Methods and Applications in Fluorescence

PAPER

RECEIVED
6 July 2016

REVISED
23 May 2017

ACCEPTED FOR PUBLICATION
13 June 2017

PUBLISHED
29 August 2017

Improved signal recovery for flow cytometry based on 'spatially modulated emission'

S Quint[1,5], J Wittek[2,3], P Spang[2], N Levanon[4], T Walther[3] and M Baßler[2]

[1] Department of Experimental Physics, Saarland University, Campus E2.6, 66123 Saarbrücken, Germany
[2] Fraunhofer ICT-IMM, Carl-Zeiss-Str. 18-20, 55129 Mainz, Germany
[3] Institute for Applied Physics, TU Darmstadt, Schlossgartenstr. 7, 64289 Darmstadt, Germany
[4] Department of Electrical Engineering-Systems, Tel Aviv University, Ramat Aviv 69978, Israel
[5] Author to whom any correspondence should be addressed.

E-mail: Stephan.Quint@physik.uni-saarland.de **and** Michael.Bassler@imm.fraunhofer.de

Keywords: flow cytometry, spatially modulated emission, digital filtering, pulse compression, SNR gain, RADAR, LABS

Abstract

Recently, the technique of 'spatially modulated emission' has been introduced (Baßler *et al* 2008 *US Patent* 0080181827A1; Kiesel *et al* 2009 *Appl. Phys. Lett.* **94** 041107; Kiesel *et al* 2011 *Cytometry* A **79A** 317–24) improving the signal-to-noise ratio (SNR) for detecting bio-particles in the field of flow cytometry. Based on this concept, we developed two advanced signal processing methods which further enhance the SNR and selectivity for cell detection. The improvements are achieved by adapting digital filtering methods from RADAR technology and mainly address inherent offset elimination, increased signal dynamics and moreover reduction of erroneous detections due to processing artifacts. We present a comprehensive theory on SNR gain and provide experimental results of our concepts.

1. Introduction

Modern flow cytometers [1] are widely used for detecting biological cells or particles in fluidic streams. Using different light sources, bulky optics and a number of detectors, both scattering and fluorescent light is observed when objects pass the excitation and detection zone of the flow cytometer. Specific cell-tagging, e.g. by attaching fluorophore-conjugated antibodies to antigens at the cell surface, is often used to distinguish between different cell populations. A narrow well-adjusted excitation spot providing a high photon flux density avoids coincidences between different objects and ensures a sufficient and detectable amount of fluorescent light during a short time interval. Due to their size and susceptibility, commercial state-of-the-art devices are restricted to laboratory use. For this reason, the concept of 'spatially modulated emission' has recently been introduced which promises a more robust operation and ease of use.

Instead of a narrow excitation spot addressing single traversing objects, the detection zone is enlarged, resulting in a less-intense but time-distributed fluorescence (figure 1). The key element is a non-periodic slit mask placed in between the fluidic cell and the optical detector. This yields a time-modulated fluorescence signal of passing objects, where the signal pattern in time corresponds to the sequence of transparent slits (fluorescence is transmitted) and opaque bars (fluorescence is blocked). Putting the sequence in a binary context, transparent slits are denoted with 1 and opaque bars are denoted with 0. Binary sequences can either be randomly chosen or derived from so called 'low auto-correlation binary sequences' (LABS) which have their origin in signal information theory [2–9] and are widely used e.g. for RADAR applications. In particular, LABS serve to increase the signal-to-noise ratio (SNR) of RADAR signals by transmitting time-extended and time-modulated pulses which offer the same delay resolution but higher energy compared to unmodulated short pulses. In conjunction with suitable digital filters, i.e. correlation sequences, received pulse trains can then be compressed in time (pulse compression) to recover the necessary time resolution.

Digital filtering in the field of RADAR technology is mainly based on correlation techniques which are here exploited for signal processing. We remark that fluorescence signals from the here discussed measurements differ significantly from RADAR signals in three aspects: (1) cell velocity is not uniform in microchannels resulting in significant variations of up ±20 % in signal duration. (2) Auto-fluorescence and scattered

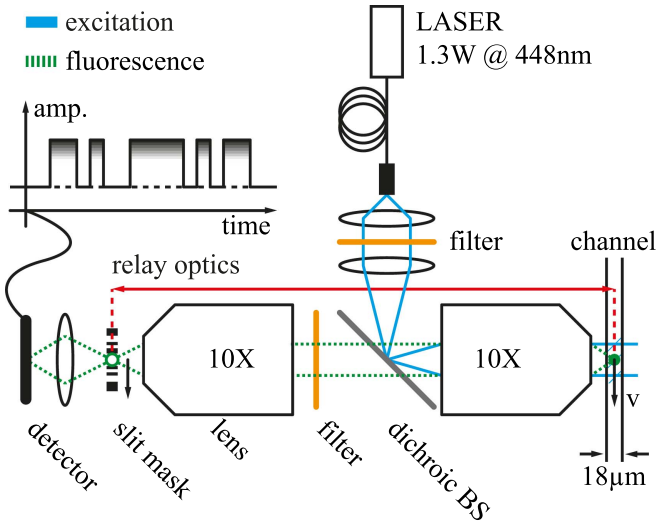

Figure 1. Flow cytometry setup implementing spatially modulated emission. Similar to an epi-fluorescence microscope, the setup consists of a relay optics that images the fluorescent light of excited cells in the channel on a slit mask. Depending on the cell velocity, the detector receives an amplitude-modulated (unipolar) signal which corresponds to the sequence of slits and bars of the mask. Excitation as well as fluorescence detection are achieved by the same lens (right). Separation of both light paths is realized by a dichroic beam splitter (BS). Additional optical filters support the dichroic beam splitter to suppress stray light from excitation. Moreover, focusing lenses serve to ensure a collimated light beam of the excitation light within the detection zone.

excitation light inherently result in an additive signal offset. (3) The transparent and opaque features of the modulation mask cause a unipolar (on/off-type) signal in contrast to the bipolar RADAR signals.

Processing methods from RADAR technology can therefore only partially be applied to our method. Parallels are e.g. found in the theoretical description of LABS and filters, especially for SNR gain prediction. In contrast, peak finding methods cannot be easily adapted. The key difference between both techniques is the method to determine the velocity of objects. Whereas RADAR applications exploit the Doppler effect (frequency shift) of the RF carrier, the here discussed principle takes advantage of the temporal dilation or compression of fluorescence signals. This results in correlation signals with fundamentally different signal properties. We discuss various designs of digital filters adapted to the specifics of the fluorescence signals. However, the discussion of the complete signal processing chain is beyond the scope of this paper and we refer to pertinent literature in the field of RADAR [2].

2. Signal theory

It may be noted that we make use of the standard definitions from related literature and textbooks [10]. E.g., the typical measure of the signal energy is the integral (sum) over the squared signal amplitude. From a physical point of view, squaring the amplitude of a voltage signal yields the unit V^2 expressing the power drop of a signal flowing through a hypothetical 1 Ω resistor. Summing this term over a certain time period therefore provides information on the energy of a time-limited signal. This is in analogy to continuous systems where the integral over a certain period of time is calculated. Giving a better impression on these abstract definitions, we explicitly added units to our derivations in brackets. However, since signals do not pass resistors and since integration is only performed by discrete means, the absolute value of signal energy is measured in units of V^2. This is a reasonable approach thoroughly found in textbooks since the numerical value of the signal energy is proportional to the squared signal amplitude in Volts.

In signal information theory, LABS are denoted in a bipolar fashion [2] with −1 and 1 and can be expressed as

$$B = (b_0 \in \{-1, 1\}, b_1, \cdots b_{N-1}) \quad [\mathrm{V}]. \quad (1)$$

Let H be a second sequence of length M with

$$H = (h_0 \in \mathbb{R}, h_1 \cdots, h_{M-1}) \quad [\mathrm{V}], \quad (2)$$

then the cross-correlation $R_k(H, B)$ between H and B at delay k is defined as:

$$R_k(H, B) = (H \otimes B)_k = \sum_{i=0}^{N-1} h_i b_{i+k} \quad [\mathrm{V}^2]. \quad (3)$$

If $H = B$, the cross-correlation becomes an auto-correlation, where the full energy $E(B)$ of B is focused into a single sharp peak (figure 2(a)) at zero delay ($k = 0$):

$$E(B) = R_0(B, B) = \sum_{i=0}^{N-1} b_i^2 \quad [\mathrm{V}^2]. \quad (4)$$

For delays $k \neq 0$, the auto-correlation function of LABS show low signal amplitudes which are referred to as 'signal side-lobes'. This so called 'pulse compression' is a main feature of LABS, where the 'matched

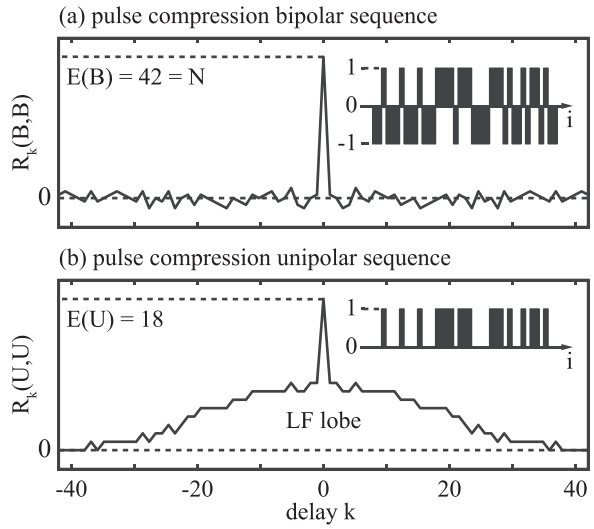

Figure 2. Comparison of the pulse compression (auto-correlation) of a bipolar modulation sequence B (a) and the respective unipolar sequence U (b). U is derived from B by setting all negative coefficients to zero resulting in a much lower signal gain of 18 to 1 compared to 42 to 1 in (a). Additionally, due to its DC characteristics, the auto-correlation signal in (b) exhibits a distinct low frequency (LF) lobe that significantly lowers the signal dynamics.

filter' $H = H_m = B$ squeezes a time distributed signal into a peak of short duration. The length of the correlation sequence always corresponds to $M + N - 1$ which is $2N - 1$ in the case of the matched filter.

However, bipolar LABS do not represent the physics of the slit mask (1 and 0) which is instead expressed by a sequence U according to

$$U = (u_0 \in \{0, 1\}, u_1, \cdots u_{N-1}) \quad \text{[V]}. \quad (5)$$

For this work, U has a length of $N = 42$ and is given with

$$U = (00010001000100011110111000011101001010110100) \quad \text{[V]}. \quad (6)$$

Its unipolar character has severe consequences on the applicability in the case of spatially modulated emission (figure 2(b)). Since negative coefficients are missing, the correlation function of the matched filter exhibits a triangular trend which is here termed 'low frequency (LF) lobe'. Such a shape is highly undesired since detection thresholds have to be increased while signal dynamics is simultaneously reduced. For this reason, the matched filter cannot be applied to unipolar sequences and must therefore be replaced by more sophisticated pulse compression filters. In the following, we take a closer look to the digital signal processing (DSP) principles that are used to evaluate the SNR gain of the used filters.

Any signal $x_a(t)$ e.g. received from photo diodes is analog at the first glance and has to be sampled by an analog-to-digital-converter in order to be fed to a digital signal processor. The analog signal then becomes a time-discrete signal X with samples x_n according to

$$x_n = x_a(n\Delta t) \quad \text{with} \quad n \in \mathbb{Z} \quad \text{[V]}. \quad (7)$$

Δt denotes the sampling interval, the sample frequency is given with $f_s = 1/\Delta t$. In our case, any detected signal is a superposition of the actual signal Au_n with scaling factor A, widely stationary signal offsets c_n and (white) noise e_n:

$$x_n = Au_n + c_n + e_n \quad \text{[V]}. \quad (8)$$

For simplification, the sequence is here sampled once per symbol.

In analogy to equation (4), *any* filter H can be normalized to exactly recover the energy of signal U:

$$E(U) = R_0(U, U) = R_0(H, U) \quad \text{[V}^2\text{]}. \quad (9)$$

Taking equation (9) into account, the expectation value $\mathcal{E}(\cdot)$ of the signal energy $E(X)$ can be calculated according to

$$\mathcal{E}(E(X)) = \mathcal{E}\left(\sum_{i=0}^{M-1} h_i x_i\right) \quad (10)$$

$$= \sum_{i=0}^{M-1} h_i \mathcal{E}(Au_i) + \sum_{i=0}^{M-1} h_i \underbrace{\mathcal{E}(c_i)}_{\mu_c} + \sum_{i=0}^{M-1} h_i \underbrace{\mathcal{E}(e_i)}_{=0} \quad (11)$$

$$= A \sum_{i=0}^{M-1} h_i u_i + \mu_c \sum_{i=0}^{M-1} h_i \quad (12)$$

$$= AE(U) + \mu_c \sum_{i=0}^{M-1} h_i \quad \text{[V}^2\text{]}. \quad (13)$$

In equation (13) we assumed that $c_n = \mu_c \,\forall\, n$, such that $\mathcal{E}(c_n) = \mu_c$. Additionally, we assumed the expectation value of the white Gaussian noise e_n to be zero.

If the filter coefficients are chosen such that

$$\sum_{i=0}^{M-1} h_i = 0, \quad (14)$$

the second part of equation (13) vanishes and it holds that

$$\mathcal{E}(E(X)) = \mathcal{E}\left(\sum_{i=0}^{M-1} h_i x_i\right) \tag{15}$$

$$= AE(U) \qquad [\mathrm{V}^2]. \tag{16}$$

Hence, if condition equation (14) is fulfilled, the signal energy can exactly be read off the correlation peak amplitude and stationary signal offsets do not influence the absolute value of the signal energy. Such filters show a high-pass, respectively bandpass behavior and exhibit no DC amplification. This important aspect of digital filtering here significantly differs from RADAR applications where stationary signal offsets are not critical in terms of signal processing. Since all considered filters are designed to obey equation (14), all stationary signal components are neglected in following calculations. In our case, deviations of the resulting signal amplitude may only be due to the remaining noise (see below).

It is common practice (spectroscopic notation) to define the SNR of a signal as the ratio of the maximal expected signal amplitude $U_{\max}$ and the standard deviation σ_{in} of the superimposed noise e_n (figure 7, top). For an unbiased input signal X, the SNR is defined as

$$\mathrm{SNR} = \frac{AU_{\max}}{\sigma_{\mathrm{in}}}. \tag{17}$$

Since $U_{\max}(= 1\ \mathrm{V})$ as well as σ are expressed in Volts, SNR is dimensionless. Through filtering procedures, the noise (standard deviation) is amplified by A_e which is defined as

$$A_e = \sqrt{\sum_{i=0}^{M-1} h_i^2} \qquad [\mathrm{V}]. \tag{18}$$

The SNR gain G of a pulse compression filter applied to a scaled unipolar binary modulation sequences U (with amplitude 1 V) and scaling factor A can then be calculated according to

$$G = \frac{\mathrm{SNR}_{\mathrm{out}}}{\mathrm{SNR}_{\mathrm{in}}} = \frac{AE(U)}{\sigma_{\mathrm{in}} A_e} \Big/ \frac{AU_{\max}}{\sigma_{\mathrm{in}}}. \tag{19}$$

$U_{\max}$ is the maximal input amplitude of the unscaled sequence and equals 1 V, A is the (dimensionless) signal scaling factor. The SNR gain G can be rewritten as:

$$G = \frac{AE(U)}{\sigma_{\mathrm{in}} A_e} \frac{\sigma_{\mathrm{in}}}{AU_{\max}} = \frac{\frac{1}{\mathrm{V}}E(U)}{A_e} = \frac{E'(U)}{A_e}. \tag{20}$$

$E'(U)$ corresponds to the signal energy $E(U)$ which is normalized by V^{-1} with regard to its dimension. Therefore, G is dimensionless.

3. Digital filter design

In the following, we discuss three different filter types which mainly differ in their SNR gain, signal side-lobes of correlation signals and cost of implementation. Since related signal theory has consistently been derived in the previous section, we further on relinquish any units.

The already discussed matched filter has been used by Kiesel *et al* [11–13] for unipolar modulation sequences. Besides disturbing LF lobes, correlation signals from this filtering technique show remarkable signals offsets. Suppressing the LF lobe as well as stationary DC components, a subsequent differentiation of filtered signals has been performed by Kiesel *et al*. However, by the use of the differential operator DIFF $= (-1, 1)$, the differentiation can directly be incorporated with the matched filter to yield the here called 'diffed filter' H_{df} as a joint pulse compression filter. It is built according to

$$\mathrm{DIFF} \otimes (U \otimes U) = \underbrace{(-\mathrm{DIFF} \otimes U)}_{H_{\mathrm{df}}} \otimes U \tag{21}$$

$$= H_{\mathrm{df}} \otimes U. \tag{22}$$

Note that the diffed filter is always longer than the modulation sequence ($M = N + 1$) resulting in a correlation signal of length $2N$. The correlation signal received from diffed filtering is point-symmetric and shows two distinct peaks (positive and negative) around $k = 0$ (figure 3(b)) with a steep zero transition in between. At zero delay it can be noticed, that the filter coefficients are highly sensitive to edge transitions of the modulation sequence, whereas constant signal parts are suppressed (multiplication with zeros). Especially for sequences with a relative low number of edge transitions, a large part of the signal energy is not taken into account using this filter type. Moreover, the output signal shows significant peak side-lobes resulting in not optimized signal dynamics. For the chosen sequence (equation (6)), the theoretical SNR gain of the diffed filter is +6.99 dB which can be further improved by more sophisticated filter architectures discussed below.

The so called 'balanced filter' H_{bal} is a rather basic filter type which has been designed specifically for the concept of spatially modulated emission. It aims to exceed the SNR gain that can be achieved with diffed filtering by exploiting a higher amount of signal energy. The filter coefficients are derived by subtracting the mean value $\overline{H}_{\mathrm{m}}$ from each filter coefficient of the matched filter:

$$H_{\mathrm{bal}} = H_{\mathrm{m}} - \overline{H}_{\mathrm{m}}. \tag{23}$$

As can be seen from figure 3(c), the filter coefficients equal the modulation sequence and are balanced around zero. This ensures higher SNR gain compared to the diffed filter as well as an effective suppression of DC components since $\sum_{i=0}^{M-1} h_i = 0$. Moreover, compared to the diffed filter, the correlation signal shows a sharp positive peak which is surrounded by relatively small side-lobes superimposed by a LF fluctuation. This fluctuation is due to the imbalanced energy distribution of the signal in time according to the modulation sequence used.

Signal side-lobes can in general be described by the so called 'peak-to-side-lobe ratio' (PSLR) and the

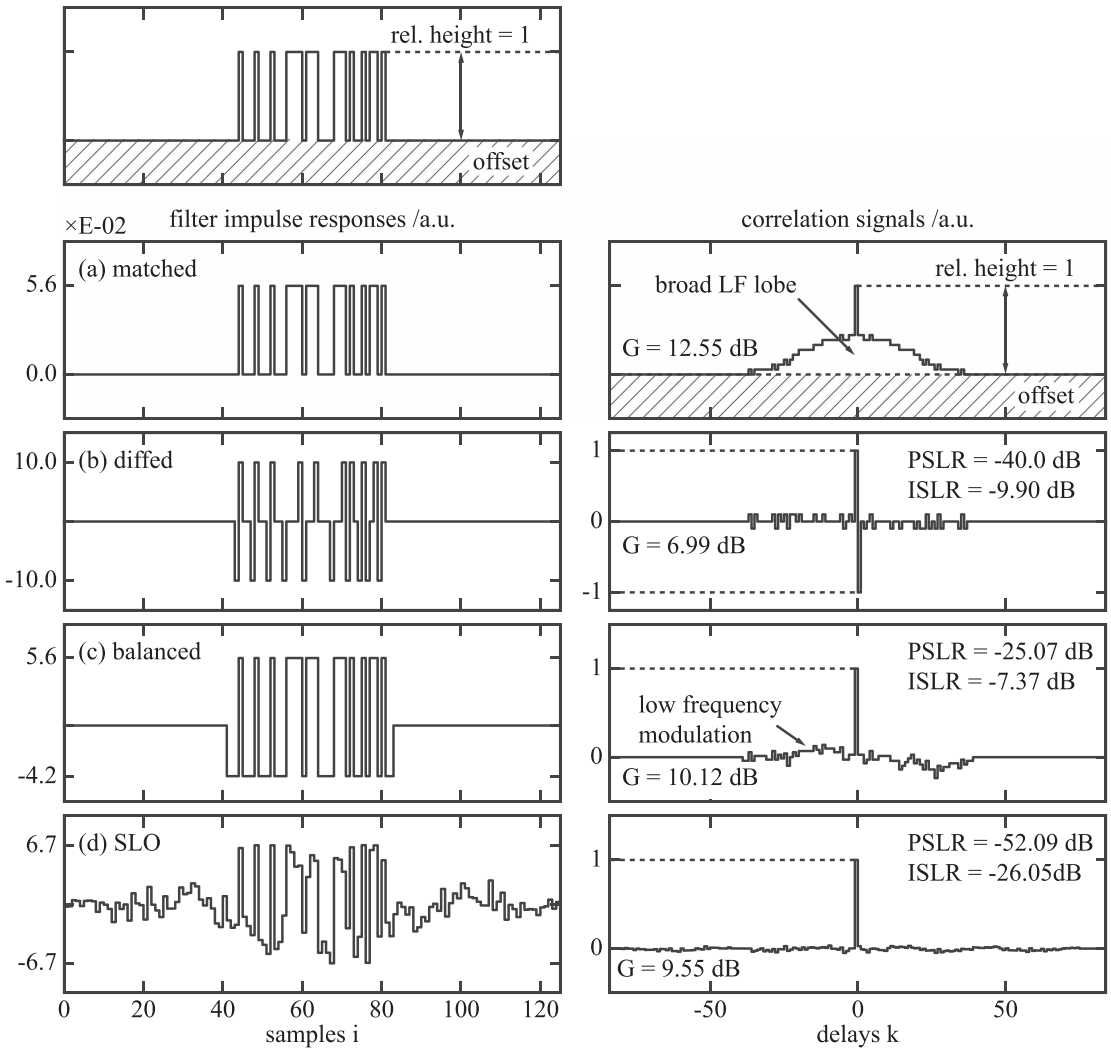

Figure 3. The received unipolar signal (top) with additional DC offset is processed with different filter types (left) to yield the corresponding output signals (right). Matched filtering (a) results in the highest achievable SNR gain (+12.55 dB), but cannot be used due to its distinct LF lobe resulting in poor signal dynamics. The diffed filter (b) shows a point-symmetric signal around zero delay ($k = 0$) with a steep zero crossing behavior, but has a poor SNR gain (+6.99 dB). With respect to SNR gain (+10.12 dB), the balanced filter (c) is an improvement compared to (b), but has a higher ISLR and PSLR than the diffed filter. An excellent compromise is the side-lobe optimized (SLO) filter (d), showing the best PSLR and ISLR behavior with a slightly lower SNR gain (+9.55 dB) compared to (c). However, this filter type has the drawback of higher computational costs (number of filter coefficients) and demands more hardware resources. Due to their different lengths and filter coefficients, all filters are normalized to yield the same peak amplitude of their output signals referred to a relative input signal amplitude of 1.

'integrated side-lobe ratio' (ISLR) [14], which are defined as:

$$\mathrm{PSLR} = \frac{1}{R_0^2}\left(\max_{k\neq 0}|R_k|\right)^2, \tag{24}$$

$$\mathrm{ISLR} = \frac{1}{R_0^2}\sum_{k\neq 0} R_k^2. \tag{25}$$

The lower these values the better the combination of a LABS and a filter is suited to maximize the dynamic range of the expected output signals as well as to reduce the crosstalk between neighboring events. Additionally, these rating criteria can be used to design optimized filters with highly suppressed side-lobes in their correlation signals. The technique is well known from RADAR applications in order to maximize signal dynamics and is persistently described in relevant literature on this subject [15–20]. In the field of RADAR technology, such filters are widely known as 'mismatched filters' and are here referred to as 'side-lobe optimized (SLO) filters' H_{SLO}.

For unipolar signals it has been shown that optimizing filters with respect to the ISLR yields homogeneous output signals with negligible LF fluctuations and highly suppressed side-lobes and is therefore the preferred optimization method for the concept of spatially modulated emission. As optimization routine, the so called 'differential evolution' by Storn and Price [21] has proven to be a suitable non-linear algorithm and is available as optimization tool e.g. in NI LabVIEW.

Sidelobe optimized filters are typically longer than the initial modulation sequence in order to regulate the side-lobe amplitudes while signals travel through the filter. In particular, a length of $M = 3N$ has shown remarkable results and has turned out to be the

Table 1. Filter characteristics. Note that the negative main lobe of the diffed filter is excluded for determining the PSLR and ISLR.

Filter type	G_{theo}	PSLR	ISLR	Length M
Matched	12.55	N.A.	N.A.	N
Diffed	6.99	−40.00	−9.90	$N+1$
Balanced	10.12	−25.08	−7.37	N
SL optimized	9.55	−52.09	−26.05	$3N$

(all values in dB)

preferred filter length for our purpose. When using longer SLO filters of length $3N$, U must also be extended by two all-zero sequences Z, both of length N, to yield the sequence U_0 according to:

$$U_0 = (ZUZ). \tag{26}$$

This ensures U to be centered in the filter at zero delay within the correlation process. The correlation peak then occurs at $R_0(H_{SLO}, U_0)$ such that the ISLR can be calculated according to equation (25). In general, the filter coefficients of H_{SLO} correspond to rational numbers and are shown in figure 3(d). Since these coefficients show minor mismatches to the modulation sequence, the SLO filter can be expected to have a slightly lower SNR gain than the balanced filter. In table 1, the values for the theoretical SNR gain G_{theo}, ISLR and PSLR are given for the sequence in equation (6).

Without experimental modifications, higher SNR gain as well as better signal dynamics can be achieved by proper DSP. The balanced filter is an easy-to-build pulse compression filter which shows a +3.13 dB gain compared to the state-of-the-art diffed filter. Drawbacks arise from higher signal side-lobes (ISLR + 2.53 dB) and are mainly due to LF fluctuations which are characteristic for the chosen sequence. However, these fluctuations can easily be reduced by choosing suitable modulation sequences with more homogeneous energy distributions. Signal side-lobes generally lower the signal dynamics and cause false detections originating from side-lobe superpositions of neighboring events. Therefore, the technique of SLO filtering has been adapted from RADAR applications to unipolar sequences with the additional constraint of DC suppression. For the sequence in equation (6), the corresponding SLO filter with $M = 3N$ shows slightly worse SNR gain (−0.57 dB) compared to the balanced filter but exhibits a much better side-lobe suppression (ISLR −18.68 dB). This allows for higher cell densities in the sample flow due to improved signal dynamics and peak discrimination. Drawbacks of this filter type arise from higher processing resources necessary for implementation (number of filter coefficients) as well as from longer signal runtimes (group delay).

4. Experiment

Besides theoretical considerations, we also describe an experimental approach for validating theoretically derived SNR gains of the various filter types. Our setup consists of a micro-fluidic channel (figure 1) with a cross-section of 480 ± 20 μm × 18 ± 2 μm through which we pump fluorescent latex spheres (Polysciences Fluoresbrite® 2.0 μm YG microspheres) with a diameter of 2 μm. For sample preparation, we highly dilute the original particle concentration ($c_0 = 5.68 \times 10^9$ ml^{-1}) by pipetting 1.174 μl aqueous particle suspension into 50 ml distilled water. Consequently, the particle concentration of the sample liquid corresponds to 133.33 μl^{-1}. For the chosen flow rate of $Q = 0.5$ μl s^{-1}, we then expect a number of ≈2000 particles within each measurement interval of 30 s. Making sure that our apparatus detects only valid signals at high SNR, we manually counted particles at the highest SNR and found a nominal value of 2000 ± 25 particles within a 30 s data-set. This interval perfectly matches the range prescribed by the standard deviation of the underlying Poisson statistics ($\pm\sqrt{2000} = \pm 45$). Uncertainties arise e.g. from signal superpositions which cannot be clearly separated by bare eye. Since all of the tested filters converge towards this manually determined value, the false negative as well as the false positive detection rate can be estimated to be ⩽1.25 % at high SNR. The concentration has been chosen for two reasons. First, adequate statistics can be generated out of 2000 nominal events. Second, the concentration is still low enough to ensure that the majority of particles is individually passing the detection zone. The latter aspect is the main requirement to precisely benchmark the used digital filters with respect to SNR gain. Our syringes exhibit a volume of 10 ml. Setting the flow rate to 0.5 μl s^{-1} and performing 22 measurements, a total volume of less than 1 ml is processed by our setup. Moreover, since the complete data set is recorded all at once, the preparation of a single sample is sufficient to serve our measurement series. Deviations in the nominal count rate might only occur due to particle sedimentation in the syringes. However, since the total measurement duration is less than 15 min, this effect is negligible.

The relay optics (consisting of two Nikon Plan-Apo λ10×, NA = 0.45) together with a dichroic beam splitter (BS) serves to reflect the excitation light (1.3 W @ 448 nm) from the light source (laser) to the channel. Fluorescent light from the particles is guided through the BS and is then focused on the non-periodic slit mask. The mask is made of glass with an etched chrome plating forming the binary sequence. A lens in between the slit mask and the detector (SiPM) collects the fluorescent light passing through the slits. Furthermore, excitation and emission filters are used to clean up the laser spectrum and to suppress excitation light in the detection.

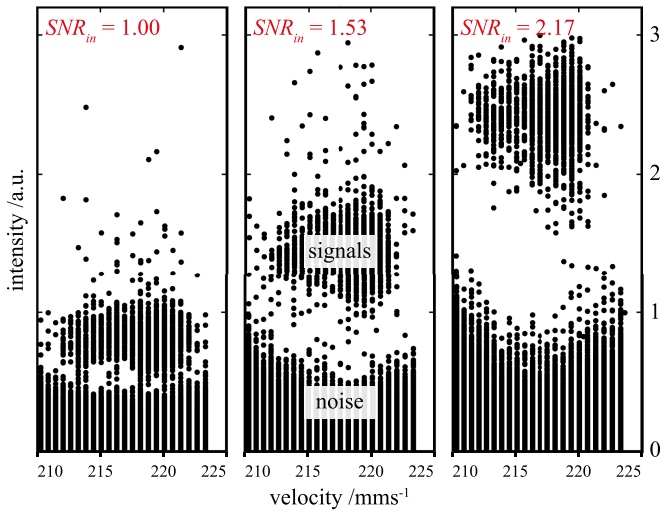

Figure 4. Detected peaks for a narrow velocity distribution of fluorescent particles at different SNR levels. The upper populations contain valid signal peaks from fluorescence signals, whereas the lower populations originate from peaks found at the respective noise level. Reducing the excitation intensity causes the signal peaks to merge into the noise population and the count rate of valid events decreases. The rate at which signal peaks converge to the noise level is a function of the input SNR and depends on the used filter. Consequently, the count rate at certain input SNRs can be used to benchmark the relative filter performance in terms of SNR gain.

Reducing spherical aberrations from the optics and ensuring a homogeneous illumination, the slit-mask and therefore the detection zone has a length of ≈1.33 mm.

With velocities of 220 ± 5 mm s^{-1}, particles pass the detection zone and the raw signal is recorded on disk. Testing the SNR gain of the different filters, a series of measurements is taken, where the excitation intensity is successively lowered to reduce the fluorescence emission. With decreasing excitation intensity, the signal strength from emission converges closer to the electronic noise level (figure 4).

Ensuring particles are individually passing the detection zone, the particle density of the sample flow is adjusted such that individual signals do not overlap (see above). This avoids coincidences and erroneous detections which corrupt the count rates used to determine the individual SNR gains of the filters. The raw data on disk is then processed with the three filter methods and the resulting count rates are determined as a function of the respective input SNR (equation (17)).

Since particles show a tight distribution of their velocities (±2.3 %), filters are adapted to match the time-scaled particle signals within this range [22, 23]. This yields a filter bank $H_{F,v}$ for each filter type F with a finite number of processing channels covering the whole velocity spectrum. Due to the small variation in filter length, counting results are not affected. Raw signals from each measurement are then post-processed for each filter type to calculate the respective correlation matrix $R_{k,v}(H_{F,v}, X)$. The correlation maximum finally occurs for the velocity matching filter (figure 5).

Consequently, the correlation matrix is subject to progressive peak search where the highest peaks with certain expected widths (depending on the particle velocity) are recorded. This record contains all valid

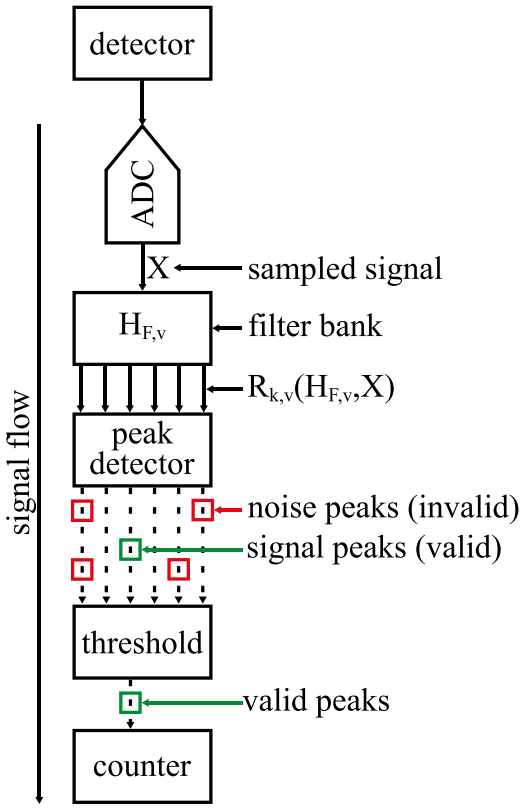

Figure 5. Digital signal processing chain of the used evaluation algorithm. Analog signals are first sampled by an analog-to-digital converter (ADC) and then filtered with respect to a certain velocity spectrum (filter bank $H_{F,v}$). The consecutively generated correlation matrix $R_{k,v}(H_{F,v}, X)$ consists of a number of signal peaks caused by noise (invalid) and true signals (valid). These peaks are discriminated by applying a constant threshold of 6.5 times the median peak amplitude. Valid signal peaks are counted if the peak amplitude exceeds this threshold. The processing chain as well as the threshold level is the same for any tested filter. Therefore, signal-to-noise performance between different filter types can be directly compared by means of count rate.

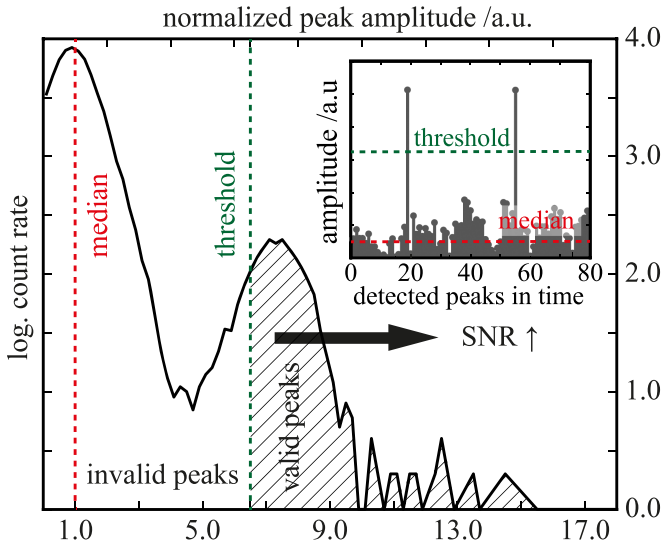

Figure 6. Discrimination of all the detected peaks within the correlation signals. The inset at the upper right shows an interval of 80 detected peaks within a certain time. These peaks originate from valid signal peaks, signal side-lobes and noise. The main picture shows the statistics of all peaks found in a data set. Since peaks from pure noise predominate the number of valid signal peaks, all peaks can be normalized according to the median peak amplitude. A threshold is then defined as a multiple of the median peak amplitude. The threshold is set to 6.5 times the median to clearly separate the valid signal peaks from noise. This relative threshold is applied to each data set and filter to ensure a fair comparison between the filters. As the input SNR increases with excitation light intensity, the peak population of valid signal peaks detaches from the noise-peak population and shifts to the right. This results in higher count rates.

signal peaks as well as false detections originating from side-lobes, side-lobe superpositions and noise. Next, for each measurement and each filter method the peak intensities are normalized to the respective median peak intensity of all recorded peaks. To all data sets the same fixed threshold is then applied to identify valid peaks (figure 6). This threshold is set to 6.5 times the median peak amplitude and is chosen to meet two criteria regarding the count rate (figure 7): (1) no false negative and false positive detections at high SNR; (2) no false positive detections in pure noise signals and for measurements with low SNR ($SNR_{in} < 0.3$). Criterion (1) is fulfilled as the particle count saturates for each filter at the expected value of 2000 (compare experimental section). The adherence to criterion (2) can be seen from the zero count rate at low SNR for all filters and from a reference experiment (not shown) with pure noise signals. Thus, the selected setting of the threshold to 6.5 times the median peak intensities ensures no false positive detections and no false negative detections for all filter types. Consequently, within the statistical error rate of ± 2.24 %, the count rate in figure 7 resembles the true positive count rate for each experiment. E.g., at half count rate (figure 7), the false negative rate equals 50 % which corresponds to a critical SNR that is individually determined for each filter type. Thus, the filter performance can be considered the same at high input SNR and filters can be directly compared for lower input SNRs. For calculating the input SNR, the average amplitude of valid peaks is used as signal strength and the noise level is derived from the raw signals of the illuminated channel without particles in the fluid. Signal strength as well s noise level depend linearly on the incident excitation intensity. Therefore, for low count rates, where it is impossible to directly measure the signal strength, the input SNR is determined from extrapolation. This evaluation procedure results in a total count rate for each measurement and each filter which is plotted as a function of the input SNR in figure 7. For high input SNR, all filter methods show almost the same count rate (approx. 2000 particles for the chosen measurement duration of 30 seconds) proving a proper selection of the threshold.

However, for low input SNR count rates vanish which demonstrates that the threshold is high enough to prevent positive false detections originating from pure noise. The count rate of the diffed filter drops at relatively high input SNR. Balanced and SLO filtering show almost comparable performance and maintain the count rate for lower input SNR than the diffed filter. Each count rate follows a Gaussian error function with the skewness depending on the noise gain of each filter. The input SNR at the turning point at the half maximum count rate is called the critical input SNR SNR_{crit}. For each filter method, the SNR gain G is associated with the respective $SNR_{crit,filter}$ and can be compared to the diffed filter with lowest $SNR_{crit,diffed}$. This yields a relative (experimental) SNR gain $G'_{exp,filter}$ for each filter type:

$$G'_{exp,filter} = 10 \log\left(\frac{SNR_{crit,filter}}{SNR_{crit,\ diffed}}\right). \qquad (27)$$

Experimental results are given in table 2. The comparison of the measured gain $G'_{exp,filter}$ and the

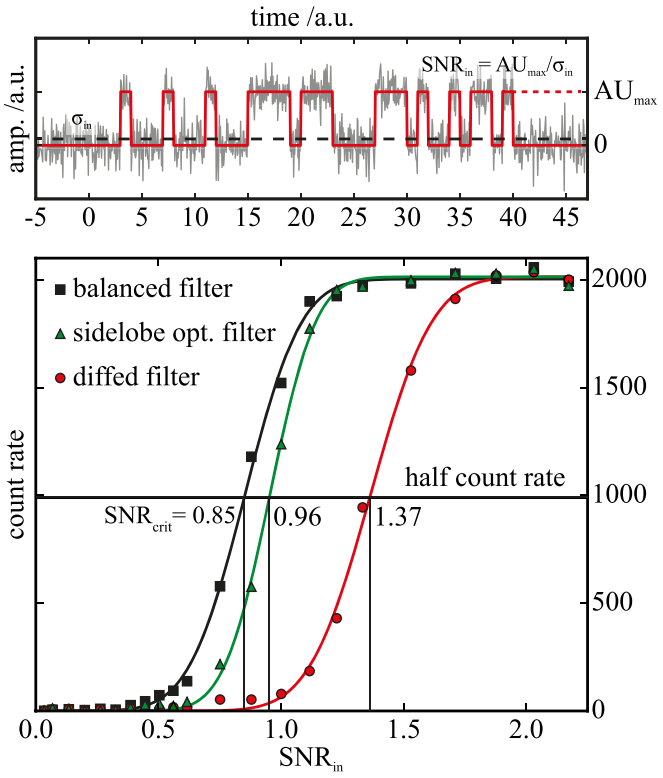

Figure 7. Count rates of found signal peaks versus input SNR. The input SNR at the half count rate indicates the critical input SNR (SNR_{crit}) of each filter and can be used to calculate relative SNR gains.

Table 2. Relative experimental SNR gains estimated from the half count rate of each filter.

Filter type	SNR_{crit}	G'_{exp} /dB	G'_{theo} /dB	$\Delta G'$ /dB
Diffed	1.37	0.00	0.00	0.00
Balanced	0.85	+4.15	+3.13	+1.02
SL optimized	0.96	+3.09	+2.56	+0.53

(relative) theoretical gain $G'_{theo,filter}$ for each filter type shows in general a good agreement between the theory and experiment. However, the experimental gain does not fully match the theoretical expectation. Deviations may be due to differences between the CAD model of the slit mask and the physical representation and design of the digital filters. Inaccurate slit width and position inevitably lead to deviations of the measured signals from the theoretical representation. Additional deteriorations may originate from factors such as inhomogeneous illumination or scratches introduced during chip fabrication. Moreover, any passing object has a size d which can be expressed as a fraction the physical mask length L. Due to this extent, the original binary sequence experiences a smoothing of its shape (smoothed edges with a minimum relative width of d/L). For each processing method, these deviations from the ideal signal lower the peak amplitude and effectiveness of pulse compression that could be achieved compared to perfect signals. However, the experimental results exceed the theoretical expectation by up to +1 dB. This suggests that the above described experimental deviations from the theoretical assumption do not affect all filters equally. As the diffed filter is most sensitive to the quality of the signal edges, we assume that its SNR gain is overestimated. Nevertheless, the experimental result confirms the theoretical expectation that the balanced filter and the SLO filter clearly outperform the diffed filter in terms of SNR gain. Moreover, the balanced filter is slightly better than the SLO filter even though a LF fluctuation and significantly larger side-lobes remain in the correlation signal. In the case of overlapping signals from neighboring particles, valid signal peaks are superimposed by side-lobes and LF fluctuations. It can be expected that the SLO filter is better suited in that situation. However, our theories as well as our experimental results do not cover that question, since particle densities have been chosen sufficiently low to avoid particle coincidences.

5. Conclusion

In conclusion, it is demonstrated that advanced signal processing methods can improve the SNR gain of the method of spatially modulated emission almost by a factor of 1.6 (balanced filter). We introduced two improved signal processing approaches, i.e. balanced filtering and SLO filtering, which can be applied to our modulation principle. Both methods enable inherent baseline correction and improved noise suppression. Moreover, SLO filtering may be a promising method

for increased signal dynamics compared to diffed and balanced filtering which is particularly interesting for high particle densities. As the demonstrated improvements in sensitivity solely gear into the signal processing chain they can be used for any technical implementation of spatially modulated emission without need for physical changes.

Acknowledgments

We thankfully acknowledge the European Research Council (ERC) for funding the project under grant # 258604.

References

[1] Shapiro H M 2003 *Practical Flow Cytometry* (New York: Wiley)
[2] Levanon N and Mozeson E 2004 *Radar Signals* (New York: Wiley)
[3] Mertens S 1996 Exhaustive search for low-autocorrelation binary sequences *J. Phys. A: Math. Gen.* **29** L473–81
[4] Coxson G E and Russo J C 2005 Efficient exhaustive search for optimal-peak-sidelobe binary codes *IEEE Trans. Aerosp. Electron. Syst.* **41** 302–8
[5] Nunn C J and Coxson G E 2008 Best-known autocorrelation peak sidelobe levels for binary codes of length 71 to 105 *IEEE Trans. Aerosp. Electron. Syst.* **44** 392–5
[6] Shanmugam S K, Mongrédien C, Nielsen J and Lachapelle G 2008 Design of short synchronization codes for use in future GNSS system *Int. J. Navig. Obs.* **2008** 246703
[7] Golay M J E 1982 The Merit factor of long low autocorrelation binary sequences *IEEE Trans. Inf. Theory* **28** 543–9
[8] Militzer B, Zamparelli M and Beule D 1998 Evolutionary search for low autocorrelated binary sequences *IEEE Trans. Evol. Comput.* **2** 34–9
[9] Gallardo J E, Cotta C and Fernández A J 2009 Finding low autocorrelation binary sequences with memetic algorithms *J. Appl. Soft Comput.* **9** 1252–62
[10] Oppenheim A V and Schafer R W 2014 *Discrete-Time Signal Processing* 3rd edn (Harlow: Pearson Education Limited)
[11] Baßler M, Kiesel P, Schmidt O and Johnson N M 2008 Means for counting particles suspended in a fluid *US Patent* 0080181827A1
[12] Kiesel P, Baßler M, Beck M and Johnson N M 2009 Spatially modulated fluorescence emission from moving particles *Appl. Phys. Lett.* **94** 041107
[13] Kiesel P, Beck M and Johnson N M 2011 Monitoring CD4 in whole blood with an opto-fluidic detector based on spatially modulated fluorescence emission *Cytometry* A **79A** 317–24
[14] Levanon N 2005 Cross-correlation of long binary signals with longer mismatched filters *IEE Proc., Radar Sonar Navig.* **152** 372–82
[15] Rihaczek A W and Golden R M 1971 Range sidelobe suppression for Barker codes *IEEE Trans. Aerosp. Electron. Syst.* **AES-7** 1087–92
[16] Ackroyd M H and Ghani F 1973 Optimum mismatched filters for sidelobe suppression *IEEE Trans. Aerosp. Electron. Syst.* **AES-9** 214–8
[17] Zoraster S 1980 Minimum peak range sidelobe filters for binary phase-coded waveforms *IEEE Trans. Aerosp. Electron. Syst.* **AES-16** 112–5
[18] Ackroyd M H 1982 Economical filters for range sidelobe reduction with combined codes *Radio Electron. Eng.* **52** 309–10
[19] Rohling H and Plagge W 1989 Mismatched-filter design for periodical binary phased signals *IEEE Trans. Aerosp. Electron. Syst.* **25** 890–7
[20] Chen X H and Oksman J 1990 A new algorithm to optimize Barker code sidelobe suppression filters *IEEE Trans. Aerosp. Electron. Syst.* **26** 673–7
[21] Storn R and Price K 1995 Differential evolution—a simple and efficient adaptive scheme for global optimization over continuous spaces *Technical Report* TR-95-012 (ICSI)
[22] Sommer C, Quint S, Spang P, Walther T and Baßler M 2014 The equilibrium velocity of spherical particles in rectangular microfluidic channels for size measurement *Lab Chip* **14** 2319–26
[23] Sommer C, Quint S, Spang P, Walther T and Baßler M 2014 Studying the Segré–Silberberg effect by velocitmetry in microfluidic channels *Adv. Fluid Mech.* X **82** 256–77