF)	12.5 nA
high	15.0 mV/ke$^-$ (C_F=10 fF)	25.0 nA

II. PROTOTYPE CHARACTERIZATION

A test structure, incorporating the readout channel shown in Fig. 1, but not including the circuitry for ToT digitization,

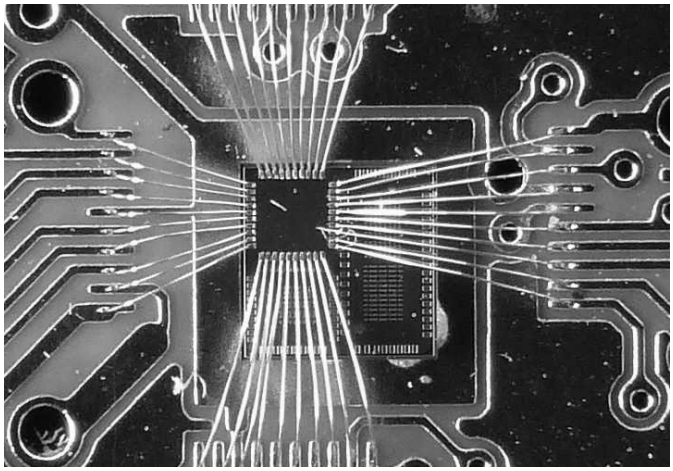

Fig. 5. Microphotograph of the 2 mm×2 mm prototype chip wire-bonded to the test board.

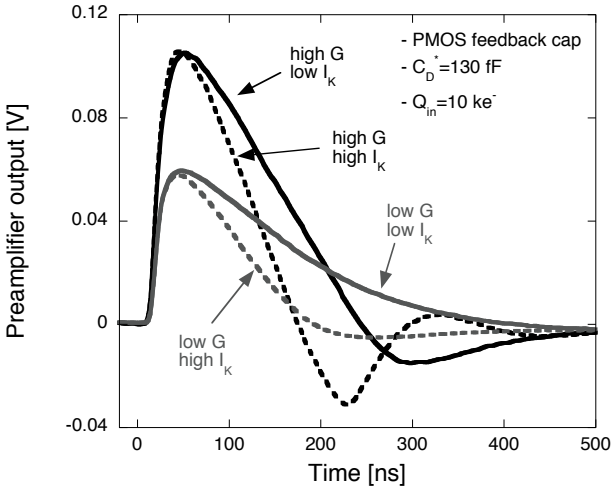

Fig. 6. Response of a charge preamplifier with a MOS capacitor in the feedback network to a 10 ke$^-$ signal for the four different possible combinations of the charge sensitivity and of the recovery current.

has been fabricated in a 65 nm CMOS technology. It contains individual channels (charge preamplifier plus threshold discriminator) and stand-alone preamplifier and discriminator stages, and was mostly conceived for testing the single blocks of the readout channel and assessing their analog performance. In the circuits, a few programming bits can be used to set the charge sensitivity G_Q (by changing the preamplifier feedback capacitance C_F, 1 bit), the recovery time (by varying the current I_K in the Krummenacher stage, 1 bit), and the detector emulating capacitance at the preamplifier input (2 bits). The nominal values of the charge sensitivity (together with the corresponding feedback capacitance value) and of the recovery current in the different configurations are shown in Table I. The different gain and recovery current options have been included in the prototype design mainly for testing purposes. The preamplifier configuration with high gain and low recovery current is the one guaranteeing that the 400 ns ToT range (corresponding to the counter full scale) and the full output dynamic range of 450 mV be simultaneously covered with an input signal range of 3×10^4 e$^-$. Therefore, in the following, it will be referred to as the standard channel configuration. The detector emulating capacitance can be made to vary from 0 to 150 fF in steps of 50 fF. The expected value for the detector capacitance in the real application is close to 50 fF. A 30 fF injection capacitance C_{inj} connected to the input terminal (see Fig. 1) is used to test the response of the channel at both the preamplifier and the discriminator output to a charge signal. The prototype preamplifier comes in two slightly different flavors. In one, the feedback capacitor is implemented with a PMOS transistor in an inversion mode configuration [8], with the gate connected to the output node of the CSA, the source and drain to the input node and the bulk to VDD. Coupling of the MOS capacitor to bulk might lead to cross-talk between cells in the same chip, due to VDD transients induced by firing pixels, therefore resulting in PSRR degradation. However, no significant effects were found in circuit simulations. In the second version of the prototype, a metal-insulator-metal (MIM) structure is used instead. According to the information available for the technology and to the simulation results, a MOS capacitor is less susceptible to process variations than a MIM one, whereas the value of a MIM capacitor tends to remain more stable within a given run. On the other hand, a significant cost reduction can be obtained if the MIM module is not included in the layer stack. Fig. 5 displays a microphotograph of the prototype chip wire-bonded to the daughter-board, to be plugged into a mother-board providing bias voltages and currents, control switches and connectors for input and output signals.

If a slight difference in charge sensitivity is excluded, no significant discrepancy was found in the performance of the two types of charge preamplifier. Therefore, the measurement results shown in the following, when referring to only one of the two, can be regarded as representative of the behavior of both versions of the CSA.

A. *Preamplifier response and charge sensitivity*

Fig. 6 shows the response of a charge preamplifier with a MOS capacitor in the feedback network to a 10^4 electron signal for the four different possible combinations of the charge sensitivity and of the recovery current. The waveforms were obtained with a detector emulating capacitance C_D at the preamplifier input of 100 fF. The capacitance C_D^* referred to in Fig. 6 (and in the subsequent ones) also includes the contribution from the injection capacitance (which, being in parallel to C_D, has to be taken into account during the channel response, charge sensitivity and noise measurements). In Fig. 6, the response of the circuit in the high gain and high recovery current configuration features a significant undershoot followed by an overshoot, both roughly proportional to the amount of input charge signal. This ringing behavior is related to the pole in $f_1 \frac{1}{1\pi\tau_1}$ in the feedback network. In the actual application, the overshoot might lead to false hits as it becomes comparable to the threshold. However, it is worth emphasizing here once again, that the options for gain and recovery current configuration were introduced in the design of this prototype merely for testing purposes. In the standard

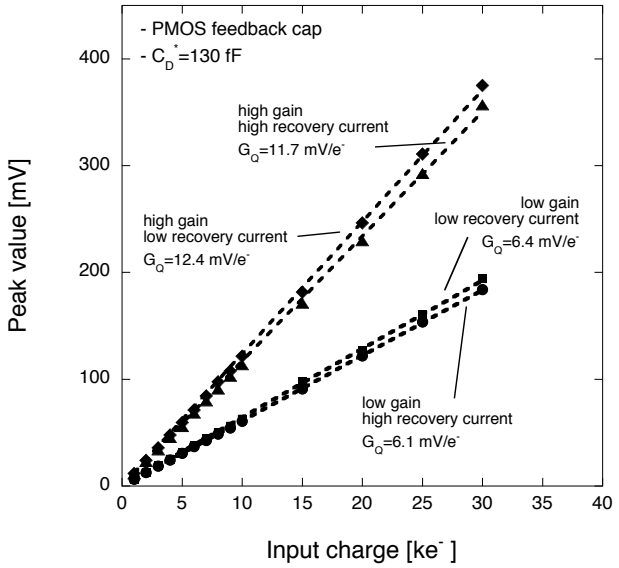

Fig. 7. Peak response of a charge preamplifier with PMOS feedback capacitor as a function of the input charge for the four different possible combinations of the charge sensitivity and of the recovery current.

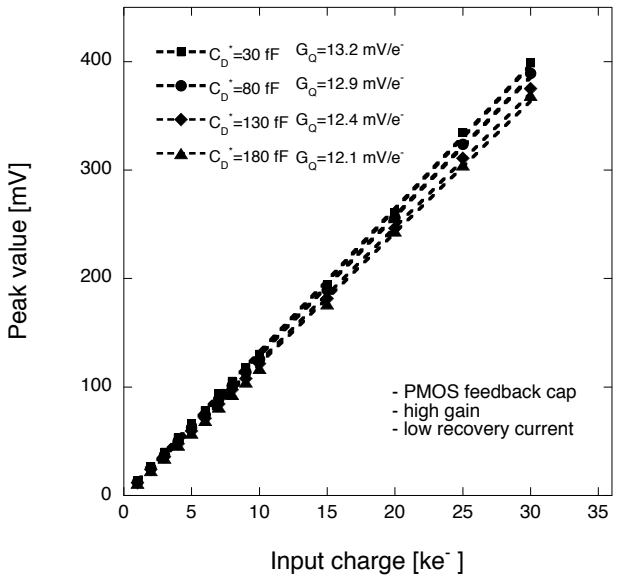

Fig. 8. Peak response of the same CSA as in Fig. 7 as a function of the input charge for four different values of C_D^*.

configuration (large gain, low recovery current), the only one which has been retained for the large scale integration, no significant ringing in the return to the baseline can be detected. Fig. 7 shows the peak amplitude of the signal at the output of a CSA with PMOS feedback capacitor as a function of the input charge in the four different available configurations. The change in the curve slope with the CSA configuration is in good agreement with the expression of G_Q in (4). In particular

- an increase in the recovery current leads to a decrease in the charge sensitivity; as a matter of fact, increasing I_K results in an increase of the transconductance g_{m_n} of M_1 and M_2 in the Krummenacher network, therefore increasing h_1 and reducing G_Q;
- reducing C_F by a factor of 2 does not multiply the charge sensitivity exactly by 2; what happens is that, when C_F is halved, h_1 increases roughly by a factor of 4; consequently, G_Q is less than doubled.

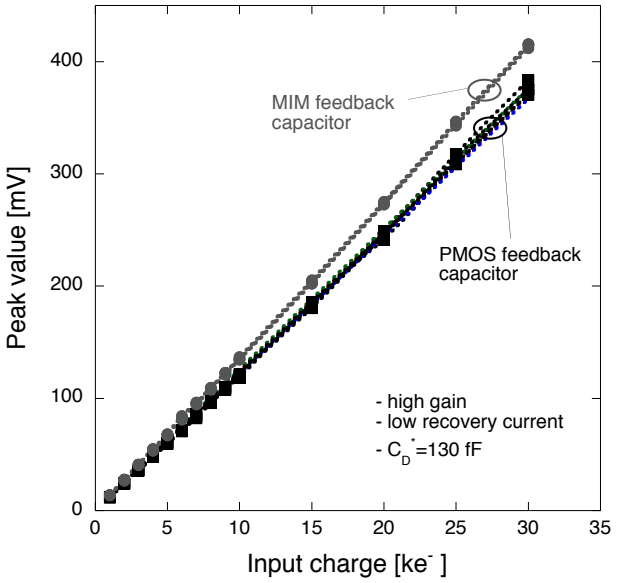

Fig. 9. Peak response of the charge preamplifier as a function of the input charge for the two versions of the stage. Each set of curves includes the results from the characterization of four different samples.

In Fig. 8, the peak amplitude of the same CSA as in Fig. 7 as a function of the input charge is plotted for four different values of C_D^*. Again, the effect of C_D^* on the charge sensitivity is consistent with (4). More specifically, an increase in the detector capacitance leads to an increase in the h_1 coefficient and to a decrease of G_Q.

Fig. 9 shows the peak amplitude in the response of the charge sensitive amplifier as a function of the input charge for the two versions of the stage and for a detector emulating capacitance of 100 fF (C_D^*=130 fF). Each of the two sets of curves includes results from the characterization of four samples. In both cases, the dispersion in the curve slope, i.e., in the charge sensitivity, is quite small, the standard deviation being less than 1% of the average value in the case of the MIM feedback capacitor and about 1% in the case of the stage with the PMOS transistor in the feedback network. Despite the very limited statistics, the measurements are in good agreement with the expectation of marginally smaller dispersion for the MIM feedback capacitor, as predicted by the foundry models. In both cases, the charge sensitivity is smaller than the design value of 15 mV/ke$^-$: about 13.7 mV/ke$^-$ for the CSA with MIM capacitor, about 12.4 mV/ke$^-$ for the one with the PMOS capacitor. This discrepancy is likely due to poor modeling of the MIM and PMOS capacitances at such small values. The non-linearity, estimated by means of the endpoint line approach, was found to be fairly low in both variants of the preamplifier: less than 4% for the one using the PMOS transistor, less than 3% for the other CSA version.

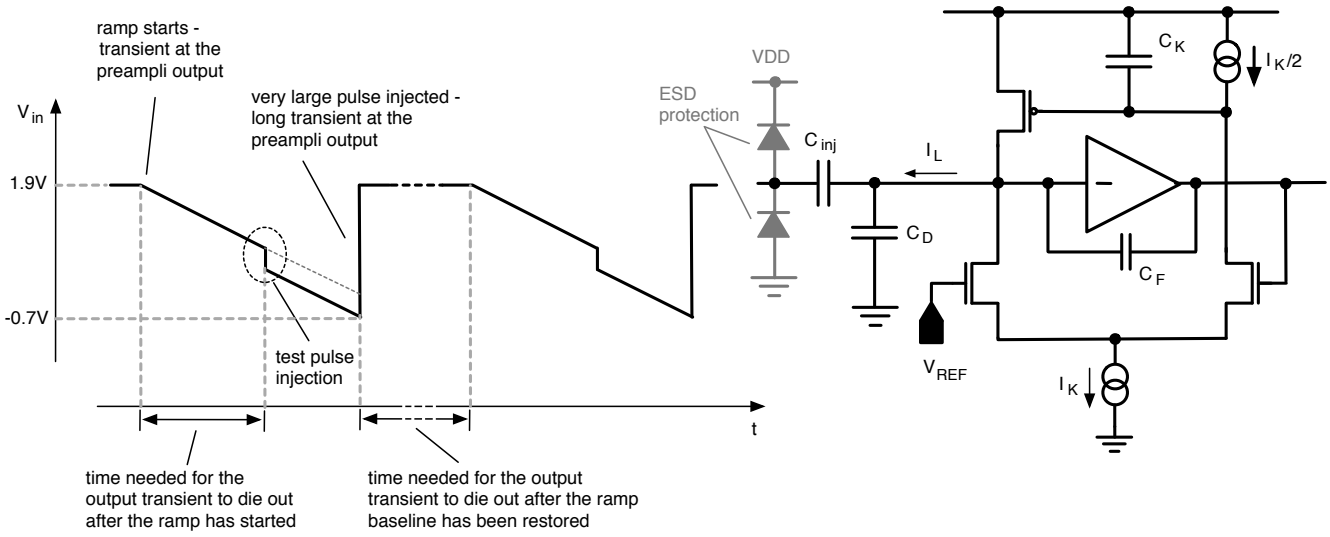

Fig. 10. Characterization of the charge preamplifier in the presence of a leakage current from the detector: input test signal.

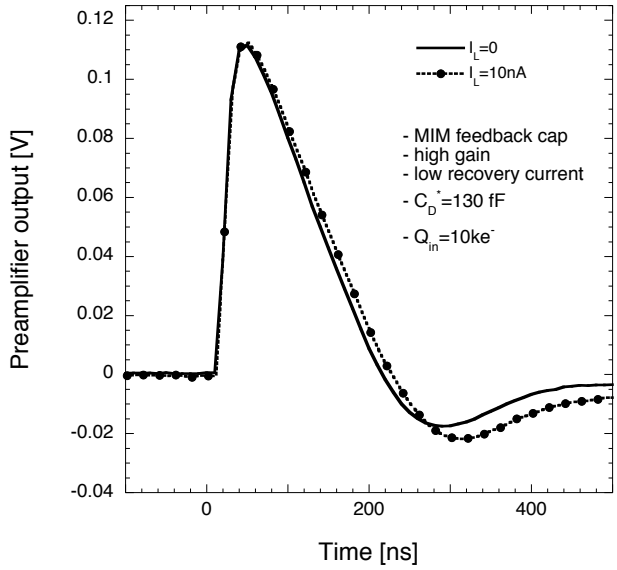

Fig. 11. Response of a charge preamplifier with MIM feedback capacitor to a 10 ke$^-$ signal with no input leakage current and in the presence of a leakage current of 10 nA.

B. Effect of leakage current at the preamplifier input

In the upgrades of the CMS and ATLAS experiments at the HL-LHC, the innermost layers of the pixel detector are expected to be struck by extremely high levels of ionizing and non-ionizing radiation, as already mentioned in the introduction to this work. One of the anticipated consequences of the detector operation in such a hostile radiation environment is a relatively large increase in the leakage current [9], [10], up to a worst case limit in the order of 10 nA. The readout channel has to ensure proper operation during its lifetime even in the presence of such a current, which has to be supplied to the detector by the readout channel itself. Therefore, a test on the charge preamplifier is mandatory to verify that the foreseen increase in the input quiescent current does not degrade its performance unacceptably.

In the device under test, due to the presence of the injection capacitor C_{inj}, the input terminal of the charge preamplifier cannot be accessed directly. However, to emulate a DC current flowing out of the CSA input node, a ramp voltage signal $V_{in}(t)$ can be applied to the injection capacitor, as suggested in Fig. 10. In this case, the emulated leakage current I_L will be given by

$$I_L = -C_{inj}\frac{dV_{in}}{dt}. \tag{23}$$

Therefore, by changing the slope of the ramp, the magnitude of the current can be adjusted. The presence of a pair of ESD protection diodes in the input pad (as in any pad of the test chip) limits the excursion of the ramp voltage signal to the $[VDD + V_\gamma, -V_\gamma]$ interval, where $V_\gamma \simeq$0.7 V is the voltage drop across a forward biased diode. Limitations close to the above ones would be set anyway by the maximum operating voltage constraints of the technology. The limit on the voltage excursion forces the test signal to be a periodic one. As a consequence, transient signals will appear at the preamplifier output as a response to a sudden change in the input signal derivative, as again indicated in Fig.10, in particular at the beginning of the ramp and when the ramp baseline is restored. Performance evaluation of the CSA in the presence of I_L has to be performed only when the stage is in its quiescent operating point, i.e., after the output transient signals have faded away. The limits on the input signal excursion and the time constraints due to the output transients set an upper bound for I_L, that in the particular setup considered in this section cannot exceed 15 nA.

Examples of the results from the tests performed in these conditions on a CSA with MIM feedback capacitor are provided in Fig. 11 and Fig. 12. Fig. 11 shows the response of a CSA to a 10 ke$^-$ signal in two different cases, no I_L and I_L=10 nA. The charge was injected by superimposing a negative step signal

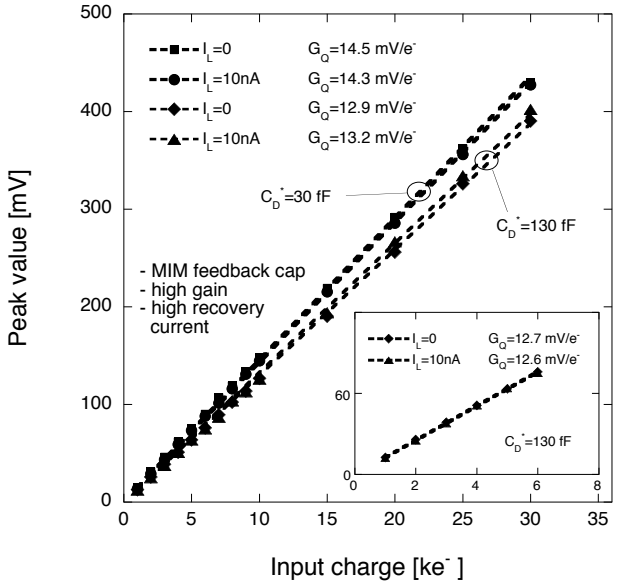

Fig. 12. Peak response of the charge preamplifier with MIM feedback capacitor as a function of the input charge with no input leakage current and in the presence of a leakage current of 10 nA. The two pairs of curves are relevant to two different values of C_D^*. The inset shows the initial part of the two curves at C_D^*=130 fF, with the relevant slopes.

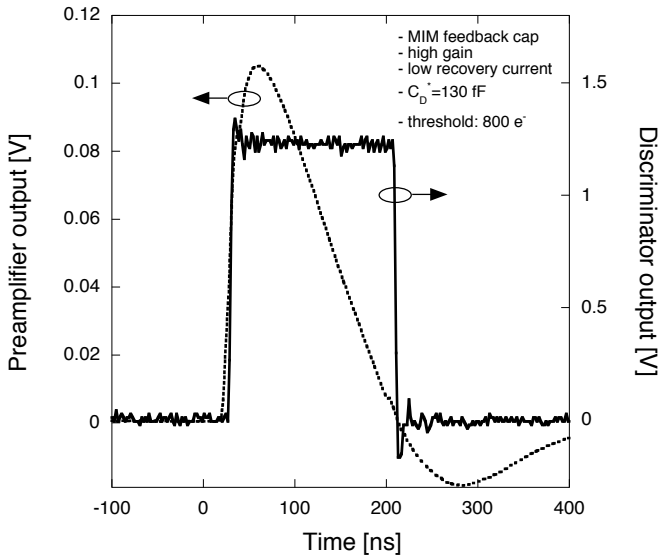

Fig. 13. Response to a 10 ke$^-$ signal of a CSA with a MIM capacitor in the feedback network and of a discriminator in the same processing chain.

onto the ramp (see Fig. 10). No difference can be detected in the peaking time. Just a small change in the slope of the falling edge and in the undershoot is visible. Fig. 12 shows the peak response in the same CSA as a function of the input charge, again with no input leakage current and for I_L=10 nA, for two different values of C_D^*. The presence of the leakage current has a very mild effect on the charge sensitivity, with a change of less than 3% in the slope of the curves. In the case of C_D^*=30 fF, a slight decrease in G_Q is detected, consistent with what discussed in Section II-A. In the case C_D^*=130 fF, instead, a larger G_Q value is detected for I_L=10 nA than for

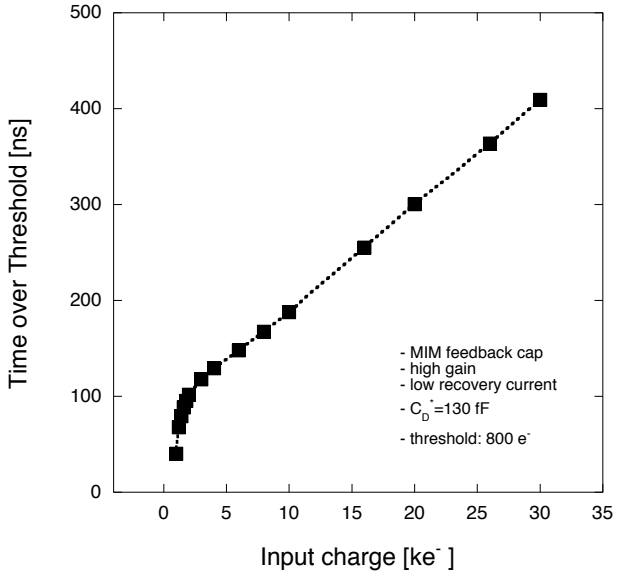

Fig. 14. Time-over-threshold as a function of the input charge, determined as the time duration of the signal at the output of the same discriminator as in Fig. 13.

I_L=0, a result which seems to be inconsistent with the above considerations. Actually, as shown by the inset in Fig. 12, G_Q is found to be smaller at larger I_L for small values of the input charge. Here the CSA is operated closer to its linear region, the same region in which (4) was derived. Very likely, for large values of the input charge, the non-linear behavior of the preamplifier significantly modifies the slope of the straight line interpolating the data points, overriding the small leakage related effect.

C. Threshold discriminator response and ToT

The response of a threshold discriminator, from one of the chips under test, in the case of an input signal of 10 ke$^-$, is shown in Fig. 13, together with the response of the CSA in the same processing chain. The duration of the signal at the discriminator output, corresponding to the time-over-threshold, is also shown in Fig. 14 as a function of the input charge. The measurements displayed in the two figures were obtained with the CSA in the standard configuration. The time-over-threshold is very close to 400 ns for an input charge of 30 ke$^-$. Moreover, the ToT vs input charge curve is fairly linear for charge values exceeding 2 ke$^-$, with a slope of about 11 ns/ke$^-$. To account for the behavior of the curve below 2 ke$^-$, one should consider that, for small input charge, the rising time of the CSA output signal cannot be neglected (as it was done in (10)) and that, in addition, the feedback capacitor is not discharged linearly, but exponentially. These two effects are responsible for the non-linearity detected in the initial portion of the characteristic.

D. Discriminator speed and time walk

The front-end channel and, in particular, the response of the threshold discriminator, are affected to some extent by time walk phenomena, i.e., the delay of the discriminator response

TABLE II
DELAY OF THE SIGNAL AT THE DISCRIMINATOR OUTPUT FOR A THRESHOLD OF 600 e^- AND A CHARGE OVER THRESHOLD OF 600 e^- ($< delay > \pm\sigma(delay)$, IN ns).

	chip 10		chip 11	
	PMOS cap	**MIM cap**	**PMOS cap**	**MIM cap**
$\mathbf{C_D^* = 30fF}$	18.7±1.5	19.3±1.4	19.0±3.1	19.0±1.4
$\mathbf{C_D^* = 80fF}$	22.8±2.3	24.2±2.3	22.8±3.5	23.4±2.4
$\mathbf{C_D^* = 130fF}$	26.9±3.6	28.7±3.8	26.5±3.6	27.8±3.6

with respect to the time of arrival of the event depends on the input signal amplitude. Time walk spoils the time correlation between the signal at the channel output and the triggering event. Therefore, at the HL-LHC, time walk must be minimized taking into account the 25 ns bunch crossing period. In particular, a 600 e^- overdrive (i.e., signal over threshold) is specified for the readout system at a detector capacitance of 50 fF. For a 600 e^- minimum threshold setting, this means that the channel has to be designed in such a way that a 1200 e^- signal will have a 50% probability of firing the discriminator within a time interval of 25 ns from the largest expected signal, 30 ke^-. Table II shows the delay (the average measured value $\pm$ the standard deviation in the measurement) of the discriminator signal with respect to the input signal for a threshold of 600 e^- and a charge over threshold of 600 e^-. This is defined as the interval between the time at which the signal at the discriminator output crosses half of the high logic level and the time at which the input step signal reaches 50% of its final level. For a CSA with PMOS feedback capacitance, the response to a 30 ke^- signal at C_D^*=130 fF features a delay of around 17 ns. This is about 10 ns smaller than the delay in the response to a 1200 electron signal, as displayed in Table II for the same C_D^*, also corresponding to the worst case among those in the table. Therefore, the prototype under test falls well within the specifications, even for detector capacitances significantly larger than 50 fF. This means that the present design of the discriminator still offers some room for reducing the power dissipation.

E. Equivalent noise charge

The equivalent noise charge as a function of C_D^* is shown in Fig. 15 for a CSA with PMOS capacitor in the feedback network. Very similar results were obtained in the case of the CSA with MIM feedback capacitor. The performance of the preamplifier was studied in the four possible combinations of the gain and recovery current settings. The high gain/low recovery current configuration results in the best ENC performance, with an equivalent noise charge of about 120 e^- rms at C_D^*=100 fF. In the same configuration, the $dENC/dC_D$ slope is about 310 e^-/pF. In Fig. 15, a non negligible increase in the ENC can be detected when switching from the high gain to the low gain configuration, i.e., when C_F is increased from 10 to 20 fF. This result is consistent with (18) and (19), showing that the equivalent noise charge, through its main contribution ENC_0, is proportional to $\sqrt{(C_F)}$. Also switching to a larger

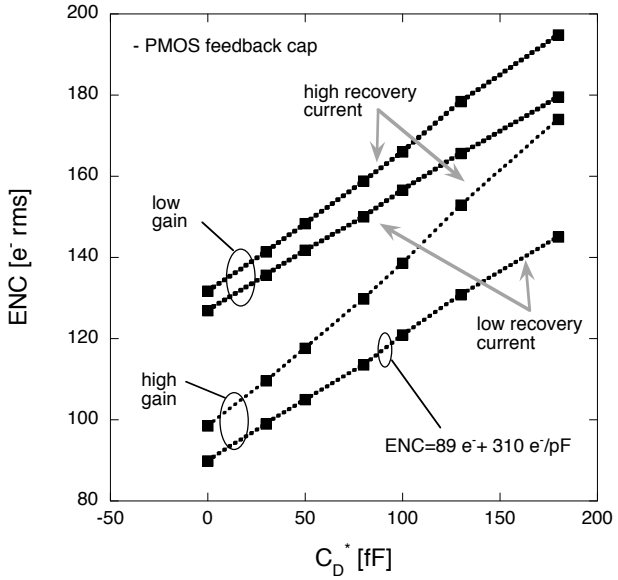

Fig. 15. Equivalent noise charge as a function of C_D^* for the four different possible configurations of a charge preamplifier with PMOS capacitor in the feedback network.

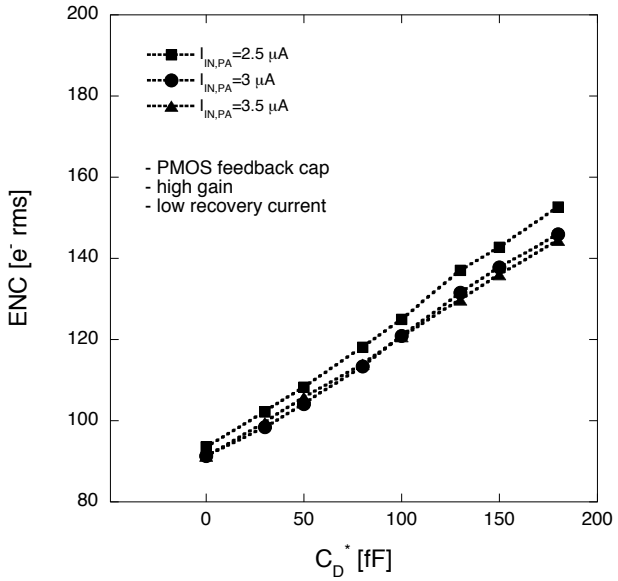

Fig. 16. Equivalent noise charge as a function of C_D^* for a charge preamplifier with PMOS capacitor in the feedback network in its standard configuration. The ENC is plotted for three different values of the current in the CSA input device.

recovery current, which implies increasing g_{m_n} (and g_{m_p}), has the effect of increasing the ENC. In particular, Fig. 15 shows that, for a given gain, the change due to the increase in I_K grows with C_D^*. This result is again in agreement with (18), where the term adding to ENC_0^2 is roughly proportional to $g_{m_n} C_D'$ for large enough values of C_D'.

The equivalent noise charge was also measured for three different values of the drain current, $I_{in,PA}$, in the CSA input device. A set of measurements, performed for a stage with a PMOS feedback capacitor in its standard configuration is shown in Fig. 16. Very small changes are detected when the

current is shifted 500 nA above or below its design value, 3 μA. As a matter of fact, a variation in $I_{in,PA}$ involves a change in the transconductance g_m of the CSA input device, which in turn affects the GBP. Such a change, however, has no significant impact on the equivalent noise charge, at least in the limited current range explored here, because the main contribution to the ENC, ENC_0, depends on the GBP/g_m ratio, which is a constant function of g_m. This points to the fact that, for a change in the input device current, the variation in the channel thermal noise is compensated for by a counteracting change in the preamplifier gain-bandwidth product.

III. Conclusions and Future Work

A readout channel for pixel detectors has been designed for the phase II upgrades of the CMS experiment at the HL-LHC. The circuit includes a charge preamplifier, based on a Krummenacher feedback network, followed by a fast threshold discriminator. Amplitude-to-digital conversion is to be performed through a time-over-threshold technique. The paper presented the main design features of the overall channel and the results from the characterization of both the charge preamplifier and the threshold discriminator. The CSA is available in two different variants, differing in the way the feedback capacitor is implemented, with a PMOS transistor or a MIM structure. The overall performance of the readout channel, for both the versions of the CSA, complies with the challenging specifications set by the HL-LHC experiment upgrades. ENC in the 100 e^- range has been measured for a detector capacitance of 50 fF. The CSA, taking advantage of a Krummenacher feedback network, can deal with relatively large leakage currents from the detector, in excess of 15 nA, with no significant change in the response. Small dispersion in the charge sensitivity, 1% or smaller, was detected for both channel variants. The time response of the overall processing chain, including the threshold discriminator, is fully compatible with the time-walk requirements of the application. The models presented and discussed in the paper for charge sensitivity and equivalent noise charge are in good agreement with the measurement results. Based on the experimental data, none of the two variants can be claimed to be clearly superior to the other. Therefore, the CSA version with MIM feedback capacitor is going to be dropped from the next iteration of the readout channel, so that MIM structures can be removed from the layer stack and lower production costs can be achieved. The performance of the readout channel presented in this work has also been validated through a couple of radiation tolerance campaigns, where a number of samples has been characterized before and after exposure to very high doses of ionizing radiation, up to 800 Mrad. The circuit has already been integrated in the CHIPIX65 demonstrator, a 64×64 readout chip now in production, and will be integrated, with a few slight modifications, in the RD53A demonstrator, a larger chip including 400×192 elements and developed by the RD-53 collaboration.

Acknowledgment

The research activity presented in this paper has been carried out in the framework of the CHIPIX65 experiment funded by the Italian Institute for Nuclear Physics (INFN). It has also received funding from the European Union's Horizon 2020 Research and Innovation program under Grant Agreement no. 654168.

References

[1] RD-53 Collaboration website [Online]. Available: https://rd53.web.cern.ch/RD53/.

[2] L. Pacher, E. Monteil, A. Rivetti, N. Demaria, M.D. Da Rocha Rolo, "A Low-Power Low-Noise Synchronous Pixel Front-End Chain in 65 nm CMOS Technology with Local Fast ToT Encoding and Autozeroing for Extreme Rate and Radiation at HL-LHC", presented to the 2015 IEEE Nuclear Science Symposium and Medical Imaging Conference, San Diego, USA, 31 October-7 November 2015.

[3] F. Krummenacher, "Pixel detectors with local intelligence: an IC designer point of view", *Nuclear Instrum. and Methods* , vol. 305, no. 3, pp. 527-532, 1991.

[4] F. Maloberti, *Analog design for CMOS VLSI systems*, Kluwer Academic Publishers, Boston, 2001.

[5] P.M. Furth, A.G. Andreou, "Linearised Differential Transconductors in Sub-threshold CMOS", *Electron. Lett.*, vol. 31, no. 7, pp. 545-547, Mar 1995.

[6] L. Ravezzi, G.-F. Dalla Betta, "Simple high-speed CMOS current comparator", *Electron. Lett.*, vol. 33, no. 22, pp 1829-1830, Oct. 1997.

[7] L. Ratti, M. Manghisoni, V. Re, G. Traversi, "Discriminators in 65 nm CMOS process for high granularity, high time resolution pixel detectors", *2013 IEEE Nuclear Science Symposium Conference Record*, p. 1-6, Seoul, South Korea, Oct. 27 2013-Nov. 2 2013.

[8] R.L. Bunch and S. Raman, "Large-Signal Analysis of MOS Varactor's in CMOS - G_m LC VCOs", *IEEE J. Solid-State Circuits*, vol. 38, no. 8, pp. 1325-1332, Aug. 2003

[9] S. Wonsak, A. Affolder, G. Casse, P. Dervan, I. Tsurin, M. Wormald, "Measurements of the reverse current of highly irradiated silicon sensors", *Nucl. Instrum. Methods*, vol. A831, pp. 105-110, 2016.

[10] R. Klingenberg, S. Altenheiner, D. Bryan, S. Dungs, A. Gisen, C. Gössling et al., "Power dissipation studies on planar n^+-in-n pixel sensors", *Nucl. Instrum. Methods*, vol. A796, pp. 126-130, 2015.