



HAL
open science

Detecting cells rotations for increasing the robustness of cell sizing by impedance measurements, with or without machine learning

Pierre Taraconat, Jean-philippe Gineys, Damien Isèbe, Franck Nicoud, Simon Mendez

► To cite this version:

Pierre Taraconat, Jean-philippe Gineys, Damien Isèbe, Franck Nicoud, Simon Mendez. Detecting cells rotations for increasing the robustness of cell sizing by impedance measurements, with or without machine learning. *Cytometry Part A*, 2021, 99 (10), pp.977-986. 10.1002/cyto.a.24356 . hal-03228679

HAL Id: hal-03228679

<https://hal.science/hal-03228679>

Submitted on 19 Oct 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

RESEARCH ARTICLE

Detecting cells rotations for increasing the robustness of cell sizing by impedance measurements, with or without machine learning

Pierre Taraconat^{1,2} | Jean-Philippe Gineys¹ | Damien Isebe¹ | Franck Nicoud² | Simon Mendez*²

¹Horiba Medical, Parc Euromédecine, Rue du Caducée, BP 7290, 34184 Montpellier Cedex 4, France

²Institut Montpellierain Alexander Grothendieck, CNRS, Univ. Montpellier, Place Eugène Bataillon, 34095 Montpellier Cedex 5, France

Correspondence

*Simon Mendez, Institut Montpellierain Alexander Grothendieck. Email: simon.mendez@umontpellier.fr

Abstract

The Coulter principle is a widespread technique for sizing red blood cells in hematology analyzers. It is based on the monitoring of the electrical perturbations generated by cells passing through a micro-orifice, in which a concentrated electrical field is imposed by two electrodes. However, artefacts associated with near-wall passages in the sensing region are known to skew the statistics for red blood cells sizing. This study presents numerical results that emphasize the link between the cell flow-induced rotation in the detection area and the error in its measured volume. Based on these observations, two methods are developed to identify and reject pulses impaired by cell rotation. In the first strategy, the filtering is allowed by a metric computed directly from the waveform. In the second, a numerical database is employed to train a neural network capable of detecting if the cell has experienced a rotation, given its electrical pulse. Detecting and rejecting rotation-associated pulses is shown to provide results comparable to hydrodynamical focusing, which enforces cells to flow in the center of the orifice, the gold standard implementation of the Coulter principle.

KEYWORDS:

Red blood cells; Computational fluid dynamics; Coulter principle ; Pulse editing; Hydrodynamical focusing; Neural Network

1 | INTRODUCTION

Complete blood count, which notably consists in counting and sizing blood cells, is widely employed to assess the patients' health condition. Such an analysis is performed by industrial hematological automata, that generally implement the Coulter principle¹. As sketched in Fig. 1, cells in suspension in an electrolytic solution are aspirated through a micro-orifice (aperture) in which electrodes maintain a concentrated electrical field. A cell passing through the aperture is detected by an increase of voltage at the electrodes. In addition, the maximum of this electrical pulse is assumed to be proportional to the cell volume. Screening pulses thus yields the count and the volume distribution of the analyzed cells.

Regarding red blood cells (RBCs), the Mean Corpuscular Volume (MCV) and the RBCs Distribution Width (RDW) are haematological parameters of first importance for clinical decision making. In particular, the combination of MCV and RDW was historically used for classifications of anemias², but RDW has more recently aroused the interest of many research groups as a marker of various human disorders³.

According to theory⁴, the resistive perturbation (ΔR) produced by the cell in the detection area is related to its volume V_p as follows, assuming a perfectly insulating particle:

$$\Delta R \propto E^2 f_s V_p \quad (1)$$

E is the electrical field and f_s the particle shape factor. Note that the voltage pulses are proportional to ΔR , as the current intensity remains constant. The shape factor characterizes how easy it is for the electrical current to go around the insulating particle, given its volume. It depends on the particle shape and

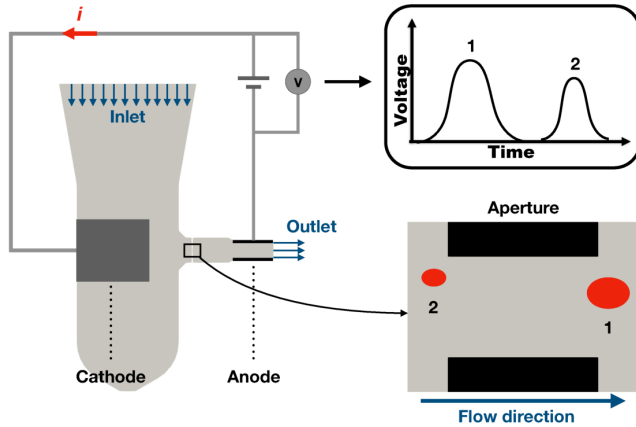


FIGURE 1 Diagram of the Coulter principle for particles counting and sizing.

orientation. For the case of deformable cells such as RBCs, f_s depends on the dynamics and deformations of the cell in the detection area. Pathologies altering the RBCs deformability and/or resting shape may thus lead to misleading diagnosis^{5,6}. For instance, RBCs with abnormally high hemoglobin contents in their cytoplasm have their volume overestimated⁵. Furthermore, artefacts inherent to the implementation of the Coulter principle alter the measurements and lead to right-skewed RBC volume distributions^{7,8} instead of the expected symmetrical distribution^{9,10}, even for normal blood samples. Indeed, measuring cells volume in a sample requires that all cells experience the same electrical field and behave identically within the aperture (in other words that all cells are associated with the same E and f_s , see Eq. 1). However, as reported in former studies, E is highly heterogeneous next to the aperture walls^{4,11,12}, and RBCs may rotate and deform as they flow across high-shear regions in the wall vicinity^{11,13,14}.

In flow cytometry, errors associated with the random orientation of asymmetrical cells with respect to the flow direction have long been understood. For instance, light scattering depends on the orientation of the cell with respect to the incident laser beam and to the position of the detectors¹⁵. Hardware modifications may be made so that asymmetrical cells always present the same orientation to the observation device¹⁶. For impedance measurements, the random orientation of cells with respect to the flow direction does not generate any error: it is the flow-induced rotation of cell that changes the shape factor^{11,13,14}.

For impedance measurements, hydrodynamical focusing¹⁷ overcomes measurement inaccuracies caused by cells flowing next to the walls, so that symmetrical and Gaussian-like distributions are retrieved. The principle is to pinch the sample flow with a secondary sheath flow, thus enforcing particles

to follow centered trajectories. This enables standardization of f_s and E for a robust measurement, but implies a more intricate implementation of the Coulter principle¹⁸. Another strategy consists in removing the distorted pulses associated with cells passing near of the aperture edges from the analysis (also called pulse-editing). We provide here some examples of the numerous techniques introduced over the years. Remarking that centered paths generate ‘bell-shaped’ and symmetrical signatures, Dunstan *et al*¹⁹ proposed to assess the pulse symmetry for rejecting spurious signals. Another method was to exclude long pulses⁸: near the walls, flow speed is lower than at the center of the orifice; thus, RBCs impacted by near-wall effects are expected to produce longer electrical prints. However, the pulse duration depending on the particle volume, there is a relationship between the duration and the maximum of the pulse, which is disregarded with a simple threshold on the pulse duration. This explains the introduction of new pulse-editing strategies taking into account both the pulse maximum and the pulse duration^{20,21}. Methods combining several widths and/or several amplitudes calculated from pulses are currently implemented in commercial instruments, as suggested by the patent from Zhang *et al*²². Note that pulse editing is a more general strategy, which has for instance been used to measure the orientation of asymmetrical cells¹⁵ or identify doublets in optical flow cytometers²³.

Easier to implement and cheaper, classical systems combined with pulse editing continue to be used, in spite of the higher robustness of hydrodynamical focusing. Among current instruments, different pulse editing methods are used, but they are not necessarily well documented. While the MCV is usually a robust measurement, RDW has been recently shown to depend on the analyzer²⁴. To avoid such discrepancies, it is essential to control volume errors and properly establish the quality of pulse editing methods. In particular, while hydrodynamical focusing can be considered as the reference²⁵, it is rarely used to assess the accuracy of pulse editing algorithms.

Industrial instruments generally suffer from accessibility issues, which limits the characterization of measurement artefacts and thus the design of pulse-editing procedures. Recently, numerical simulation has been found to reproduce the electrical signatures observed in industrial instruments. Building upon this proof of concept (described in the Material and Methods, Sec. 2), the present work presents a series of simulations of electrical signatures for different trajectories of RBCs inside the sensing region of a YumizenTM H2500 (HORIBA Medical, Montpellier, France), which is based on the Coulter principle. Results highlight the link between measurement errors and the flow-induced rotations of the cells which yields a secondary peak on the pulses (see Sec. 3.1). This observation is the foundation for two new pulse-editing methods²⁶, one

based on the use of a neural network and the other on a metric explicitly calculated on the pulses, designed to reject pulses associated with RBCs that rotate in the sensing region (see Sec. 2). When applied to experimental data (Sec. 3.2), both methods yield symmetrical volume distributions in agreement with results obtained with hydrodynamical focusing.

2 | MATERIALS AND METHODS

2.1 | Numerical simulations

Numerical pipeline

Computing the dynamics of an RBC in the detection area is extremely challenging due to the separation of scales: RBCs are first stretched for tens of milliseconds, then pass through the sensing zone in a few microseconds¹². To overcome this difficulty, an innovative numerical pipeline has been designed¹⁴ to predict the dynamics of RBCs in the sensing zone and the corresponding electrical pulse, using the in-house solver YALES2BIO (<https://imag.umontpellier.fr/~yales2bio/>). The RBC dynamics is predicted by solving a fluid-structure interaction problem^{27,28,29} between the fluid outside the RBC, the cytoplasm, and the RBC membrane. Then, from a resulting sequence of RBC states along time inside the orifice, as many electrostatic computations are performed. In each of them, one RBC position/shape is used, the cell being assumed insulating. The electrical field around the insulating cell is computed, and the additional resistance due to the presence of the cell (ΔR) is obtained by comparison with a calculation without RBC. Hence, calculating ΔR for the sequence of RBC consecutive positions leads to the resistive pulse over time, associated with the RBC path and deformation in the aperture. This numerical setup has been detailed and validated previously¹⁴. More details on the numerical setup can be found in the Supplementary Materials.

Simulated cases

In this work, the role of numerical simulations is twofold: understanding measurement inconsistencies induced by cells flowing near the aperture edges ; building a numerical database for the training of a Neural Network (NN) capable of rejecting pulses irrelevant to volume assessments.

With the aim of investigating the impact of RBC paths on the impedance measurements, several trajectories are simulated while maintaining cell parameters constant. The RBC is represented as a discoid of volume $93 \mu\text{m}^3$, whose shape has been defined analytically³⁰. The RBC has an aspect ratio Q of 0.65, with Q defined as the ratio between the cell volume and the volume of the sphere having the same surface area as the RBC. The mechanical model of the RBC membrane reproduces its

composite structure, with different sub-models representing the lipid bilayer of the membrane and its cytoskeleton^{31,32}. This modeling is now classical and has been successfully applied and validated in numerous situations^{14,27,28,29}. The membrane mechanics is notably parametrized by the shear modulus G_s , the area modulus E_a , the bending modulus E_b . In agreement with measurements available in the literature^{33,34}, we set: $G_s=2.5\times 10^{-6} \text{ N m}^{-1}$; $E_a=2.5\times 10^{-1} \text{ N m}^{-1}$; $E_b = 6.0\times 10^{-19} \text{ N m}$. The viscosity of the suspending medium is $\nu_{ext} = 10^{-6} \text{ m}^2 \text{ s}^{-1}$ and its conductivity $\sigma_{ext} = 2.27 \text{ S m}^{-1}$. Regarding the fluid inside the RBC membrane, a viscosity ν_{in} of $18.0\times 10^{-6} \text{ m}^2 \text{ s}^{-1}$ is set which is representative of viscosity values of hemoglobin solutions at ambient temperature³⁴. Ten different trajectories are investigated, from the axis of the orifice to the near-wall region.

A numerical database is built for training a NN capable of detecting if an RBC rotates in the aperture, given the associated electrical pulse. In order to make the NN robust with respect to cells rheology and morphology, RBC parameters were varied in the following ranges to build the database: $\nu_{in} \in [15.0\times 10^{-6} \text{ m}^2 \text{ s}^{-1}, 20.0\times 10^{-6} \text{ m}^2 \text{ s}^{-1}]$; $G_s \in [2.5\times 10^{-6} \text{ N m}^{-1}, 160.0\times 10^{-6} \text{ N m}^{-1}]$; $V_p \in [85 \mu\text{m}^3, 100 \mu\text{m}^3]$; $Q \in [0.65, 0.85]$. Altogether, the database is composed of 78 cases.

All simulations are performed by considering the industrial geometry and operating regime provided by HORIBA Medical and illustrated in Fig. 1, which characterize the YumizenTM H2500 (HORIBA medical, Montpellier, France). More precisely, the aperture has a diameter of $50 \mu\text{m}$, a length of $75 \mu\text{m}$ and the pressure drop ensuring the flow is of 200 mbar.

2.2 | Detection of the particle flow-induced rotation

As reported in the literature^{10,11}, complex pulses shapes are obtained when deformable and aspherical particles flow near the aperture wall, while ‘bell-shaped’ signatures are obtained for centered paths. Previously¹⁴, we directly showed that RBCs deform and, most importantly, rotate when flowing near the wall, which induces a peak on the electrical pulse at the exact moment of the flow-induced rotation. Golibersuch¹³ also reported pulses presenting several peaks when the aperture is long enough to allow RBCs to rotate several times in the sensing region. Our pulse editing procedures consist in detecting such rotations on the electrical pulses, in order to only conserve ‘bell-shaped’ signatures from centered trajectories.

In the first method, ‘bell-shaped’ signatures are isolated by thresholding a metric directly calculated from the pulses. With \mathcal{H} the maximum of the pulse, whatever the quantity analyzed (voltage or resistance), let us define \mathcal{W}_u and \mathcal{W}_d , two pulse widths computed with thresholds $\mathcal{H} \times p_u$ and $\mathcal{H} \times p_d$, respectively (see Fig. 2). p_u and p_d are included in $[0,1]$, with $p_u >$

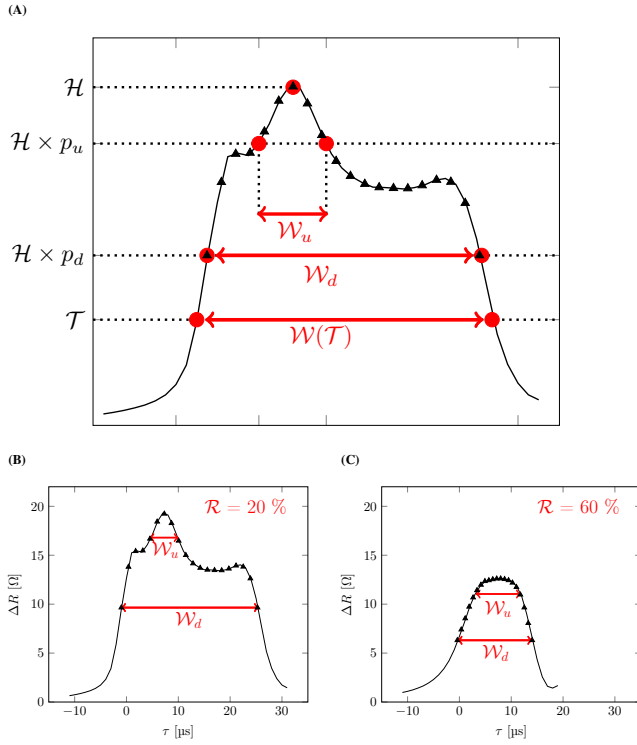


FIGURE 2 (A) Illustration of quantities computed from an electrical pulse. In particular, \mathcal{W}_u and \mathcal{W}_d required for assessing the metric \mathcal{R} are depicted. Let us define in general $\mathcal{W}(\mathcal{T})$ as the pulse width calculated with a threshold \mathcal{T} . Then, we define $\mathcal{W}_u = \mathcal{W}(\mathcal{H} \times p_u)$ and $\mathcal{W}_d = \mathcal{W}(\mathcal{H} \times p_d)$, with \mathcal{H} the pulse maximum. Calculations of \mathcal{R} for a complex signature presenting a peak and a ‘bell-shaped’ pulse are given in (B) and (C), respectively. Black triangles depicted in the graphs illustrate the resampling performed on the waveforms in order to make them intelligible by the NN. This is required to represent all pulses with the same number of variables.

p_d . We introduce \mathcal{R} , the ratio between \mathcal{W}_u and \mathcal{W}_d :

$$\mathcal{R} = \frac{\mathcal{W}_u}{\mathcal{W}_d} \times 100 \quad (2)$$

The parameter p_d is chosen in such a way that $\mathcal{H} \times p_d$ intersects the ascending and descending slopes of the pulse, and \mathcal{W}_d informs on the time spent by the particle in the micro-orifice. If a secondary peak is present on the pulse, \mathcal{W}_u is meant to measure its duration. Hence, p_u must be defined so that $\mathcal{H} \times p_u$ crosses the peak. Note that in the absence of peak, \mathcal{W}_u should be closer to \mathcal{W}_d . Hence, \mathcal{R} is expected to be small in the case of a flow-induced rotation peak (see Fig. 2B) and larger for a ‘bell-shaped’ pulse without peak (see Fig. 2C). The filtering then consists in applying a rejection threshold on the metric \mathcal{R} . In practice, we have used $p_u = 7/8$ and $p_d = 1/2$. However,

tests shown as Supplementary Materials support that the quality of the results does not depend on the exact values of these thresholds.

In the second method, the filtering is performed by a Neural Network (NN). A simple Multi-Layer Perceptron (MLP) of 3 layers is sufficient to address this clustering problem. The first layer is composed of 21 neurons and takes the pulse as an input. Pulses are resampled into 20 values by taking into account the part of the curve located above $\mathcal{H} \times p_d$, as illustrated by black triangles in Fig. 2A. These 20 values are normalized by \mathcal{H} and given as inputs to the NN, in addition to \mathcal{W}_d . This way, pulses of different durations are all expressed with the same number of variables (see Fig. 2B and C). The reworked pulse is propagated into the NN through the single hidden layer of 20 neurons toward the output layer, composed of a sole neuron. Sigmoid activation functions are associated to all neurons of the model. Hence, the NN renders a score included between 0 and 1. A database of 78 numerical pulses is considered for the model training. Pulses are associated to target values of 0 or 1 whether the RBC has turned in the aperture or not. Then, numerical pulses are split into a training dataset made of 62 observations and a test dataset of 16 pulses. The learning step is done on the training dataset in 10000 epochs with a batch size of 62 (the size of the training dataset) with the ADAM algorithm (a variant of the gradient descent method) which aims at minimizing the RMS error between NN predictions and target values. When applied to the test dataset, the NN is found to properly isolate pulses associated to rotating RBCs from others by applying a threshold of 0.5 on NN scores. More information about the database and the preliminary tests is provided as Supplementary Materials. The NN implementation is done by means of the python library TensorFlowTM³⁵ (GOOGLE Inc.) which is publicly available (<https://www.tensorflow.org>).

2.3 | Experimental tests

Experiment overview

To assess the accuracy of the pulse-editing methods, results from an analyzer with the classical implementation of the Coulter principle combined with pulse editing are compared for each blood sample with measurements from a device using hydrodynamical focusing. Distribution of measured volumes, MCV and RDW are considered for the comparisons.

Flow sample and specimen description

Human peripheral blood samples were collected from 20 healthy donors in K3EDTA tubes (VACUETTETM) and analyzed within the 6 hours after withdrawal. They were obtained from Etablissement Français du Sang, Montpellier, France.

Instrument

The analyses are performed with a Yumizen™ H2500 commercialized by HORIBA Medical. In the present work, two units of the Yumizen H2500 are considered: the device dedicated to the RBCs analysis based on the Coulter principle, and the system normally dedicated to the optical platelet detection, in which hydrodynamical focusing is used. In the first one, RBCs are counted and sized by the use of a classical implementation of the Coulter principle (as in Fig. 1). Here, the second unit is not used for its original purpose, but to count and size RBCs with hydrodynamical focusing, to provide reference results. For each blood sample, electrical pulses coming from the two units are recorded separately. This is done by a dedicated in-house LabVIEW™ (National Instruments) code, detailed in the Supplementary Materials. Acquisitions from the unit classically dedicated to RBCs are labelled noHF, while the acquisitions from the hydrofocused platelet detector used here for sizing RBCs are referred to as HF. Numerical simulations are relevant to the classical implementation of the Coulter principle, thus to measurements labelled by noHF.

3 | RESULTS

3.1 | Detection of the particle flow-induced rotation from the associated pulse

This section introduces the fundamental concept from which the filters of Sec. 2.2 are designed. More precisely, the impact of the flow-induced cell rotation on the measured volume and its relationship with the trajectory are explained, keeping the intrinsic properties of the RBC strictly identical. Considering the 10 possible RBC trajectories depicted in Fig. 3A, electrical signatures are computed (Fig. 3B). The non-dimensional resistive perturbations ΔR^* are obtained by scaling ΔR by the pulse maximum (viz. \mathcal{H}) of the centered path (case 1), which is supposed to be a correct measure for the RBC volume, free of edge effects. As the same RBC is considered for all cases, \mathcal{H}^* , the maximum of ΔR^* pulses, informs about the errors in the volume measurement associated with the RBC trajectory. In an ‘ideal’ system, $\mathcal{H}^* = 1.0$ independently of the trajectory.

The closer to the orifice edges is the particle path, the longer the pulse duration, which makes it easy to order the pulses. For the most centered paths (cases 1 to 4), the pulses are "bell-shaped" while more complex signatures are observed for the near-wall trajectories (cases 5-10). Figures 3C and 3D show RBC consecutive positions in the orifice for cases 1 and 10, respectively. The RBC flowing along the axis (Fig. 3C) displays an elongated shape aligned with the orifice axis. In contrast, an RBC evolving near the aperture edges rotates, which induces a peak on the electrical perturbation, as highlighted

by the black arrows in Fig. 3D and 3E. In industrial instruments, high-speed flows are used to ensure a high throughput. Hence, high shear stresses result in rotation and deformation of the RBCs, which cause an increase of shape factor f_s and a peak on the electrical pulse. The time of exposure to the shear and the shear level depend on the trajectory, which explains the variety of electrical signatures (cases 5-10, Fig. 3B).

In the Coulter principle, the pulse maximum is supposed to be proportional to the particle volume. Figure 3B thus shows that the apparent volume of an RBC depends on the RBC trajectory. For instance, comparing the maxima of cases 1 and 10 in Fig. 3B demonstrates that an RBC flowing near the wall is seen 1.53 times bigger than if it had flowed near the aperture axis. Figure 4A displays the different non-dimensional measured volumes \mathcal{H}^* as a function of the cell distance to the aperture axis (D), for different internal viscosities ν_{in} . Regarding case $\nu_{in} = 18.0 \times 10^{-6} \text{ m}^2 \text{ s}^{-1}$ in Fig. 4A, \mathcal{H}^* is close to 1.0 in the core region of the aperture ($D \in [0 ; 13 \mu\text{m}]$). With the perfectly centered path as the reference, the error at $D = 13 \mu\text{m}$ is about 3%. \mathcal{H}^* increases with D between 13 μm and 17 μm , then stabilizes around a value of 1.53. These volume overestimates for near-wall trajectories ($D > 13 \mu\text{m}$) explain the typical ‘right-skewed’ volume distributions observed experimentally^{7,8}. Simulations also show that for near-wall trajectories, the apparent volume increases with the internal viscosity (see Fig. 4A). This confirms the experiments of Mohandas *et al*⁵ which showed an overestimation of the MCV when the haemoglobin content of RBCs is abnormally high (which corresponds to high cytoplasm viscosity). Consequently, removing near-wall trajectories has two advantages: avoiding sizing errors in general and making measurements robust to variations in the haemoglobin content of the RBCs.

As defined in Fig. 4B, the orientation of the cell when the pulse maximum is reached, θ_m , is computed and displayed in Fig. 4C. If $D > 16 \mu\text{m}$, the time spent by the cell in the detection area is sufficiently long and the shear undergone by the particle large enough for the RBC to reach a 90° orientation inside of the aperture: the RBC is then perpendicular to the electrical field. On the contrary, in the core region of the micro-orifice, the cell barely rotates and $\theta_m \approx 0^\circ$ when $D < 13 \mu\text{m}$. Between 13 μm and 16 μm , the cell starts rotating, but does not spend enough time in the aperture to reach $\theta_m = 90^\circ$. The fact that both \mathcal{H}^* and θ_m display a similar sigmoid-like profile (see Fig. 4A and C) supports the approach of detecting the flow-induced RBC rotation for pulse edition to improve the volume assessment.

The evolution of the metric \mathcal{R} (see Eq. 2) with D is shown in Fig. 4D. When θ_m is small ($D < 13 \mu\text{m}$), high values of \mathcal{R} are obtained (between 60 and 70%), while for cases where the RBCs may reach $\theta_m = 90^\circ$ ($D > 16 \mu\text{m}$), \mathcal{R} is below 30%. By comparing Fig. 4A and D, it is seen that rejecting pulses for

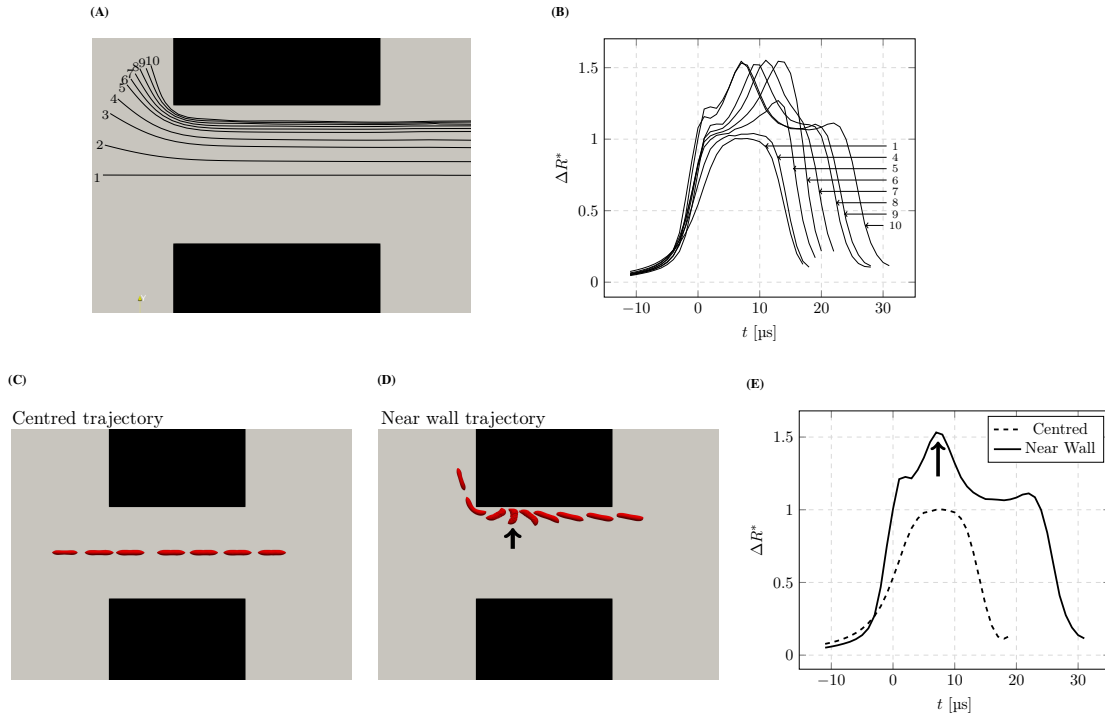


FIGURE 3 Simulated impedance pulses for different virtual trajectories of an RBC in the sensing region. All results in this figure were obtained with the same RBC characteristics: $v_{in}=18.0 \times 10^{-6} \text{ m}^2 \text{ s}^{-1}$, $G_s=2.5 \times 10^{-6} \text{ N m}^{-1}$, $E_a=2.5 \times 10^{-1} \text{ N m}^{-1}$, $E_b = 6.0 \times 10^{-19} \text{ J}$, $V_p=93 \text{ } \mu\text{m}^3$ and $Q = 0.65$. The considered trajectories are shown in (A) and the corresponding pulses in (B). The pulses amplitudes in (B) are scaled with the maximum of case 1 (the centered path) which equals $12.6 \text{ } \Omega$. For sake of clarity, cases 2 and 3 are not shown in (B). Pictures (C) and (D) show a series of RBC consecutive positions inside the aperture for cases 1 and 10, respectively. Figure (E) recall the electrical signatures related to RBC dynamics of (C) and (D). The cell rotation (D) generates a peak on the associated impedance pulse, as highlighted by black arrows in (D) and (E).

which \mathcal{R} is below 55% would for instance allow reducing the volume overestimation (\mathcal{H}) from 53 % to 10 %. Thus, a pulse editing method with a threshold on \mathcal{R} is expected to separate pulses from rotating RBCs and ‘bell-shaped’ pulses, for which RBCs do not rotate. For the NN-based pulse editing method, each numerical pulse is associated with a target NN score of 1 or 0, whether θ_m is lower or higher than 15° , respectively.

Our pulse editing algorithms have been designed using simulated data. As a consequence, their success depends on how accurate and representative the numerical results are with respect to real data. In a previous study, we have proven that simulation is able to predict relevant pulses both for RBCs and rigid spherical beads¹⁴. Further comparisons for RBCs are presented in the Supplementary Materials to show the excellent agreement between pulse shapes. Overall, all simulated pulse shapes are found in the experiments, and they are representative of the vast majority of the actual pulses (about 92 %). However, experiments may feature other pulses, corresponding for instance to doublets or other blood cells, that can be easily separated from the RBC pulses.

Pulse editing methods based on the detection of the flow-induced RBC rotation are next tested on experimental data to quantify the improvement in the accuracy of the haematological parameters related to the RBCs volumes, the MCV and the RDW.

3.2 | Application to experimental data

Let us consider two acquisitions performed on the same blood sample, but obtained with the classical (noHF) and the hydro-focused systems (HF). For each pulse coming from the classical acquisition, the maximum \mathcal{H} of the voltage perturbation, the width ratio \mathcal{R} and the NN score are computed. Figure 5 shows the scatter plot of \mathcal{R} as a function of \mathcal{H} by colouring the data points according the NN scores, for the acquisition noHF. Two distinct clusters separated by a threshold of $\mathcal{R} = 52 \%$ are observed in Fig. 5. This threshold is close to the value of 55 % that was predicted numerically. Note that for the cluster located below $\mathcal{R} = 52 \%$, \mathcal{H} is between 1.6 V and 3.0 V, while for the cluster above $\mathcal{R} = 52 \%$, \mathcal{H} is between 1.0 V and 2.25 V.

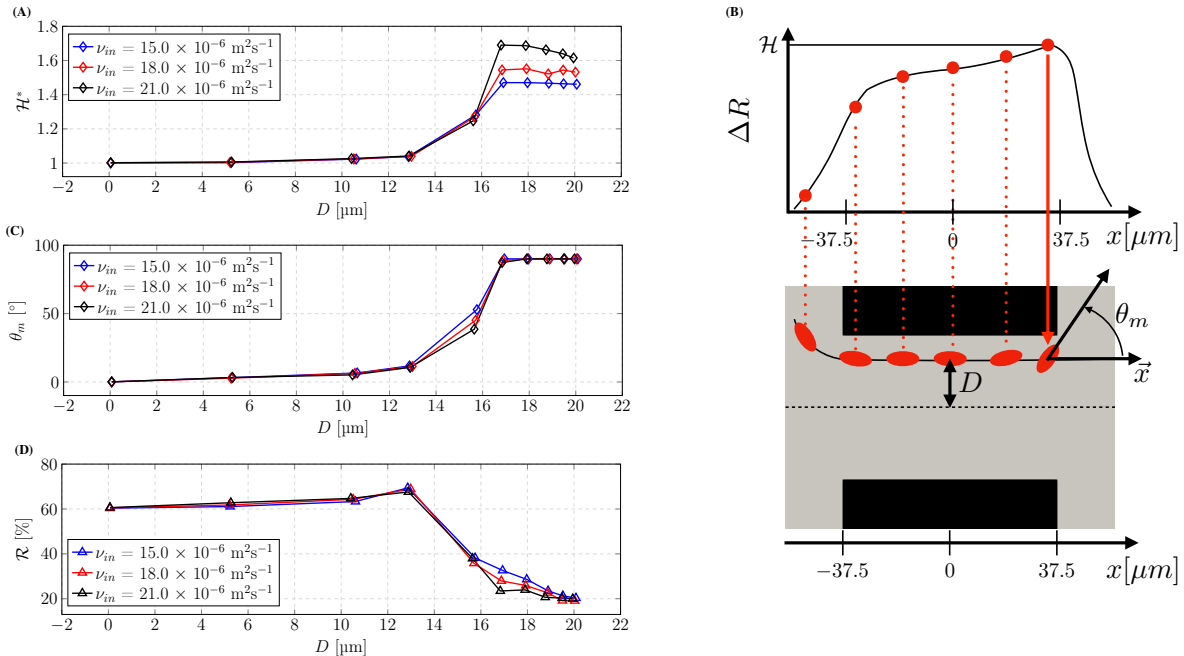


FIGURE 4 Evolution of the measured apparent volume, presented here with the non-dimensional form \mathcal{H}^* (A), the particle orientation θ_m when the pulse reaches its maximum value (C), and the \mathcal{R} metric (D), with the distance from orifice centerline D . Picture (B) illustrates how the distance from the aperture center D and the orientation θ_m are defined. In (A), (C) and (D), results are shown for different internal viscosities ν_{in} , the remaining RBCs parameters being set to: $G_s=2.5 \times 10^{-6} \text{ N m}^{-1}$; $E_a=2.5 \times 10^{-11} \text{ N m}^{-1}$; $E_b = 6.0 \times 10^{-19} \text{ J}$; $V_p=93 \text{ } \mu\text{m}^3$; $Q = 0.65$.

This is consistent with the numerical results, which show that pulses with low \mathcal{R} are associated with cell rotation and volume overestimation (see Fig. 4 A, C and D). The NN modelling provides a clustering globally equivalent to \mathcal{R} . Indeed, for \mathcal{R} above 52 %, NN scores are close to 1, while below $\mathcal{R} = 52 \%$, values close to 0 are predicted by the NN.

Minor populations are also visible in Fig. 5. First, measurements at \mathcal{H} below 0.75 V are associated with platelets. Then, the widespread population around $\mathcal{H} = 3.5 \text{ V}$ results from doublets or bubbles in the aperture. Finally, measurements at $\mathcal{R} \approx 80 \%$ may be explained by pulses presenting a small peak so that \mathcal{W}_u does not measure the time spent in the peak. This could be induced by doublets or quasi-spherical RBCs, such as echinocytes. Note that the parameter p_u may be increased in order to handle pulses with a small peak as well. Otherwise, such minor populations could be discarded by applying relevant thresholds on \mathcal{R} or \mathcal{H} .

In the following, we investigate the effect of rejecting pulses associated with rotating RBCs from the statistical analysis. Two removal methods are assessed: $\mathcal{R} < 52 \%$ and NN score < 0.5 . By doing so, roughly 46 % of the pulses are discarded. The distributions of \mathcal{H} obtained with the classical counter and the hydrofocused counter are shown in Fig. 6A. Note that the signals are treated and amplified differently in

the two counters, which explains the offset between the distributions, and that the number of RBCs measured is different. Nonetheless, a first observation is that hydrodynamical focusing yields a symmetrical and Gaussian-like distribution while a right-skewed volume distribution is obtained with the classical system. However, rejecting pulses with \mathcal{R} below 52 % or with NN scores below 0.5 is shown to render symmetrical distributions (cases noHF- \mathcal{R} and noHF-NN in Fig. 6A).

In order to provide a quantitative comparison between both systems, measurements are calibrated before computing the probability density functions of cells volumes depicted in Fig. 6B. Voltage pulses are calibrated in such a way that the averaged \mathcal{H} for cases HF and noHF- \mathcal{R} (see Fig. 6A) correspond to the MCV rendered by a ABX Pentra™ DX120 (HORIBA Medical), used as a reference machine. As shown by Fig. 6B, a good comparison is observed between hydrofocused and classical systems, provided that a rotation filter is used (see noHF- \mathcal{R} and noHF-NN in Fig. 6B).

Electrical pulses from the classical and the hydrofocused devices were recorded for 20 blood samples in total. After calibration, Mean Cells Volumes (MCV) and RBCs Distribution Widths (RDW) were calculated. The MCV is the mean of V_p and the RDW is computed as the standard deviation of V_p , divided by the MCV. MCV and RDW are calculated from the

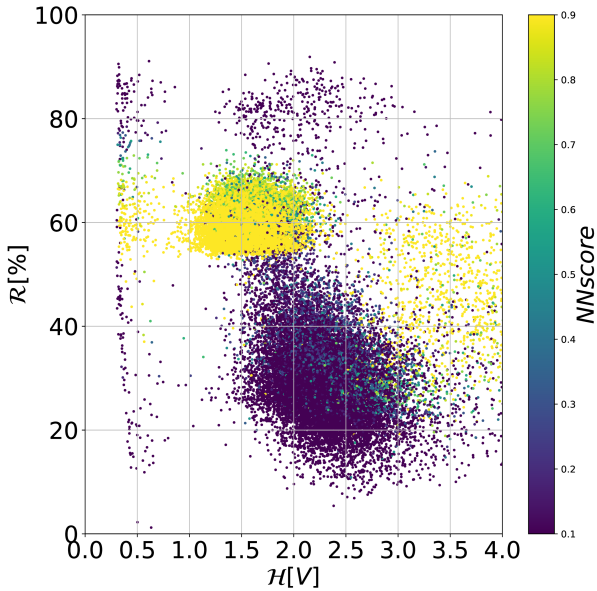


FIGURE 5 Scatter plot of the width ratios (\mathcal{R}) according to the pulses maximums (H) by colouring points according to the NN predictions (NN score).

amplitudes distribution, by discarding the part of the curve that falls below 5% of the maximum occurrence, which naturally excludes platelets and doublets associated with large volumes. The distribution curve is smoothed with a moving average algorithm prior the calculation of MCV and RDW. Correlations of MCV and RDW between the hydrofocused system and the classical system with different filters are shown in Fig. 6C and D, respectively. Regarding the MCV (Fig. 6C), good Pearson correlation coefficients (R) are obtained for all cases, even without sorting the pulses: when pulses are not sorted, the MCV is consistently overestimated by 20-25 % in our setup. Considering the RDW, results show that unfiltered data from the classical system are poorly correlated with hydrodynamical focusing (see case noHF in Fig. 6D, $R = 0.75$). The correlation is significantly improved to 0.88 and 0.89 with the use of the \mathcal{R} based filter and the NN modelling, respectively.

4 | DISCUSSION

For counting and sizing RBCs, impedance-based systems using the Coulter principle without hydrodynamical focusing remain widely used, because of their low fluid consumption, the simplicity of their implementation and their lower cost. The present study provides new insights on inaccuracies in

the sizing of RBCs in such analyzers. Using numerical simulations, we showed ‘bell-shaped’ pulses when cells flow in the aperture core. On the opposite, a variety of complex signatures are obtained when cells flow near the wall: there, velocity gradients make RBCs rotate, which generates a peak on the electrical pulse, leading in turn to overestimated RBCs volumes and right-skewed distributions. Furthermore, this volume overestimation is shown to depend on the RBCs rheology. Novel pulse editing methods are then proposed to reject pulses affected by near-wall effects. In the first approach, the rotation-associated peak is detected by computing and thresholding a metric (\mathcal{R}) calculated from the pulse. In the second, “bell-shaped” signatures of central trajectories are isolated using a neural network (NN) trained from numerical simulations. These methods retrieve symmetrical RBC volume distributions, in agreement with results from a system with hydrodynamic focusing, which can be considered as the reference. Besides, very good correlations are obtained with the reference in terms of mean corpuscular volume (MCV) and RBCs distribution width (RDW) with the two methods.

Further assessments of our approach are intended in the near future. Since both filters provide close results, we will focus on the \mathcal{R} -based one, which is easier to implement on the automata as compared with the NN. Actually, we are currently integrating the \mathcal{R} -based filter in the analyzer to apply it automatically and avoid time-consuming data transfer and application of the filter on the personal computer, as it was done in this proof of concept. In addition, we plan to apply this method to other blood samples with wider variations in MCV and RDW (for instance using pathological samples) to challenge the filter.

The proposed filters are meant to detect flow-induced rotations of particles. Note that in the case of more spherical particles, the problem of rotation-associated overestimation of the volume disappears. However, small electrical peaks^{36,4,14} may be present at the corner of the orifice. The parameters of the \mathcal{R} -based filter could be modified to detect such peaks, as discussed in Sec. 3.2.

Some blood analyzers use surfactants to modify the RBCs and make them spherical, so that the sizing is more robust. The use of a rotation-based filter may make the use of such complex and possibly costly external fluids useless.

Filtering the data using \mathcal{R} , which is a ratio of pulse widths, has advantages and drawbacks: first, such a filter has two relative threshold parameters that define the widths used to calculate \mathcal{R} . These parameters, p_u and p_d , need to be tuned and may depend on the external fluid used, the geometry, the flow speed and the type of particles to size. In general, when adapting the \mathcal{R} -based filter to a new machine, preliminary tests are needed to check and adapt p_u and p_d and the threshold value of \mathcal{R} to reject or accept the pulses. However, we have seen that the filter is not very sensitive to the values of those thresholds

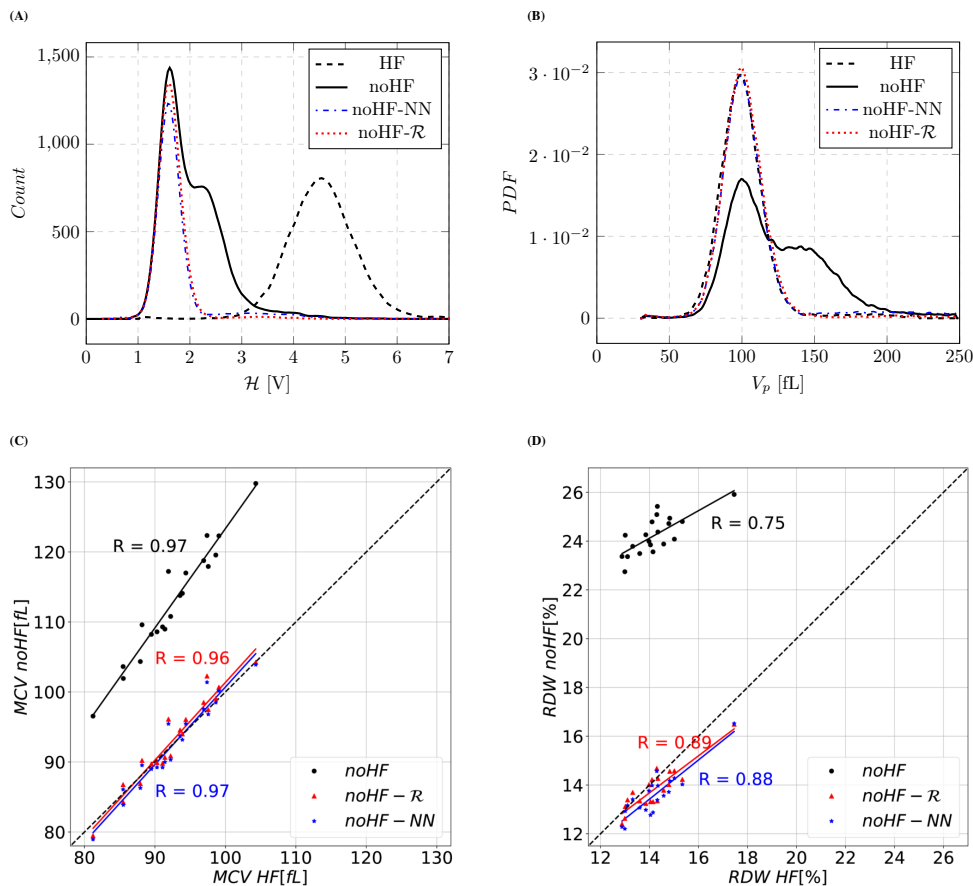


FIGURE 6 (A) Histograms of pulses maxima obtained from the classical system (noHF) and the system with hydrodynamic focusing (HF) when analysing a blood sample from a healthy donor. Case noHF- \mathcal{R} is obtained by applying $\mathcal{R} < 52\%$ as a rejection criterion, while noHF-NN results from the removal of pulses whose NN score is below 0.5. (B) Probability density functions of RBCs volumes, derived from pulses heights \mathcal{H} , after calibration. Correlations of Mean Cell Volumes (C) and RBCs Distribution Widths (D), between the classical system and the hydrofocused analyzer, with and without filters.

(as illustrated in the supplementary materials). In addition, the tuning process is easy, as the \mathcal{R} -based filter should separate the pulses into two well-defined subpopulations, as shown in Fig. 5. Such a property of the \mathcal{R} -based filter is actually a major advantage compared to other pulse editing methods^{8,19,21,22}. One reason for this is that \mathcal{R} is actually independent of the RBC volume, which is not the case of many other metrics used for pulse editing, such as the pulse width⁸. This also guarantees a better robustness of the filter. In addition, note that some filters based on the symmetry of the pulse would not reject some near-wall symmetric pulses, while the \mathcal{R} -based filter does.

While many analyzers implement pulse editing methods, they are rarely documented in a thorough way, and a stricter quality assessment would be beneficial. In particular, we stress that it is possible to have a true validation of pulse editing methods by comparing the pulse edited data with data obtained in systems with hydrodynamic focusing for the same blood, as

presented in this paper, and by comparing not only the MCV but also the RDW, which is more sensitive to errors and is more and more employed as a diagnostic parameter³.

Finally, cells orientation in flow cytometers has been reported to alter the optical measurements^{15,16}. Numerical simulations of cells dynamics in such devices could bring new understanding on such artefacts and potentially assist the development of more robust flow cytometers.

References

1. Coulter WH. Means for counting particles suspended in a fluid. US Patent 2,656,508. 1953.
2. Evans TC, Jehle D. The red blood cell distribution width. *The Journal of Emergency Medicine* 1991; 9: 71–74. doi: 10.1016/0736-4679(91)90592-4

3. Salvagno GL, Sanchis-Gomar F, Picanza A, Lippi G. Red blood cell distribution width: A simple parameter with multiple clinical applications. *Critical Reviews in Clinical Laboratory Sciences* 2014; 52(2): 86–105. doi: 10.3109/10408363.2014.992064
4. Kachel V. Sizing of Cells by the Electrical Resistance Pulse Technique: Methodology and Application in Cytometric Systems. *Cell Analysis* 1982; 1: 195–331.
5. Mohandas N, Clark MR, Kissinger S, Bayer C, Shohet SB. Inaccuracies associated with the automated measurement of mean cell hemoglobin concentration in dehydrated cells. *Blood* 1981; 56(1).
6. Strauchen JA, Alston W, Anderson J, Gustafson Z, Fajardo LF. Inaccuracy in automated measurement of hematocrit and corpuscular indices in the presence of severe hyperglycemia. *Blood* 1981; 57(6).
7. Brian SB. On the Distribution of Red Cell Volumes. *Blood* 1968; 31: 503–515.
8. Waterman CS, Atkinson EEJ, Wilkins BJ, Fischer CL, Kimzey SL. Improved Measurement of Erythrocyte Volume Distribution by Aperture-Counter Signal Analysis. *Clinical Chemistry* 1975; 21(9): 1201-1211.
9. Price-Jones C. The diameters of red cells in pernicious anæmia and in anæmia following hæmorrhage. *The Journal of Pathology and Bacteriology* 1922; 25(4): 487–504. doi: 10.1002/path.1700250410
10. Grover NB, Naaman J, Ben-Sasson S, Doljanski F. Electrical Sizing of Particles in Suspensions: III. Rigid Spheroids and Red Blood Cells. *Biophysical Journal* 1972; 12: 1099-1116.
11. Kachel V. Electrical Resistance Pulse Sizing: Coulter sizing. In: No. 4. Wiley-Liss, Inc.; 1990: 45-80.
12. Isèbe D, Nérin P. Numerical simulation of particle dynamics in an orifice-electrode system. Application to counting and sizing by impedance measurement. *International Journal for Numerical Methods in Biomedical Engineering* 2013; 29(4): 462–475.
13. Golibersuch D. Observation of Aspherical Particle Rotation in Poiseuille Flow via the Resistance Pulse Technique: I. Application to Human Erythrocytes. *Biophysical Journal* 1973; 13(3): 265 - 280.
14. Taraconat P, Gineys JP, Isebe D, Nicoud F, Mendez S. Numerical simulation of deformable particles in a Coulter counter. *International Journal for Numerical Methods in Biomedical Engineering* 2019. doi: 10.1002/cnm.3243
15. Loken MR, Parks DR, Herzenberg LA. Identification of cell asymmetry and orientation by light scattering.. *Journal of Histochemistry & Cytochemistry* 1977; 25(7): 790-795. doi: 10.1177/25.7.330730
16. Stovel R, Sweet R, Herzenberg L. A means for orienting flat cells in flow systems. *Biophysical Journal* 1978; 23(1): 1 - 5. doi: [https://doi.org/10.1016/S0006-3495\(78\)85427-7](https://doi.org/10.1016/S0006-3495(78)85427-7)
17. Spielman LGSL. Improving Resolution in Coulter Counting by Hydrodynamic Focusing. *Journal of Colloid and Interface Science* 1968; 26: 175–182.
18. Tatsumi N, Tsuda I, Fujimoto K, Matsumoto H. Evaluation of a new blood cell counter with sheath flow system. *Cytometry: The Journal of the International Society for Analytical Cytology* 1985; 6(5): 395–400.
19. Dunstan HJ, Harfield JG, Knight P, Podgorney HJ. Method and apparatus for editing particle produced electrical pulses.US Patent 4,797,624. 1989.
20. Lombarts AJPF, Koevoet AL, Leijnse B. Basic Principles and Problems of Haemocytometry. *Annals of Clinical Biochemistry* 1986; 23(4): 390-404. doi: 10.1177/000456328602300404
21. Ishizuka K, Matsubayashi H, Nishimori M. Particle counting method and particle counting apparatus.WO Patent 2015099116 A1. 2015.
22. Zhang S, Zheng M, Lin D, Huo Z. Systems and methods for particle counting.US Patent 7,457,709 B2. 2008.
23. Donnenberg AD, Donnenberg VS. Rare-Event Analysis in Flow Cytometry. *Clinics in Laboratory Medicine* 2007; 27(3): 627 - 652. doi: <https://doi.org/10.1016/j.cll.2007.05.013>
24. Lippi G, Pavesi F, Bardi M, Pipitone S. Lack of harmonization of red blood cell distribution width (RDW). Evaluation of four hematological analyzers. *Clinical Biochemistry* 2014; 47(12): 1100–1103. doi: 10.1016/j.clinbiochem.2014.06.003
25. Behrens WV, Edmondson S. Comparison of techniques improving the resolution of standard Coulter cell sizing systems. *Journal of Histochemistry & Cytochemistry* 1976; 24(1): 247-256. doi: 10.1177/24.1.1254920
26. Taraconat P, Isebe D, Mendez S. Device for medical analyses with impedance signal processing.WO Patent 2020216952. 2020.

27. Lanotte L, Mauer J, Mendez S, et al. Red cells' dynamic morphologies govern blood shear thinning under micro-circulatory flow conditions. *Proceedings of the National Academy of Sciences USA* 2016; 113(47): 13289–13294.
28. Sigüenza J, Mendez S, Nicoud F. How should the optical tweezers experiment be used to characterize the red blood cell membrane mechanics?. *Biomechanics and Modeling in Mechanobiology* 2017; 16: 1645-1657.
29. Mendez S, Abkarian M. In-plane elasticity controls the full dynamics of red blood cells in shear flow. *Physical Review Fluids* 2018; 3(101101(R)).
30. Evans EA, Fung YC. Improved Measurements of the Erythrocyte Geometry. *Microvascular Research* 1972; 4: 335–347.
31. Skalak R, Tozeren A, Zarda RP, Chien S. Strain energy function of red blood cell membranes. *Biophysical Journal* 1973; 13: 245–264.
32. Helfrich W. Elastic Properties of Lipid Bilayers: Theory and Possible Experiments. *Z. Naturforsch* 1973; 28 c: 693–703.
33. Abkarian M, Viallat A. Fluid-Structure Interactions in Low-Reynolds-Number Flows. *On the importance of red blood cells deformability in blood flow. Royal Society of Chemistry, London* 2016: 347-462.
34. Mendez S, Abkarian M. *Dynamics of Blood Cell Suspensions in Microflows*.ch. Single Red Blood Cell Dynamics in Shear Flow and their Role in Hemorheology; CRC Press . 2019.
35. Abadi M, Agarwal A, Barham P, et al. TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems. *arXiv preprint arXiv:1603.04467*, 2016.
36. Grover NB, Naaman J, Ben-Sasson S, Doljanski F. Electrical Sizing of Particles in Suspensions: I. Theory. *Biophysical Journal* 1969; 9: 1398-1414.
37. Hurley J. Sizing Particles with a Coulter Counter. *Biophysical Journal* 1970; 10: 74–79.
38. Gibaud E. *Numerical simulation of red blood cells flowing in a blood analyzer*. PhD thesis. Université de Montpellier, 2015.
39. Peskin CS. The immersed boundary method. *Acta Numerica* 2002; 11: 479–517.
40. Iss C, Midou D, Moreau A, et al. Self-organization of red blood cell suspensions under confined 2D flows. *Soft Matter* 2019; 15: 2971-2980.

SUPPLEMENTARY MATERIALS

Numerical pipeline

This section provides more details on the numerical setup used in this work, which is the main topic of a previous publication¹⁴. Simulating the RBC behavior in the counting chamber of the Yumizen™ H2500 (HORIBA Medical, Montpellier, France) is challenging because of the large separation of scales: far from the sensing aperture, RBCs are transported at low velocities over long distances, while they are detected inside the micro-orifice of a few tens of micrometers, in which its velocity is of the order of 5 m s^{-1} . To overcome this separation of scales issue, a specific numerical pipeline has previously been developed and validated¹⁴. Fig. S1 illustrates the entire procedure, which consists in a specific series of numerical computations. First, the flow field is computed in the industrial geometry, without cells. The fluid domain corresponds to the Yumizen H2500 automaton, more precisely the part dedicated to the counting and sizing of RBCs. The aperture for particles detection has a diameter of $50 \mu\text{m}$ and a length of $75 \mu\text{m}$. In agreement with Yumizen H2500 operating regime, a pressure drop of 200 mbar is imposed. The electrolyte being mostly water (incompressible fluid with a kinematic viscosity ν_{ext} that equals $10^{-6} \text{ m}^2\text{s}^{-1}$), such conditions leads to a Reynolds number of the order of 200, in the aperture¹⁴. Furthermore, the flow becomes stationary and axisymmetric after $50 \mu\text{s}$, a zero velocity being imposed as initial condition. As shown by Fig. S1A, the aperture features high velocities, due to the constriction. For a given trajectory of RBCs, the corresponding streamline is extracted from the flow in the industrial geometry (See Fig. S1A) and the strain rate along this streamline is computed ($\frac{\partial u_s}{\partial s}$, with s the curvilinear abscissa), and expressed as a function of time to assess the strain rate seen by the RBC during its motion. The time evolution of the strain rate is used as an income of a simulation focusing on the dynamics of the RBC before the aperture, during which RBCs are oriented and stretched (see Fig. S1 B). This second simulation aims at elongating the RBC in an extensional flow in a way similar as if it had followed the streamline, before entering the aperture. The cell stretching is performed from time T_A to T_B corresponding to points A and B, belonging to the selected streamline (see Fig. S1A). Once elongated the RBC is dropped at point B in a reduced configuration of the industrial geometry (See Fig. S1C) and the RBC behavior in the aperture is solved. This reduced domain is restricted to a close zone around the aperture in order to reduce the computational costs. Applying suitable boundary and initial conditions, the flow in the reduced configuration is equivalent to that of the full industrial geometry (for more details please refer to the setup used in¹⁴). Separately, the electrical perturbation induced by

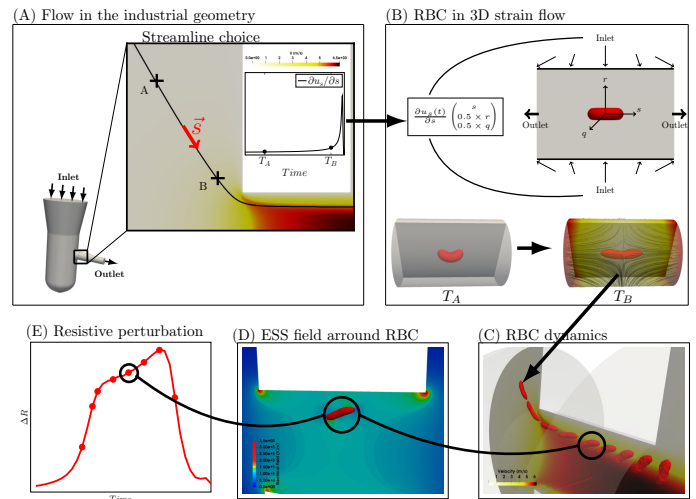


FIGURE S1 Numerical pipeline for calculating the RBC dynamics and the associated electrical perturbation in an industrial automaton counter on the Coulter principle. The pipeline is described in detail in by Taraconat *et al.*¹⁴

the presence of the cell in the aperture is computed by solving the electrostatics equation, which is a Laplace equation. Given a membrane shape coming from the simulation of the RBC dynamics (Fig. S1C), a variable conductivity coefficient σ is imposed when solving the Laplace equation. In order to reproduce the non conducting behavior of the particle stated in the analytical development^{37,36}, an infinitely small conductivity is imposed inside the membrane³⁸. Solving the Laplace equation thus leads to the electrical field disturbed by the cell (See Fig. S1D). The resistance of the system is deduced from the electrical field and compared with the resistance of the fluid domain without RBC (result of a computation with a constant conductivity in the domain). This provides ΔR (see Fig. S1E), the resistance variation induced by the cell in the electrical field. When performing the simulation of the particle dynamics inside the micro-orifice (Fig. S1C), RBC consecutive positions are stored (typically every microsecond). Repeating the electrostatic simulation described above for each membrane location allows the reconstruction of the impedance pulse (See Fig. S1E).

All simulations of the pipeline are performed with the in-house YALES2BIO software (<https://imag.umontpellier.fr/~yales2bio/>). The suspending fluid and the cytosol are considered as Newtonian and incompressible fluids. Hence, the flow field is calculated by solving Navier-Stokes equations:

$$\frac{\partial \vec{u}}{\partial t} + \vec{u} \cdot \nabla \vec{u} = -\frac{\nabla p}{\rho} + \nabla \cdot [\nu(\nabla \vec{u} + (\nabla \vec{u})^T)] + \vec{f} \quad (3)$$

$$\nabla \cdot \vec{u} = 0 \quad (4)$$

with ν denoting the kinematic viscosity, \vec{u} the fluid velocity, p the pressure and ρ the density. The source term \vec{f} accounts for the forces applied by the RBC membrane on the surrounding fluid. The computation of the flow in the industrial geometry is performed without RBCs (see Fig. S1A). Hence, \vec{f} equals $\vec{0}$ in that simulation.

Calculations of the RBC extension and RBC dynamics (illustrated in Fig. S1B and C) involve the resolution of a Fluid-Structure Interaction (FSI) problem, between the RBC solid membrane and fluids inside and outside the membrane. In that case, \vec{f} is not zero. Briefly, in YALES2BIO, the RBC membrane is assumed infinitely thin and elastic. The membrane resists to bending, shear and area deformations. This is modelled by the Skalak law³¹:

$$W_{sk} = \frac{G_s}{4} [(\lambda_1^2 + \lambda_2^2 - 2)^2 + 2(\lambda_1^2 + \lambda_2^2 - \lambda_1^2 \lambda_2^2 - 1)] + \frac{E_a}{4} (\lambda_1^2 \lambda_2^2 - 1)^2, \quad (5)$$

which account for area and shear resistance, and with the Helfrich bending energy³²:

$$\epsilon_b = \frac{E_b}{2} \int (2H - c_0)^2 dS, \quad (6)$$

which models the membrane resistance to bending deformations. G_s and E_a in Eq. 5 denote the shear modulus and the area modulus, respectively, while λ_1 and λ_2 are the principal values of strain in the membrane plane. In Eq. 6, the terms E_b , H and c_0 refer to the bending modulus, the mean membrane curvature and the spontaneous curvature (set to 0 in our simulations), respectively. The membrane is displaced and deformed by the flow field (displacement are performed according to the velocity field \vec{u} , see Eq. 3) and consequently stores elastic energy as prescribed by Eqs. 5 and 6. In return, the solid membrane applies an elastic force to the fluid, which is accounted in the source term \vec{f} . The coupling between the membrane and the fluid is allowed by the Immersed boundary Method³⁹. This type of methods is classical and the present solver has been used and validated in many different situations^{27,28,29,40}.

Finally, calculations of the electrical perturbation induced by the RBC (see Fig. S1D) are performed by solving a Laplace equation for the electrical potential ϕ , with a variable conductivity coefficient σ :

$$\nabla \cdot [\sigma(x) \vec{\nabla} \phi] = 0 \quad (7)$$

As discussed previously, the RBC is assumed insulating. Hence, σ is set to zero inside the membrane for these simulations.

The geometries and meshes used in the calculations are the same that were used in¹⁴.

Experimental acquisitions

Electrical pulses from the YumizenTM H2500 (HORIBA Medical, Montpellier, France) are recorded. The embedded treatment of this instrument aims at registering pulses maxima and pulses durations (commonly called heights and widths, respectively). This is done by defining an acquisition threshold above which a peak detector is activated, in order to assess the pulse maximum. The pulse duration is calculated as the time spent by the electrical perturbation above the predefined threshold.

In order to apply the filters presented in this work, the entire pulse must be stored, which is not the case in the Yumizen H2500 embedded system. Registering the waveforms is required because pulses widths involved in the calculation of \mathcal{R} (see Eq. 2) are relative to the pulse maximum and the NN takes the entire resampled pulse (see Fig. S4 and S5) as an input. A LabVIEWTM (National Instruments) code monitors the physical signal by keeping in the cache memory the most recent signal values. Parts of the signal higher than a predefined acquisition threshold are written in an acquisition file for future post-processing. The recording of small signal parts occurring before and after the signal is higher than the acquisition threshold is also allowed by the LabVIEW code. This is parametrized by the acquisition offset illustrated in Fig. S2. The signal is made processable by the computer on which the LabVIEW code is executed by a PicoScopeTM 4262 (Pico Technology) which is connected to the Yumizen H2500 during the experiments. Pulses are recorded at a constant time step Δt of 0.5 μs .

Signals from two different units of the automaton are recorded when analysing a blood sample. The first unit is dedicated to RBCs counting and sizing with the Coulter principle, while the second is normally used for detecting platelets among RBCs. The later device combines optical with impedance measurements for the differentiation of platelets. Making cells crossing the beam for the optical assessment is done by hydrodynamical focusing, which involves a sheath flow for centering cells. It also ensures centered paths of cells in the aperture of impedance measurements. It should be noted that the optical measurements of the second unit are not considered in this works and only impedance pulses are recorded with the LabVIEW code.

Comparison of numerical and experimental pulses

The understanding of the pulse shapes in an industrial counter based on the Coulter principle is the foundation stone of the two pulse editing methods proposed in this paper. Both methods are based on the results of numerical simulations. In this supplementary material, we show the relevance of the numerical pulses with respect with experimental measurements in the

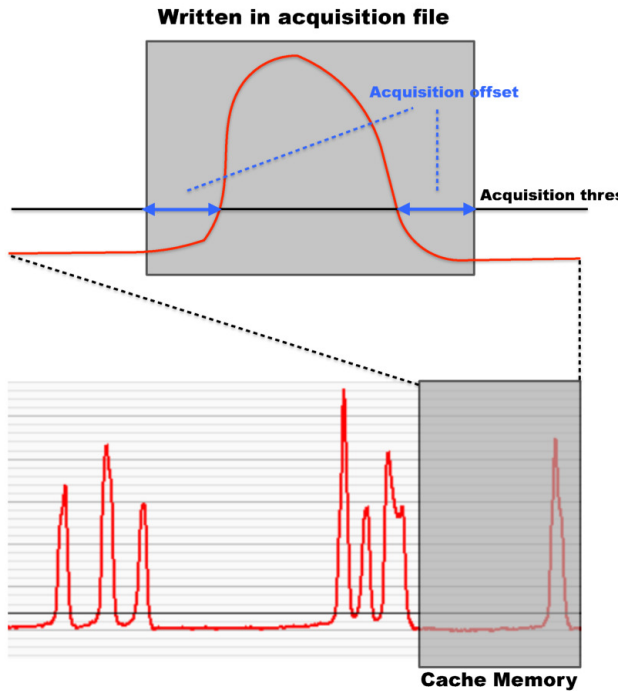


FIGURE S2 Principle of the LabVIEW code for registering voltage pulses from a blood analysis.

Yumizen H2500 automaton (see Sec. 2.1 and 2.3). The comparison is depicted in Fig. S3. More precisely, numerical pulses of Fig. S3A, B, C and D correspond to trajectories 1, 6, 7 and 10 of Fig. 3 (see main document). Note that pulses are superimposed in graphs of Fig. S3 so that their maximum coincides with time $t = 0 \mu s$. Experimental pulses (ΔU) are divided by the averaged maximum of bell-shaped signatures (viz. waveforms illustrated in Fig. S3A). This leads to ΔU^* and allows a direct comparison between resistive pulses from the simulations, with voltage pulses of the experiments. Comparisons presented in Fig. S3 show that electrical signatures predicted numerically are retrieved in the experiments. This validates the assumptions made in Sec. 3.1, which are based on numerical observations. Note that not only RBC signatures are well reproduced, but that simulations accurately predict pulses associated with other shapes, such as spheres, as shown by Taraconat *et al.*¹⁴

Concerning pulses of Fig. S3A, \mathcal{R} values around 65% are calculated, while the NN predicts scores close to 1. Regarding pulses of Fig. S3B, C and D, the NN predicts scores close to 0.0 and \mathcal{R} values are below 35%. This supports the capability of both the \mathcal{R} -based filter and the NN modelling (see Sec. 2.2) to reject rotation associated signals which present a secondary peak.

Note that experiments also feature other pulses potentially associated with platelets, white blood cells, doublets, bubbles or unusual shapes of RBCs. These cases are not handled in

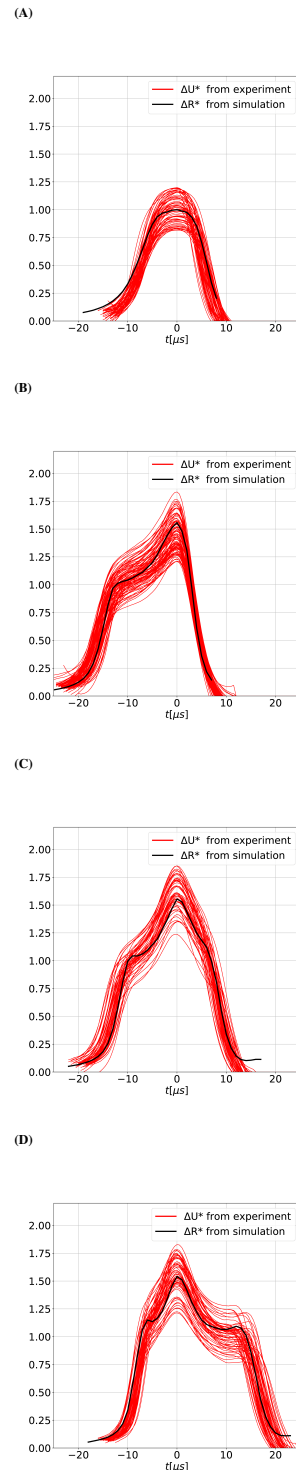


FIGURE S3 Comparison of numerical pulses (ΔR^*) and experimental results (ΔU^*). Experimental signatures are recorded when analysing a blood sample with the YumizenTM H2500 (HORIBA Medical, Montpellier, France). These experimental pulses are recorded from the unit dedicated to RBCs analysis of the Yumizen H2500, which uses the Coulter principle without hydrodynamical focusing.

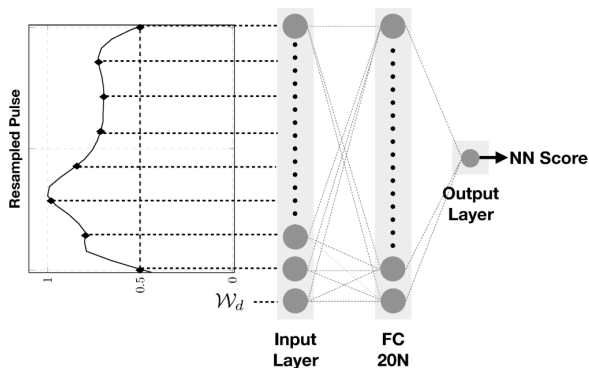


FIGURE S4 Neural network architecture used for detecting the particle rotation from the electrical pulse. The hidden layer is Fully-Connected (FC) to input and output layers and contains 20 Neurons (20N). The input layer contains 21 neurons and handles the resampled pulse, in addition to \mathcal{W}_d , expressed in μs . The output layer, composed of a sole neuron, renders the NN score included in $[0 ; 1]$, since all activation functions are sigmoids.

the simulations. Nevertheless, the majority of the experimental waveforms can be associated with a simulated pulse.

Neural network filtering

In this study, a Neural Network (NN) is shown to classify pulses that are impaired by cells rotations in the sensing region, from those associated to RBCs aligned with the aperture axis and relevant for volume measurements. The NN architecture is illustrated in Fig. S4. It is composed of three layers: (1) the input layer which handles the pulse values in addition to a calculated pulse width \mathcal{W}_d ; (2) the hidden layer composed of 20 neurons; (3) the output layer made of a single neuron which render a score between 0 and 1. A threshold of 0.5 on NN scores is used to isolate pulses suited for the volume measurement (NN score >0.5).

Pulse durations may vary according to several factors, such as the particle volume or the trajectory in the aperture. Hence, being recorded at a constant time step Δt , pulses may be re-presented with different numbers of variables. However, a NN must take a given number of variables as an input. In this respect, pulses are resampled in order to make them all processable by the NN. More precisely, after scaling the pulse by its maximum \mathcal{H} , the part of the curve located above 0.5 is resampled regularly with 20 values. The input layer (see Fig. S4) is made of 21 neurons, since it takes the 20 resampled values in addition to \mathcal{W}_d .

The NN of Fig. S4 is trained from a numerical database composed of 78 pulses. These pulses are associated to target scores

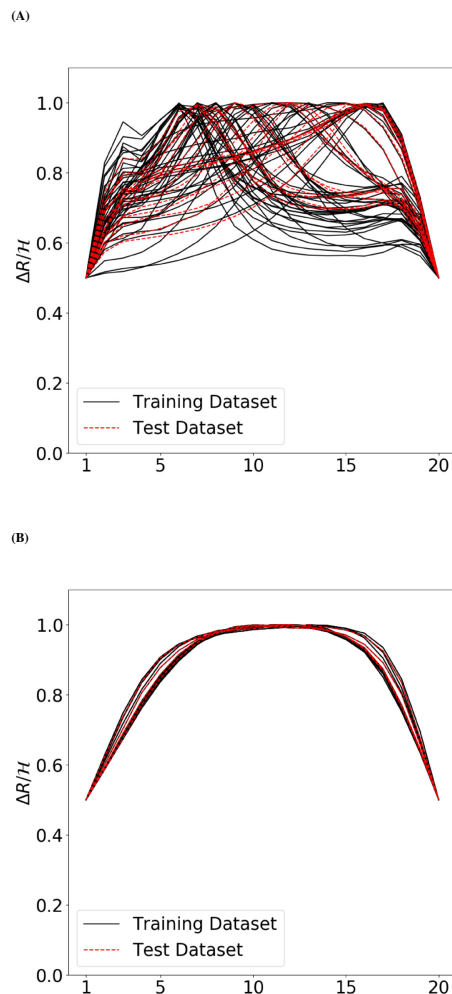


FIGURE S5 Numerical pulses resampled with 20 values and used for the NN training. (A) Waveforms associated to target values of 0 (viz. the RBC has turned in aperture). (B) Pulses for which RBCs has not turned in the micro-orifice and thus associated to target values of 1. The numerical database is split in two parts: training dataset and testing dataset.

of 0 or 1 whether the cell has experienced a rotation in the aperture or not. Figure S5A and B depict resampled pulses of the numerical database with target scores of 0 and 1, respectively. From the simulated RBC dynamics in the aperture the particle orientation (θ_m) at the pulse maximum is assessed. The corresponding pulse is then affiliated to 0 or 1 depending on if θ_m is higher or lower than 15° . The class of pulses linked to non-rotating cells is under-represented, since it represents only 20% of the numerical database (see Fig. S5B). This was a voluntary choice when building the numerical database because this specific class was found to be robust with respect to RBC parameters. Still, good results are provided by the NN on the numerical testing dataset, after the learning step performed

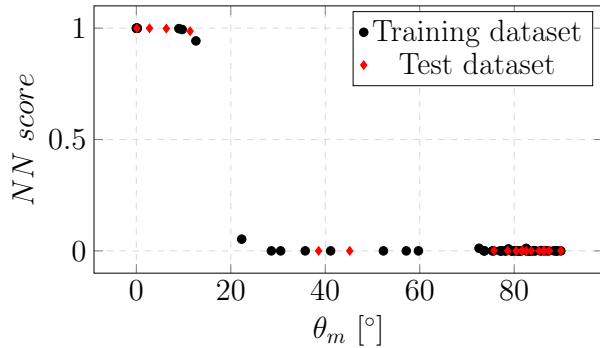


FIGURE S6 Prediction of the neural network after the learning process (NN scores) as a function of θ_m , when considering the numerical database.

with the training dataset. Figure S6 shows NN scores according to θ_m for numerical data of Fig. S5, once the model is trained. NN scores are close to 1 when θ_m is below 15° and around 0 otherwise, for both the training and the test datasets. Tests have been performed when changing test and training datasets illustrated in Fig. S5 and the results remain similar. Furthermore, the NN modelling is shown to provide relevant results when applied on experimental data, as discussed in the main text.

Effects of p_u and p_d on the \mathcal{R} based filter

One of our pulse-editing methods consisting in thresholding a metric calculated from the voltage pulses is presented in Sec. 2.2. This metric (\mathcal{R}) is parametrized by p_u and p_d , which define the thresholds for the widths computation (see Fig. 2). \mathcal{R} is defined as the ratio of two pulse widths.

Values of p_u and p_d are set to $7/8$ and $1/2$ in this work for computing \mathcal{R} . The filtering is done by rejecting voltage pulses whose \mathcal{R} is below 52% and about 46% of pulses are removed from the analysis this way. In Fig. S7, the dependency of the filter according to parameters p_u and p_d is illustrated. In particular, the graph of \mathcal{R} according to \mathcal{H} is shown for different couples p_u/p_d . Overall, the general trends are identical for all cases of Fig. S7: two clusters separated along \mathcal{R} axis; the cluster of higher \mathcal{R} values spreads along \mathcal{H} axis between 1.0 V and 2.25V, while for the cluster of low \mathcal{R} values, \mathcal{H} is broadly included in [1.6V ; 3.0V]. Then, it is observed that \mathcal{R} is more sensitive to p_u than to p_d . On the one hand, increasing and decreasing p_u by 5% change the location of the two clusters separation to 40% and 60%, respectively, as Fig. S7A and B show. On the other hand, applying the same variations on p_d broadly conserves the clusters' arrangement with a separation around 50% for both cases of Fig. S7 C and D.

Of course, the relevant choice of the threshold for \mathcal{R} depends on p_u and p_d , but it is easy to make, as clusters are clearly distinct. Adapting the threshold for each case, we obtain that

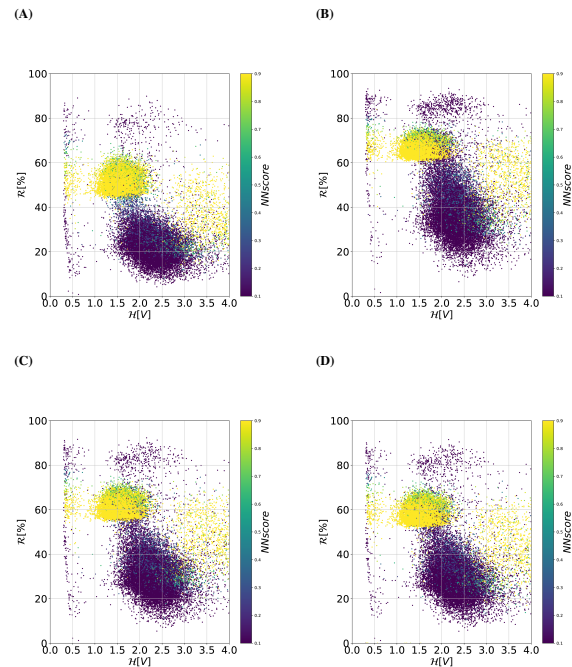


FIGURE S7 Impact of p_u and p_d on the \mathcal{R} metric. Couples p_u/p_d used for the different graphs are as follows: (A) 0.92/0.5 ; (B) 0.83/0.5 ; (C) 0.875/0.525 ; (D) 0.875/0.475.

rejecting the cluster of low \mathcal{R} performs almost the same filtering whatever p_u and p_d , since almost 46% of the signatures are discarded in each configuration. This is also supported by results of the NN (represented in color in Fig. S7) which depict that rotation associated pulses (with NN scores ≈ 0.0) are in the lower cluster, for all couples p_u/p_d .

