Supplementary Materials for

High-precision, low-complexity, high-resolution microscopy-based cell sorting

Tobias Gerling, Neus Godino, Felix Pfisterer, Nina Hupf and Michael Kirschbaum*

*Corresponding author. Email: michael.kirschbaum@izi-bb.fraunhofer.de

Video S1 (separate file)

Video S1 shows a sequence of a sorting of red and green coloured cells using a 4X objective. The video shows the electrode structures of the sorting electrode array E3 (small white rectangles) and the beginning of the separating electrode E4 (pointed triangle at the right edge, see also Figure 1a and b of the main text as well as Figure S1). The large rectangle on the left edge of the image indicates the observed ROI. The scale bar corresponds to 50 µm.
The red cells serve as target cells. They are detected in the ROI and receive an identification number, which is displayed as an overlay in the video together with the radius of the cell. From the continuous tracking over several frames, a velocity is determined for each target cell, which is used to precisely switch on the individual electrode segments. Active electrode segments are visualised by a red colouring. The sequential activation of the electrodes leads to a gradual deflection of the target cells so that they can be sorted out from the stream of green non-target cells even at small cell distances.

Video S2 (separate file)

Video S2 shows a sequence of membrane- and cytosol-stained cells as they appear in the ROI at 60X magnification during real-time morphology-based cell sorting. The cells are imaged and classified in real-time into target (i.e., membrane-stained) and non-target cells (i.e., cytosol-stained). In addition, their velocity is determined, which together with the classification forms the basis for the subsequent sorting process (see text for details on methodology). Note that the video is slowed down considerably for visualisation.

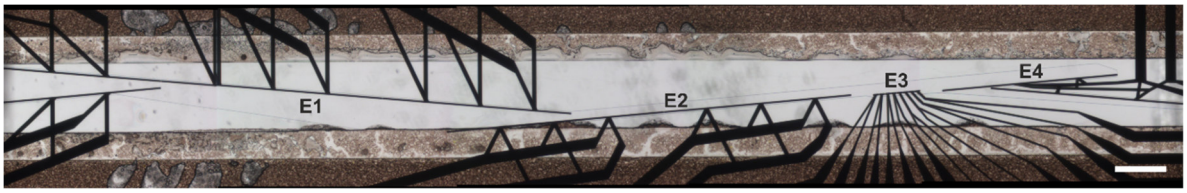

Fig. S1: Microscopic image of the sorting channel. Electrodes are labelled according to Figure 1a of the main text. Scale bar 500 µm.

Theoretical estimation of the achievable purity based on Poisson statistics

A critical feature for the performance of any cell sorting method is the achievable purity of the sorting result and the yield of target cells as a function of sample throughput. In microfluidic sorting, where cells are driven through a microchannel in a pearl chain configuration and are deflected vertical to the flow direction over a certain distance (i.e., the sorting window) in course of their selection, the performance is closely related to the effective length of this window. In an ideal system, the target and non-target cells that pass through the channel are totally uncorrelated and thus their occurrence in the sorting window follows a Poisson distribution with k the number of occurrences and μ the mean expected number of cells in the sorting window at any given time.

$$P_{\mu}(k) = \frac{\mu^{k}}{k!} e^{-\mu} \tag{1}$$

Accordingly, it cannot be ruled out that multiple cells appear in the sorting window at the same time (e.g., a target and a non-target cell), which cannot be processed independently of each other. In terms of purity and yield, this problem can be addressed in two ways. Either discard both cells to maximise the purity of the sorting result at the expense of yield ("purity sort"), or sort the target cell together with the non-target cell according to the opposite motivation ("yield sort"), the latter also forming the basis for our studies. In case of a "yield sort", where all target cells are sorted independently of the neighbouring non-target cells, the theoretical yield is by definition 100%.
The purity p achieved at a given throughput can now be used to characterize the performance of the system. It is calculated from the ratio of target cells (N_{T}) to the total number of target and non-target (N_{nT}) cells sorted out.

$$p = \frac{N_{\mathrm{T}}}{N_{\mathrm{T}} + N_{\mathrm{nT}}} \tag{2}$$

Let x be the percentage of target cells in the sample. Then the mean expected number of target (μ_T) and non-target cells (μ_{nT}) in the sorting window at any given time point calculates to $x\mu$ and $(1\text{-}x)\mu$, respectively. Accordingly, each time a target cell is sorted in the sorting window, on average there are also μ_{nT} non-target cells in the sorting window, which are sorted out together with the target cell. According to this, the purity of the sorting result calculates to

$$p = \frac{N_{\mathrm{T}}}{N_{\mathrm{T}} + N_{\mathrm{sort}} * \mu_{\mathrm{nT}}} \tag{3}$$

with N_{sort} = total number of sorting events necessary for sorting N_T target cells. It is important to note that fewer than N_T sorting events are required to sort out N_T target cells from the sample, as more than one target cell appear in the sorting window in certain cases due to the Poisson statistics described above. As can be easily seen, this value is a function of the expected number of target cells in the sorting window μ_T.
In a stochastic sorting procedure, one would expect N_T target cells after N_T/μ_T sorting operations. In a reactive, targeted sorting procedure, however, a sorting process is only triggered if there is actually at least one target cell in the sorting window, hence

$$N_{\mathrm{sort}} = \frac{N_{\mathrm{T}}}{\mu_{\mathrm{T}}}\left(1 - P_{\mu_T}(0)\right) \tag{4}$$

$$= \frac{N_{\mathrm{T}}}{\mu_{\mathrm{T}}}(1 - e^{-\mu_{\mathrm{T}}}) \tag{5}$$

The purity p of the sorted cell fraction is therefore

$$p = \frac{N_{\mathrm{T}}}{N_{\mathrm{T}} + \frac{N_{\mathrm{T}}}{\mu_{\mathrm{T}}}(1 - e^{-\mu_{\mathrm{T}}})\,\mu_{\mathrm{nT}}} \tag{6}$$

Since $\mu_T = x\mu$ and $\mu_{nT} = (1\text{-}x)\mu$, the purity can be expressed as

$$p = \frac{1}{1 + \frac{(1-x)}{x}(1 - e^{-x\mu})} \tag{7}$$

Moreover, μ can be expressed as the ratio between the width of the sorting window d and the average cell distance, while the average cell distance in turn is the ratio of mean velocity of the cells v and throughput θ. For a given x, d and v (here: 580 µm s^{-1}) the maximum achievable purity of a yield sort can be expressed as a function of θ (see Fig. S2):

$$p = \frac{1}{1 + \frac{(1-x)}{x}\left(1 - e^{-\frac{xd\theta}{v}}\right)} \tag{8}$$

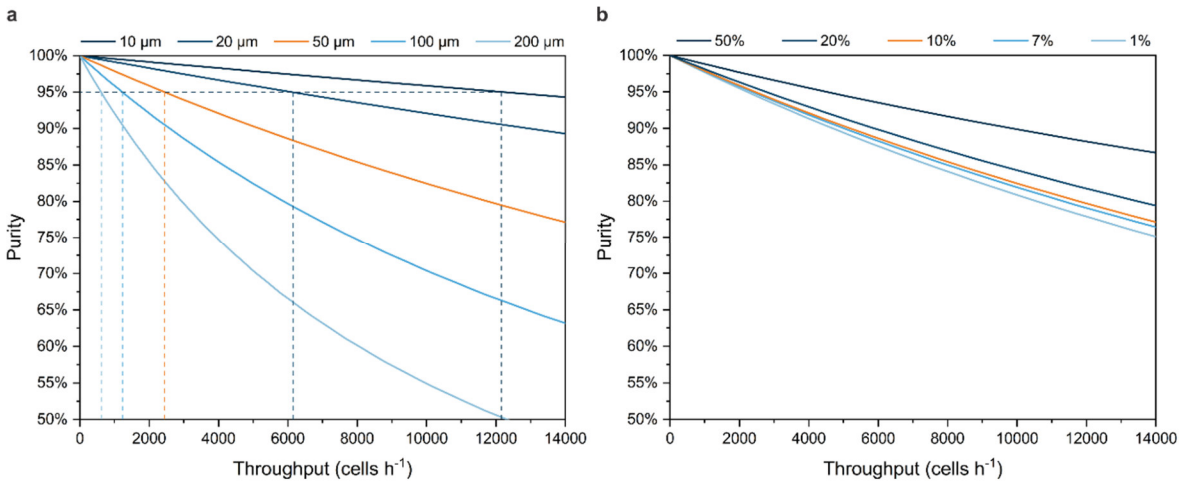

Fig. S2: Maximum achievable purity of a yield sort at a mean cell velocity of 580 µm s^{-1} and either (a) constant target cell concentration of 10% and variable sorting window width or (b) constant window width of 50 µm and variable target cell concentration. While the width of the sorting window has a significant impact on the achievable purity or throughput of a sorting process, the target cell concentration in the lower range of up to 20 % is of rather minor importance. Conditions reflecting the experimental situation in the current study are marked in orange.

Development of criteria for image-based cell classification

The cell model we chose to characterise our image-based analysis consisted of a mixture of cytosol- and membrane-stained Jurkat T cells. The use of dyes of the same colour was to enforce a classification of the cells based on morphological characteristics rather than different fluorescent colours. For the classification of a cell into target or non-target cell, sixteen radially evenly distributed brightness profiles across the outer cell boundary were collected, normalised and analysed for their shape.

We distinguished two types of profiles: one with gradually increasing values toward the centre of the cell (C-type profile) and another one, in which the values reach a maximum and then drop below a certain threshold (M-type profile).

We based our classification on the number of M-type profiles in a cell (see Fig. 4, main text). To establish the exact criteria for distinguishing between the two cell classes, we analysed images from a self-generated dataset containing 458 images of cytosolic-stained cells and 500 images of membrane-stained cells (200x200 pixels, 0.5 ms exposure time and dynamic range of 100-5000). The threshold for distinguishing between C- and M-type profiles was fixed at 0.9.

We then analysed each of the cell images in the dataset using different classification criteria, identifying membrane-stained cells (i.e., target cells) via a varying number of M-type profiles among the sixteen profiles (n >= 1; n >= 2; n >= 3; n >= 4). We then created an error matrix for the different classification criteria and quantified sensitivity, selectivity and the expected purity of a sorted cell sample of different mixing ratios (see Table S1 and S2).

In order to define the most appropriate classification decision, we calculated sensitivity, specificity and the resulting maximal sorting performance (i.e., purity) when using each classification criteria:

$$Sensitivity\ (i.e., yield) = \frac{true\ positives}{true\ positives + false\ negatives} \quad (9)$$

$$Specificity = \frac{true\ negatives}{true\ negatives + false\ positives} \quad (10)$$

$$Purity = \frac{x * sensitivity}{x * sensitivity + (1 - x) * (1 - specificity)} \quad (11)$$

where *x* and *(1-x)* are the ratios of target cells and non-target cells in the initial sample.
Based on the values for sensitivity, specificity and purity shown in Table S2, we decided to use Criteria III as our decision for the image based sorting. At least three M-type profiles out of sixteen profiles are necessary for considering a cell a membrane-stained cell. The values of expected purity, for both target/non-target initial sample ratio, are over 95%, so the error in the sorting associated to the classification in both cases is below 5%.

Table S1: Recovery rate. Recovery rate is the ratio between the number of target cells identified in the microchannel upstream of the sorting electrodes and the number of cells collected downstream of the sorting outlet and counted externally.

	Magnification	Cell counting in channel	Cell counting in well	Recovery Rate
Test 1	4X	323	308	95%
Test 2	4X	296	260	88%
Test 3	4X	344	280	81%
Test 4	4X	407	375	81%
Test 5	60X	180	153	85%
Test 6	60X	147	114	78%
Test 7	60X	125	108	86%
Test 8	60X	129	106	82%
Average		85%		
Standard deviation		5%		

Table S2: Error matrix for three different classification criteria. Each of the criteria uses a different number (n) of M-type profiles among the sixteen profiles to identify membrane-stained cells. The threshold for discriminating M- and C-type profiles was set at 0.9 for any criteria.

	Criteria I: n >= 1		*Criteria II:* n > = 2		*Criteria III:* n >= 3		*Criteria IV:* n >= 4	
	cytosolic	membrane	cytosolic	membrane	cytosolic	membrane	cytosolic	membrane
Cytosolic-stained cells	*true negatives* 312	*false positives:* 21	*true negatives:* 330	*false positives:* 3	*true negatives:* 332	*false positives:* 1	*true negatives:* 333	*false positives:* 0
Membrane-stained cells	*false negatives:* 42	*true positives:* 404	*false negatives:* 70	*true positives:* 376	*false negatives:* 89	*true positives:* 357	*false negatives:* 121	*true positives:* 325

Table S3: Sensitivity, specificity and purity of our cell classification. Sensitivity, specificity and expected maximal purity associated with the three different classification criteria. The values of purity (10/90) correspond with a sample with 10% of target cells and 90% non-target cells, and purity (25/75) with a sample with 25% of target cells and 75% of non-target cells.

	Sensitivity	Specificity	Purity (10/90)	Purity (25/75)
Criteria I: n >= 1	90.6%	93.7%	61.5%	82.7%
Criteria II: n > = 2	84.3%	99.1%	91.2%	96.9%
Criteria III: n >= 3	80.0%	99.7%	96.7%	98.9%
Criteria III: n >= 4	72.9%	100%	100%	100%

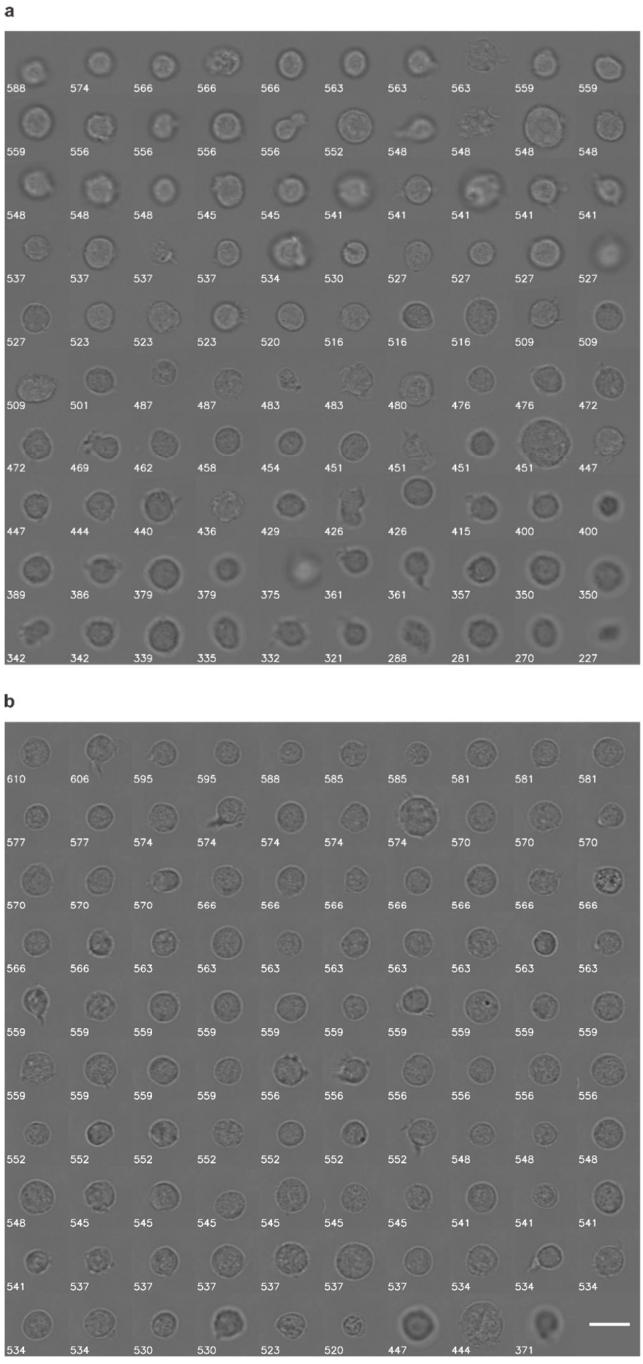

Fig. S3: Focal alignment and velocities of dielectrophoretically-guided cells. a, Exemplary images of cells flowing in the microchannel with inactive electrodes. Due to the absence of electric fields, the cells do not experience a force directed toward the centre of the channel and are therefore in different vertical positions outside the focal plane. **b**, Exemplary images of cells under the same experimental conditions but with active electrodes. Due to the interaction with the electric fields, the cells experience a vertical force and arrange in the centre of the channel in a common focal plane. The numbers in white represent the measured velocity of each cell in $\mu m\ s^{-1}$. Scale bar, 20 µm.

Table S4 (next page): Important characteristics of image-activated cell sorting approaches.
*theoretically maximum possible resolution at 600 nm wavelength based on the objective characteristics
**ratio between sample volume flow rate and flow velocity, describing how much sample volume can be processed per metre travelled distance of the cells.

Publication title	Objective	Optical resolution	Fluorescence imaging	Flow velocity	Sample volume flow rate	Flow efficiency**	Cell throughput	Purity	Recovery rate	Post sort vitality	Reference / DOI
Intelligent image-activated cell sorting	20x/0.75	400 nm*	yes	1000 mm/s	2500 µl/h	0.7 µl/m	92 eps	> 79.5%	not shown	not shown	10.1016/j.cell.2018.08.028
Intelligent image-activated cell sorting 2.0	20x/0.75	400 nm*	yes	1000 mm/s	2500 µl/h	0.7 µl/m	1133 eps	81.9%	not shown	not shown	10.1039/D0LC00080A
Raman image-activated cell sorting	20x/0.75	400 nm*	yes	40 mm/s	100 µl/h	0.7 µl/m	86 eps	81.0%	not shown	not shown	10.1038/s41467-020-17285-3
Low-latency label-free image-activated cell sorting	10x/0.28	1071 nm*	yes				200 eps	98.4%	not shown	not shown	10.1016/j.bios.2022.114865
User-friendly image-activated microfluidic cell sorting	10x/0.3	1000 nm*	not shown	106 mm/s	60 µl/h	0.2 µl/m	83 eps	95.1%	not shown	not shown	10.1039/d0lc00747a
High-speed fluorescence image–enabled cell sorting	104 px / 60 µm	1154 nm	yes	1100 mm/s			15,000 eps	96,0%	not shown	not shown	10.1126/science.abj3013
Intelligent image-based deformation-assisted cell sorting	20x/0.8	375 nm*	not shown	100 mm/s	36 µl/h	0.1 µl/m	100 eps	96.3%	not shown	yes	10.1038/s41592-020-0831-y
Image-based cell sorting using focused travelling surface acoustic waves	20x/0.8	375 nm*	not shown	89 mm/s	36 µl/h	0.1 µl/m	400 eps	94.1%	72%	yes	10.1039/D2LC00636G
Our approach	60x/1.42	216 nm	yes	0.5 mm/s	35 µl/h	20 µl/m	0.5-3 eps	72-95% (yield sort)	92%	yes	