Review of Scientific Instruments ARTICLE scitation.org/journal/rsi

Rapid and high-precision sizing of single particles using parallel suspended microchannel resonator arrays and deconvolution

Cite as: Rev. Sci. Instrum. 90, 085004 (2019); doi: 10.1063/1.5100861
Submitted: 21 April 2019 • Accepted: 12 August 2019 •
Published Online: 30 August 2019

View Online Export Citation CrossMark

Max A. Stockslager,[1] Selim Olcum,[2] Scott M. Knudsen,[2] Robert J. Kimmerling,[3] Nathan Cermak,[4] Kristofor R. Payer,[5] Vincent Agache,[6] and Scott R. Manalis[1,2,3,a)]

AFFILIATIONS

[1]Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA
[2]Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA
[3]Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA
[4]Program in Computational and Systems Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA
[5]Microsystems Technology Laboratories, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA
[6]CEA-LETI, Minatec Campus, Grenoble 38054, France

[a)]**Author to whom correspondence should be addressed:** srm@mit.edu

ABSTRACT

Measuring the size of micron-scale particles plays a central role in the biological sciences and in a wide range of industrial processes. A variety of size parameters, such as particle diameter, volume, and mass, can be measured using electrical and optical techniques. Suspended microchannel resonators (SMRs) are microfluidic devices that directly measure particle mass by detecting a shift in resonance frequency as particles flow through a resonating microcantilever beam. While these devices offer high precision for sizing particles by mass, throughput is fundamentally limited by the small dimensions of the resonator and the limited bandwidth with which changes in resonance frequency can be tracked. Here, we introduce two complementary technical advancements that vastly increase the throughput of SMRs. First, we describe a deconvolution-based approach for extracting mass measurements from resonance frequency data, which allows an SMR to accurately measure a particle's mass approximately 16-fold faster than previously possible, increasing throughput from 120 particles/min to 2000 particles/min for our devices. Second, we describe the design and operation of new devices containing up to 16 SMRs connected fluidically in parallel and operated simultaneously on the same chip, increasing throughput to approximately 6800 particles/min without significantly degrading precision. Finally, we estimate that future systems designed to combine both of these techniques could increase throughput by nearly 200-fold compared to previously described SMR devices, with throughput potentially as high as 24 000 particles/min. We envision that increasing the throughput of SMRs will broaden the range of applications for which mass-based particle sizing can be employed.

Published under license by AIP Publishing. https://doi.org/10.1063/1.5100861

INTRODUCTION

Measuring the size distributions of nanoscale and microscale particles is of central importance for a wide range of scientific and industrial applications. In the physical sciences, applications of particle sizing have included quantifying the aggregation of protein drug products[1] and measuring the volume fraction and size dispersity of colloidal suspensions.[2] In biology, cell size is fundamentally linked to the cell state, and particle sizing tools have been employed both to investigate basic questions about how individual cells regulate their size and growth[3] and for practical applications such as evaluating *ex vivo* the susceptibility of primary patient tissue to cancer therapeutics.[4]

There are several single-particle approaches that are routinely used for sizing particle suspensions in the 5–20 μm size range. Resistive-pulse sensing instruments (such as the Coulter counter) achieve volumetric precision on the order of 1%–10% in this size range, with throughput of thousands of cells per minute

due to the continuous flow-through nature of the measurement.[5,6] Transmitted- and reflected-light microscopy can measure particle diameter and infer particle volume with precision similar to the Coulter counter, provided that the particles are approximately spherical; commercial "microflow imaging" systems perform these measurements on flowing particles at rates as high as tens of thousands of particles per minute.[7] To size nonspherical particles and cells, other imaging modalities such as quantitative phase microscopy have been used.[8]

In contrast with these techniques, which size particles by electrical or optical methods, suspended microchannel resonators (SMRs) directly measure particle buoyant mass by detecting a shift in resonance frequency as particles flow through a vacuum-packaged cantilever beam containing a U-shaped microfluidic channel [Fig. 1(a)].[9] In the 5–20 μm size range, SMRs can achieve mass precision on the order of 0.1%–1%.[10] However, to date, the throughput of SMRs has been limited to tens of particles per minute, restricting their use to applications requiring highly precise measurements of relatively small numbers of particles.

Here, we present two independent technical advancements that increase SMR throughput without degrading precision.

The first is a method for obtaining accurate and precise mass measurements from particles flowing through the resonator with transit times up to ~16 times faster than previously possible. In normal operation, a phase-locked loop (PLL) is used to keep the SMR vibrating at its resonance frequency by forming a resonance-tracking SMR-PLL feedback loop[11] (supplementary material, Fig. 1). This resonance-tracking loop can be configured with arbitrary response speed, but due to the fundamental trade-off between frequency noise and bandwidth, diminishing signal-to-noise ratio sets an upper limit on the achievable resonance-tracking bandwidth. To surpass this throughput limitation, we developed a model-based deconvolution algorithm that extracts mass measurements from measured resonance frequency signals that were created by particles flowing too fast to be fully resolved by the resonance-tracking loop [Fig. 1(b)]. Intuitively, this operation corresponds to "deblurring" of the measured resonance frequency signal, which was "blurred" due to the limited bandwidth of the resonance-tracking loop, analogous to the use of deconvolution in microscopy to deblur images that were blurred due to diffraction. Since the deconvolution operation is performed in postprocessing, this approach can be immediately applied to existing SMR systems that are read out using SMR-PLL resonance-tracking loops, without any need for additional hardware modifications.

The second advancement is a microfluidic device containing 16 SMRs connected fluidically in parallel and operated simultaneously [Fig. 1(c)]. Using techniques described previously,[10,11] shifts in the resonance frequency of each cantilever can be tracked

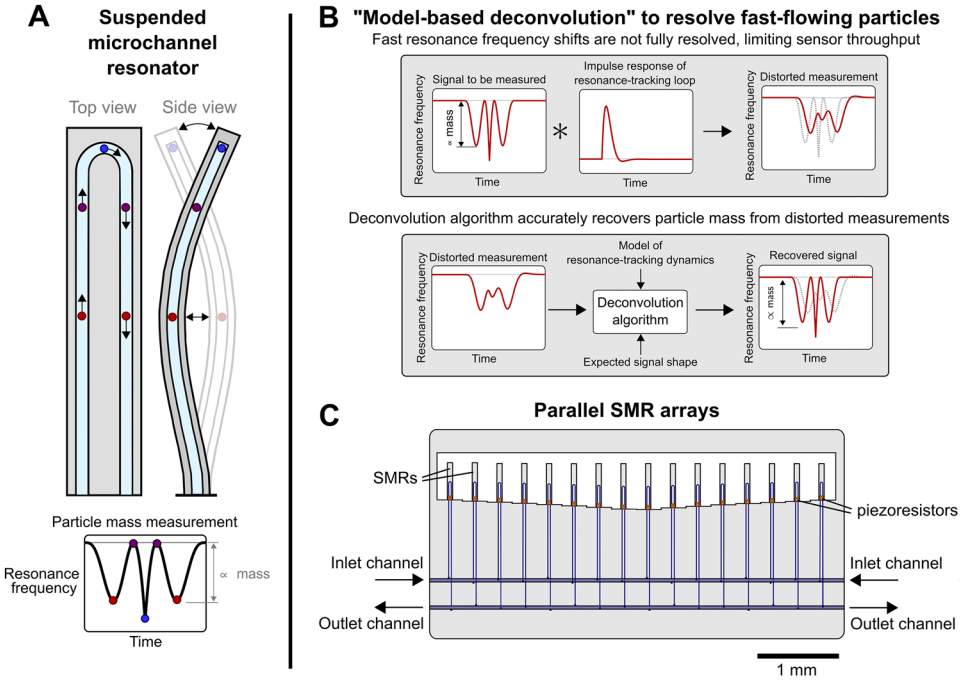


FIG. 1. Two methods for increasing the throughput of suspended microchannel resonators. (a) Suspended microchannel resonator (SMR) operating principle. Particle mass is measured by detecting a position-dependent shift in resonance frequency as particles flow through the vacuum-packaged cantilever beam with an embedded microfluidic channel. The cantilever is shown operated in the second vibrational mode, resulting in three local minima in resonance frequency as a particle flows to the tip of the cantilever and back. (b) A model-based deconvolution algorithm increases the maximum particle speed for which accurate mass measurements can be obtained. Limited resonance-tracking bandwidth constrains the maximum throughput of SMRs since fast resonance frequency shifts cannot be fully resolved by the sensor, resulting in distortion of the measured resonance frequency signal (top). A deconvolution-based algorithm that uses knowledge of the expected resonance frequency signal is used to "deblur" the distorted resonance frequency signals and recover particle mass measurements (bottom). (c) Schematic of parallel SMR array devices, which contain 16 SMRs connected fluidically in parallel and operated simultaneously to further increase throughput.

Published under license by AIP Publishing

independently, and frequency-multiplexing allows each cantilever to be continuously driven at resonance using a single actuation channel and single detection channel. We evaluate the precision of the parallel SMR arrays by measuring suspensions of monodisperse polystyrene beads, obtaining coefficients of variation up to ~4 times lower than a commercial Coulter counter configured for a similar size range.

RESULTS

Model-based deconvolution

The throughput of SMRs is fundamentally limited by the temporal resolution with which changes in resonance frequency can be tracked. To increase volumetric throughput, particles must transit the cantilever more quickly, and so wider resonance-tracking bandwidths are required to fully resolve the position-dependent shift in resonance frequency caused by the added mass of the moving particle. Specifically, for a particle flowing at constant speed through a cantilever driven in the second bending mode, the resonance-tracking loop must have bandwidth (in Hz) at least 24 times greater than the characteristic frequency $1/T_{transit}$ in order for the transient resonance frequency shift to be fully resolved with >99.9% energy recovery (a criterion we have used previously[11]).

When this bandwidth requirement is not met, the measured resonance frequency signal becomes distorted by the limited bandwidth of the resonance-tracking loop. The degree of distortion depends on the specific resonance-tracking transfer function, mostly by the relative values of the loop bandwidth and transit time. Configuring the resonance-tracking loop with a wider bandwidth enables the measurement of faster particles without distortion of the resonance frequency signal but at the expense of widening the noise bandwidth (supplementary material, Fig. 1).

In order to increase the maximum particle speed that can be measured at a particular resonance-tracking bandwidth (and therefore with constant measurement noise), we explored whether one could computationally analyze distorted resonance frequency signals to estimate the true, fully resolved resonance frequency signal, which encodes the particle's mass. Since the distorted resonance frequency measurement is mathematically the convolution of the true resonance frequency signal with the closed-loop impulse response of the resonance-tracking loop, one could potentially deconvolve the known resonance-tracking impulse response from the distorted measurement to arrive at an estimate of the original resonance frequency shift signal. There exist various computational approaches for performing this deconvolution operation, including classical Tikhonov-Wiener approaches and more modern statistical methods.[12] However, this application differs from the classical deconvolution problem since more information is available; the shape of the true resonance frequency signal is known, while only its amplitude and duration are unknown.

We developed a deconvolution-based signal recovery algorithm (Fig. 2) that uses knowledge of the expected shape of the true resonance frequency signal to estimate a particle's mass and transit time by comparing the measured peak shape against a precomputed library of distorted peak shapes. The library of distorted peak shapes was generated by convolving theoretical (nondistorted) peak shapes of varying transit time with the impulse response of the measured resonance-tracking transfer function to predict the distorted peak

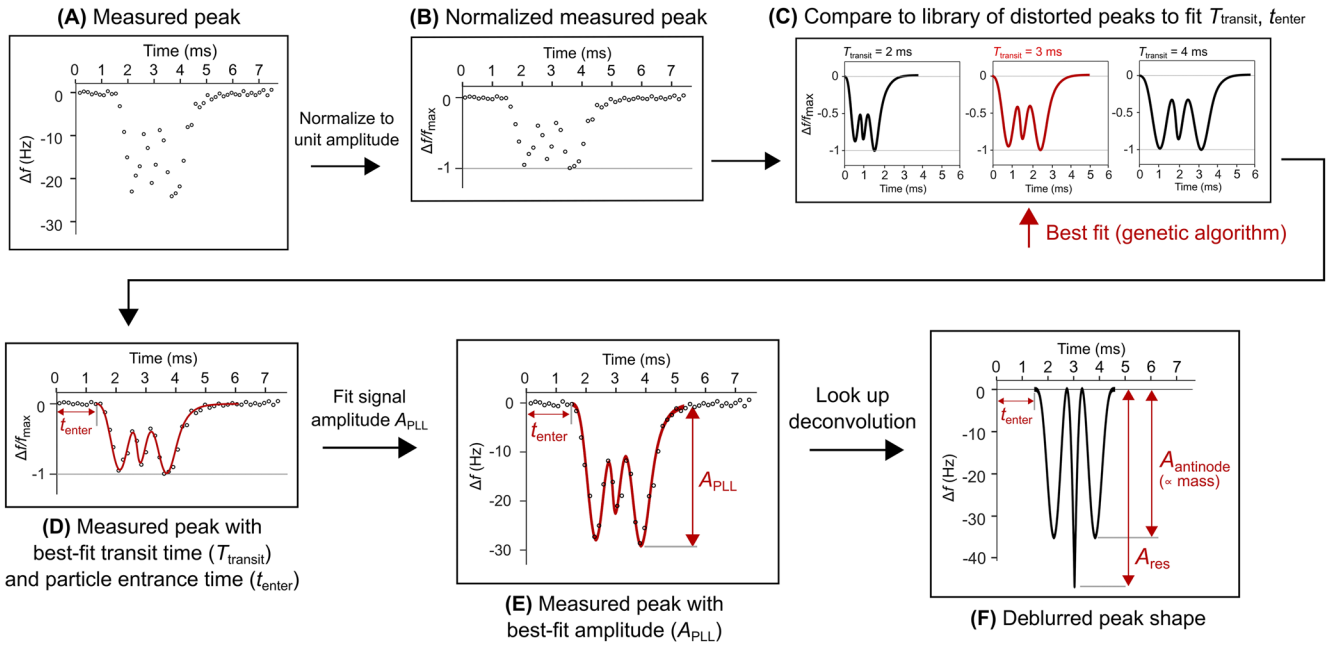


FIG. 2. Model-based deconvolution algorithm. [(a) and (b)] Measured signals are normalized such that the maximum deviation from the baseline has unit amplitude; [(c) and (d)] normalized signals are compared against a precomputed library of distorted peak shapes to estimate the particle's transit time ($T_{transit}$) and cantilever entrance time (t_{enter}); (e) the signal amplitude is fit to minimize deviation between the (scaled) fit peak shape and the measured signal; and (f) a precomputed lookup table of deconvolved peak shapes is used to recover an estimate of the deblurred signal, for which the peak resonance frequency shift at the antinode is proportional to particle mass.

Published under license by AIP Publishing

shapes that would be observed for particles with a particular transit time. Measured distorted peak shapes are compared against this library, and an optimization procedure is used to fit three parameters to the measurement: the time at which the particle entered the cantilever, its transit time, and the amplitude of the distorted peak. Modeling the distorted measurement as the convolution of the theoretical peak shape with the resonance-tracking impulse response, these three parameters (the particle's entrance time, transit time, and signal amplitude) are sufficient to fully recover an estimate of the true resonance frequency signal, which encodes the particle's mass.

Details of deconvolution algorithm

More specifically, the "model-based deconvolution" algorithm works as follows. Generating a library of blurred peak shapes first requires knowledge of the theoretical peak shape. The position-dependent resonance frequency shift that occurs when a particle flows through the resonating cantilever, normalized to unit maximum amplitude, is given by

$$f_{SMR}(t, T_{transit}) = \frac{-u_n^2(z(t))}{(u_n^2)_{\max}}, \tag{1}$$

where $u_n(z)$ is the cantilever deflection at position z (normalized such that $z = 0$ at the cantilever base and $z = 1$ at the cantilever tip) when driven in mode n (here, $n = 2$) and $\left(u_n^2\right)_{\text{antinode}}$ is the maximum squared deflection of the cantilever. For the particle trajectory $z(t)$, we approximate the particle as moving at a constant speed to the end of the fluidic channel (96.5% of the total cantilever length for our devices) and then instantly turning and returning to the base of the cantilever at the same speed.

The PLL frequency signal, i.e., the time-varying estimate of the resonance frequency generated by a phase-locked loop in feedback with an SMR, is modeled as the convolution of this theoretical resonance frequency signal with the (user-specified) resonance-tracking impulse response given by h_{SMR},

$$f_{PLL}(t, T_{transit}, h_{track}) = f_{SMR}(t, T_{transit}) * h_{SMR}(t), \tag{2}$$

i.e., the theoretical resonance frequency signal is distorted, or "blurred," by the resonance-tracking impulse response.

Any distorted resonance frequency signal (and its corresponding deblurred resonance frequency signal) can be uniquely specified by three parameters: the time at which the particle enters the cantilever (t_{enter}), the particle's transit time through the cantilever ($T_{transit}$), and the signal amplitude (A_{PLL} and A_{res} for the distorted and deblurred signals, respectively). Therefore, assuming the particle travels at a constant speed and that the SMR-PLL impulse response is known accurately, determining these three parameters (t_{enter}, $T_{transit}$, and A_{PLL}) from the distorted signal is sufficient to recover an estimate of the deblurred resonance frequency signal.

We determine these three parameters (t_{enter}, $T_{transit}$, and A_{PLL}) from distorted peak measurements as follows.

First, we precompute a library of distorted peak shapes (normalized to unit maximum resonance frequency shift) as a function of particle transit time using the measured resonance-tracking transfer function. The library contains precomputed peak shapes for the full range of transit times expected to be observed in the data—in our case, 1–100 ms, in steps of 0.1 ms.

Second, we fit an entrance time (t_{enter}) and transit time ($T_{transit}$) to the measured peak shape by normalizing it to unit maximum amplitude and searching the library of distorted peak shapes for the best fit, minimizing the following objective function:

$$\chi^2(t_{enter}, T_{transit}) = \sum_i \left[\frac{f_{measured}(t_i)}{|f_{measured}|_{max}} - \frac{f_{library}(t_i - t_{enter}, T_{transit}, h_{track})}{|f_{library}|_{max}} \right]^2, \tag{3}$$

where the summation index runs over the length of the measured signal. Since this objective function contains many saddle points and local minima, we found that we achieved the best performance using a genetic algorithm for the optimization (the MATLAB implementation *ga*). We evaluated the performance of the genetic algorithm using simulations and found that it correctly estimated the transit time (within ±1 ms of the true value for a 20 ms peak) in approximately 99% of cases (supplementary material, Fig. 2). Using this approach, the transit time fit to a particular peak is limited to the finite set of transit times for which peak shapes have been precomputed and stored in the library of peak shapes; however, one can fit the transit time with arbitrary precision by including more transit times in the peak shape library. Typically, we found 0.1 ms resolution to be more than sufficient.

Third, after fitting an entrance time and transit time to the normalized measured peak shape, we fit the distorted signal amplitude (A_{PLL}) to the non-normalized measured peak shape by minimizing the following objective function:

$$\chi^2(A_{PLL}) = \sum_i \left[f_{measured}(t_i) - A_{PLL} \frac{f_{library}(t_i - t_{enter}, T_{transit}, h_{track})}{|f_{library}|_{max}} \right]^2 \tag{4}$$

using the MATLAB nonlinear programming solver *fmincon*.

Finally, after fitting an entrance time, transit time, and signal amplitude to the distorted peak, we use the precomputed library of distorted peaks to look up the deblurred peak shape corresponding to the observed distorted peak shape, yielding a recovered estimate of the true resonance frequency signal, for which the resonance frequency shift at the antinode is proportional to the particle's mass. Specifically, we implement this procedure in two steps. First, we use the peak library to look up the maximum amplitude of the deblurred signal (A_{res}) as a function of transit time and the maximum amplitude of the distorted signal (A_{PLL}). However, for SMRs operated in the second vibrational mode, the resonance frequency shift at the *antinode* ($A_{antinode}$) is proportional to the particle's mass. Fortunately, $A_{antinode}$ is related to the maximum amplitude A_{res} by a constant multiplicative factor, which depends only on the length of the fluidic channel relative to the cantilever ($A_{antinode} = 0.809A_{res}$ for our devices, for which the fluidic channel extends 96.5% of the length of the cantilever). Scaling A_{res} by this factor results in an estimate of the resonance frequency shift when the particle is at the antinode, which is directly proportional to particle mass, regardless of the length of the fluidic channel relative to the cantilever.

Applying this deconvolution algorithm is computationally more complex than the simple case where the resonance frequency signal is fully resolved and particle mass can be estimated simply by

Published under license by AIP Publishing

measuring the peak amplitude. Typically, the rate-limiting step for deconvolution is applying the genetic algorithm to fit the distorted peak's transit time and entrance time. However, in our experience, our implementation of the algorithm running on a midrange desktop computer can complete this operation in 50–100 ms per peak, which has been sufficient for current applications. In the future, optimizing the implementation of the algorithm may be beneficial if computation becomes a rate-limiting step for data collection.

Linear resonance-tracking model

This approach relies on a linear model of the resonance-tracking dynamics of the SMR-PLL loop, i.e., the observed distorted peak shapes are accurately modeled as a convolution of the theoretical peak shape with the resonance-tracking impulse response. To test this assumption experimentally, we flowed a sample of nominal 1.1 μm polystyrene particles (Thermo Scientific 4000 Series) at multiple speeds (transit times of approximately 5, 10, and 20 ms) through an SMR with channel dimensions 3 × 5 × 120 μm^3, while the resonance-tracking loop was configured with a first-order transfer function with one of several bandwidths (100, 300, or 500 Hz). We observed good agreement between the measured peak shapes and the blurred peak shapes fit by the deconvolution algorithm across the grid of transit times and loop bandwidths [Figs. 3(a) and 3(b)], suggesting that a linear model accurately predicts the distortion of fast resonance frequency shifts by the SMR-PLL resonance-tracking loop.

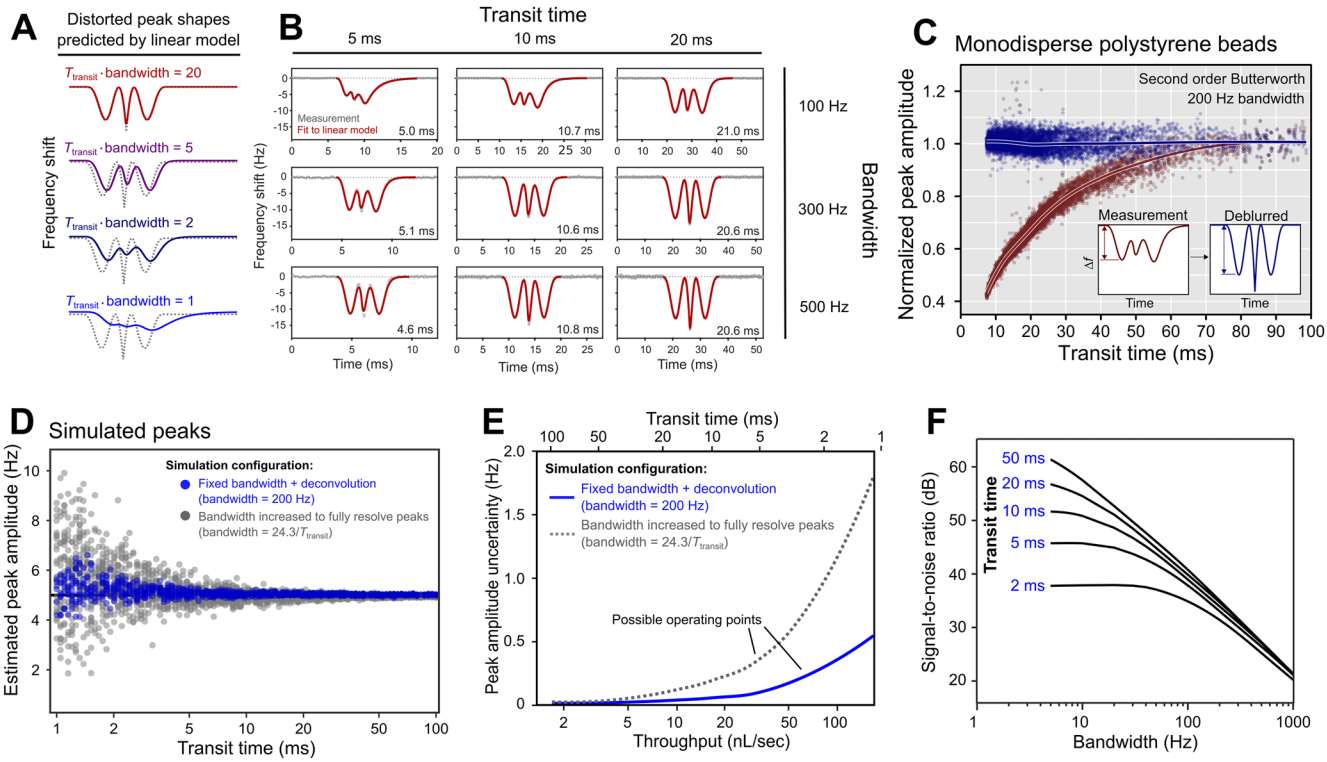


FIG. 3. Characterizing performance of the model-based deconvolution algorithm. (a) Distorted peak shapes predicted by a linear model of the SMR-PLL resonance-tracking loop. Distorted peak shapes (solid lines) are modeled as the convolution of the theoretical peak shape (dashed lines) with the nominal impulse response of the resonance-tracking loop. The degree of distortion depends on the specific resonance-tracking transfer function, particularly on the dimensionless transit time ($T_{transit}$ × bandwidth). (b) Measured peak shapes (gray points) and fits to the linear resonance-tracking model (red line) for 1.1 μm polystyrene beads measured using an SMR with 3 × 5 × 120 μm^3 channel dimensions, as a function of resonance-tracking bandwidth and particle transit time. Deconvolution was performed by comparing measured peak signals against a library of distorted peak shapes to fit three parameters: entrance time, transit time, and signal amplitude. Good agreement was observed, indicating that the linear model accurately predicts the observed distorted peak shapes. (c) Measured peak amplitudes (amplitude of the first antinode peak; red) and recovered peak amplitudes (blue) for 1.1 μm polystyrene beads measured across a range of transit times, with LOESS (locally estimated scatterplot smoothing) fits overlaid as solid lines. The model-based deconvolution algorithm accurately recovered peak amplitude (and therefore particle mass) independent of transit time. (d) Peak measurements simulated by generating theoretical peak shapes with 5 Hz amplitude and varying transit time, convolving with a specified resonance-tracking transfer function to simulate distortion and then corrupted with additive white noise. At a particular transit time, measuring the peaks with a fixed, narrow bandwidth (200 Hz) followed by model-based deconvolution (blue points) results in more precise peak amplitude estimates than the alternative approach of increasing the bandwidth to fully resolve the peak signal (gray points). Noise spectrum modeled as white, with σ = 0.25 Hz at 200 Hz bandwidth. (e) Peak amplitude uncertainty as a function of transit time (i.e., standard deviation of peak height estimates at a particular transit time) for the simulations in (e). At faster transit times, the deconvolution algorithm (while measuring peaks at a fixed 200 Hz bandwidth) provides more precise peak amplitude estimates than simply increasing the resonance-tracking bandwidth to fully resolve the signal. Throughput estimates are for a cantilever with channel dimensions 3 × 5 × 120 μm^3. (f) For particles flowing through the cantilever with a particular transit time, the signal-to-noise ratio (signal energy/measured noise power) increases when the resulting resonance frequency signal is detected using a narrower-bandwidth resonance-tracking loop.

Published under license by AIP Publishing

Accuracy and precision

After confirming that a linear model accurately predicts distorted peak shapes, we next asked whether applying the model-based deconvolution algorithm to measured distorted peak shapes could accurately estimate particle mass. Although the true, nondistorted resonance frequency signal is not available for measurement, the performance of this approach can be evaluated by measuring particles of the same size at different speeds and then confirming that after the deconvolution algorithm has been applied, the estimated mass does not depend on the particle's transit time. We configured the resonance-tracking loop with a first-order transfer function of 200 Hz bandwidth and then flowed monodisperse 1.1 μm polystyrene particles (nominal volume coefficient of variation 3.3%) through the 3 × 5 × 120 μm^3 cantilever with transit times ranging from 5 to 100 ms [N = 7685 particles; Fig. 3(c)]. As expected, the amplitude of the measured peak shapes varied with transit time due to distortion, even though all particles had approximately the same mass. In particular, peak amplitude was attenuated for particles with transit times faster than approximately 90 ms, with more attenuation for faster-flowing particles [Fig. 3(c); red]. However, applying the deconvolution algorithm to these peaks accurately recovered the true peak amplitude; the recovered peak amplitude was approximately independent of transit time [Fig. 3(c); blue], confirming that the deconvolution algorithm accurately recovered the mass of particles with transit times as low as 7.5 ms. As a measure of the bias introduced by the deconvolution algorithm, the maximum absolute deviation of the locally estimated scatterplot smoothing (LOESS) fit (blue, solid line) from the true mean peak amplitude was 2.5%. However, these measurements only provide an upper bound on the uncertainty introduced by the deconvolution algorithm; since the measured size dispersity is approximately constant across the range of transit times, the degradation of precision introduced by deconvolution is less than the true size dispersity of the particles, which have a nominal volume coefficient of variation of 3.3%.

Next, we turned to simulations to better quantify the effect of the signal recovery algorithm on measurement precision. We simulated noisy, distorted peak shapes with 5 Hz amplitude and transit times ranging from 1 to 100 ms. The distorted peaks were simulated by convolving the theoretical peak shape with a resonance-tracking impulse response of interest (here, a 200 Hz second-order Butterworth impulse response), downsampling this distorted resonance frequency signal to 12.5 kHz to simulate measurement and data acquisition through the digital SMR-PLL loop, then adding white noise (σ = 0.25 Hz at 200 Hz bandwidth). We then applied the deconvolution algorithm to each of these simulated distorted, noisy peak signals and recorded the amplitudes of the recovered "deblurred" peaks [Fig. 3(d)]. As expected, the resulting peak amplitude estimates were symmetrically distributed about the true value of 5 Hz, with larger uncertainty for faster-flowing particles.

We compared the performance of the deconvolution algorithm to the alternative approach of simply widening the resonance-tracking bandwidth to fully resolve each particle's resonance frequency signal. To maximize precision in this measurement configuration, the resonance-tracking bandwidth should be as narrow as possible while still fully resolving the signal; for a first-order loop, the minimum bandwidth is approximately $24.3/T_{transit}$ to ensure >99.9% energy recovery. To evaluate the precision of this approach, we again simulated peak measurements of 5 Hz amplitude with transit times ranging from 1 to 100 ms but with the loop bandwidth for each peak set to $24.3/T_{transit}$ to ensure that all peak shapes were fully resolved. Each peak was corrupted with white noise appropriate to the loop bandwidth; we assumed the spectrum was white with power increasing proportional to resonance-tracking bandwidth, with σ = 0.25 Hz at 200 Hz bandwidth.

Across the entire range of transit times (1–100 ms), the deconvolution algorithm achieved better precision than the alternative approach of widening the loop bandwidth to fully resolve the peak shape [Fig. 3(e)]; for 1 ms peaks, the deconvolution-based approach improved precision by more than three-fold with this configuration and noise spectrum. In summary, our deconvolution approach allows accurate, precise mass measurements to be made on particles flowing up to 16 times faster than previously possible at a particular resonance-tracking bandwidth and improves precision beyond the alternative approach of increasing resonance-tracking bandwidth to fully resolve particles flowing at these higher speeds. Table I summarizes the throughput achieved by this approach, compared to existing devices configured with the same resonance-tracking bandwidth (500 Hz) and measuring a sample with the same concentration (one particle per 10 cantilever volumes).

Optimal bandwidth

After determining that the deconvolution algorithm allows particles to be measured up to 16× faster at a particular resonance-tracking bandwidth, we next asked how the resonance-tracking bandwidth should be selected to optimize the trade-off between signal tracking and noise rejection for particles flowing at a particular speed. In particular, narrower loop bandwidths (of order $1/T_{transit}$) result in lower noise at the expense of more significant signal distortion, while wider loop bandwidths ($\gg 1/T_{transit}$) result in improved

TABLE I. SMR throughput comparison, for cantilevers of channel dimensions 12 × 20 × 350 μm^3.

System	Minimum transit time (ms)	Maximum throughput (particles/min)	Fold improvement
Single SMR	49	120	1×
Single SMR, deconvolution	3	2 000	16×
Parallel SMR array	10.6	6 800	55×
Parallel SMR array, deconvolution (estimated)	3	24 000	197×

Published under license by AIP Publishing

temporal resolution at the expense of increased frequency noise. To quantify the trade-off between these two objectives, we calculated the signal-to-noise ratio of measured peaks as a function of resonance-tracking bandwidth. First, we quantified signal tracking as a function of resonance-tracking loop bandwidth by calculating the recovered energy of the blurred signal as a function of resonance-tracking bandwidth and transit time. [supplementary material, Figs. 3(a) and 3(b)]; greater energy recovery corresponds to better tracking of the resonance frequency signal. Next, we quantified frequency noise as a function of bandwidth by recording noise samples from an SMR configured with a range of resonance-tracking bandwidths, and from these, noise samples calculated total noise power as a function of bandwidth [supplementary material, Fig. 3(c)] Taking the ratio of signal energy to noise power, we observed that for our devices, the signal-to-noise ratio decreases monotonically with bandwidth for particles with a particular transit time, i.e., using narrower loop bandwidths increases the signal-to-noise ratio, with diminishing returns at very narrow loop bandwidths [Fig. 3(f)]. This result provides additional support for the notion that when measuring particles with a particular transit time, better signal-to-noise ratios are obtained by configuring the resonance-tracking loop with a narrow bandwidth and applying the deconvolution algorithm, rather than simply widening the bandwidth to fully resolve the signal. The ideal resonance-tracking transfer function has as narrow a bandwidth as possible while still tracking the signal well enough for the deconvolution algorithm to succeed. We have found via simulations that the algorithm begins to fail (by beginning to fit incorrect peak shapes) when the dimensionless transit time $T_{transit}$ × bandwidth (Hz) falls below approximately 1.5.

Parallel SMR arrays

Next, we explored to what extent throughput could be increased by operating multiple SMRs simultaneously on the same microfluidic chip. Previously, it has been shown that our detection approach can be used to track multiple resonances simultaneously, whether those resonances are multiple vibrational modes of a single cantilever[11] or multiple cantilevers with different resonance frequencies.[10] Due to the high quality factors of the resonators, a single actuation channel and a single detection channel summing the piezoresistor currents can be used to drive multiple SMRs at resonance simultaneously with minimal cross talk.

We designed devices containing 16 SMRs connected fluidically in parallel between two large bypass channels [Fig. 4(a); device shown in Fig. 1(c)]. Particles are loaded into the device through the first bypass channel, pass through one of the 16 cantilevers, and then are flushed off the chip in the opposite bypass channel. Cantilever deflections are read out via embedded piezoresistors, as described previously.[13] The fluidic channels embedded in the cantilevers have a cross-sectional area of 12 × 20 μm^2 and the cantilever lengths vary from 461 to 573 μm, resulting in second-mode resonance frequencies in the range of 600 kHz–1 MHz, with typical quality factors in the range of 2000–4000 [Fig. 4(b)]. Of note, these devices are larger (and therefore less sensitive) than the 3 × 5 × 120 μm^3 channels described above. We chose to operate either 10 or 12 of the16

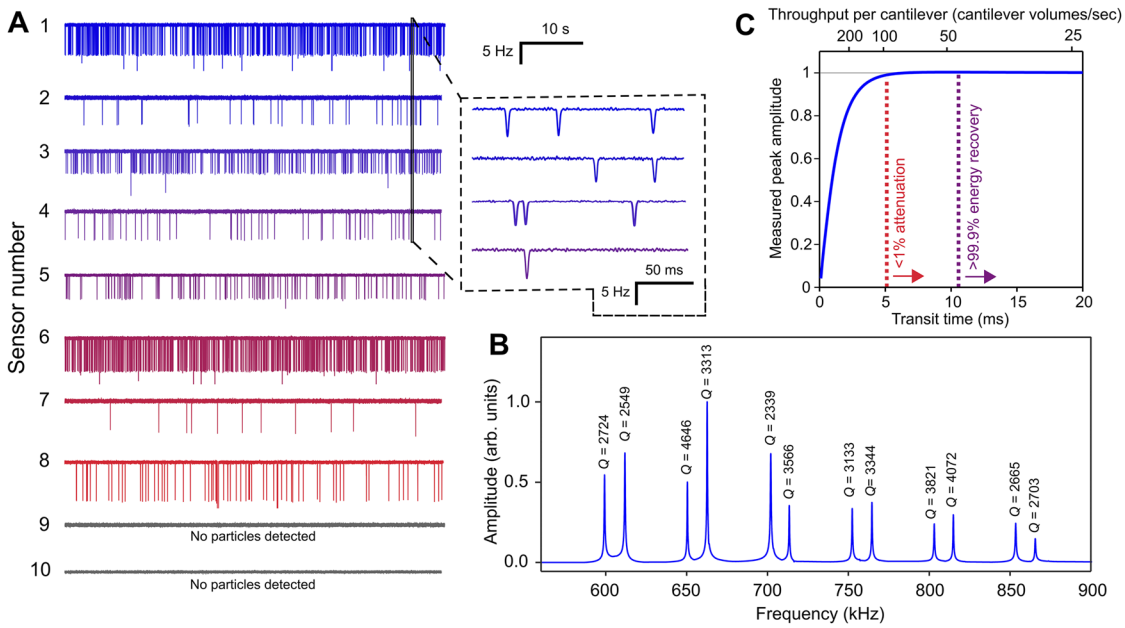


FIG. 4. Parallel SMR arrays. (a) Resonance frequency signals from 10 SMRs operated in parallel on the same chip (out of a maximum of 12). Plots show a sixty-second sample of the resonance frequency signal from each cantilever, measuring a suspension of 7 μm polystyrene beads in phosphate-buffered saline. For this particular device, we chose to operate only 10/12 cantilevers due to coupling between SMR-PLL loops, which we occasionally observe. (b) Open-loop transfer function amplitude for the parallel SMR array. Resonators are multiplexed in the frequency domain to enable independent control. Quality factors ranged from 2549 to 4646. (c) With the specified resonance-tracking transfer function (second-order Butterworth, 500 Hz bandwidth), particles can be weighed with cantilever transit times as low as 10.6 ms while still fully resolving the resonance frequency signal (>99.9% energy recovery) or with transit times as low as 5.1 ms before the measured peak amplitude becomes attenuated by 1%.

Published under license by AIP Publishing

sensors simultaneously to maintain compatibility with existing readout electronics. We confirmed that each SMR-PLL resonance-tracking loop was configured with the desired dynamics by directly measuring the resonance-tracking impulse response for each SMR-PLL loop [supplementary material, Fig. 4(a)] and, from this, derived the corresponding resonance-tracking transfer functions [supplementary material, Fig. 4(b)].

Pressure-driven flow is used to load particles from both ends of the bypass channel simultaneously for measurement. This sample-loading approach results in similar numbers of particles being loaded into each cantilever; however, we have observed slight day-to-day drift in the fraction of particles entering each sensor, even when the device is configured with nominally the same pressure settings [supplementary material, Fig. 4(c)].

We designed the fluidic channels inside the cantilevers to extend only to the antinode when driven in the second vibrational mode (~48% of the cantilever length), which results in an altered transit resonance frequency signal with a single peak, compared to the signal with three peaks when the fluidic channel extends to the tip of the cantilever[14] (supplementary material, Fig. 5). This design has several advantages. First, the transient resonance frequency signal can be fully resolved (e.g., >99.9% energy recovery, with no need for deblurring) for particles flowing at ~2.3 times higher speeds since the spectral energy of the modified peak shape is concentrated at lower frequencies. Second, since the fluid volume inside each cantilever is reduced, the system can measure more concentrated particle suspensions (by a factor of ~2.0) without instances of two particles occupying a cantilever at the same time. Finally, compared to operating SMRs in the first mode, this approach eliminates position-dependent error resulting from variation in the trajectory a particle takes when flowing through the cantilever[14] (supplementary material, Fig. 5). However, as a trade-off, we found that the modified peak shapes generated by these cantilevers cannot be reliably deblurred by the model-based deconvolution algorithm since the peak shapes do not contain enough features to reliably determine the particle's transit time and fit a distorted peak shape.

Throughput

We evaluated the maximum throughput that could be achieved using the parallel SMR arrays while still precisely measuring particle masses. Our standard criterion is to require that the peak shape be fully resolved (i.e., >99.9% energy recovery); for a resonance-tracking loop with second-order Butterworth dynamics, this requires that the dimensionless transit time ($T_{\text{transit}} \times$ bandwidth) be least 5.3. At 500 Hz bandwidth, this corresponds to a minimum transit time of 10.6 ms and a maximum throughput of 8.1 μl/min through 12 cantilevers, representing a 27-fold improvement in volumetric flow rate over a single SMR with the same resonance-tracking transfer function (Table I). As an additional criterion for determining the maximum throughput, we required that the measured peak amplitude should be attenuated by no more than (say) 1% from the true value [Fig. 4(c)]; however, this constraint was less restrictive, only requiring that the dimensionless transit time be greater than 2.6. To summarize, the designed parallel SMR arrays increase volumetric throughput by a factor of up to 27 compared to a single SMR operating with the same resonance-tracking bandwidth.

Although the peak shapes from these devices are not compatible with the model-based deconvolution algorithm, future device designs could include fluidic channels extending the full length of the cantilever, enabling the use of the deconvolution approach for even further increases in throughput. Based on the algorithm's approach for single-cantilever SMR devices, i.e., accurately recovering signal amplitude for dimensionless transit times $T_{\text{transit}} \times$ bandwidth as low as 1.5, we estimate that future parallel SMR arrays could achieve throughput as high as 59 μl/min through 12 cantilevers, a 197-fold improvement in volumetric throughput over previously described single-cantilever SMRs and a further 7-fold improvement over the current generation of parallel SMR arrays (Table I).

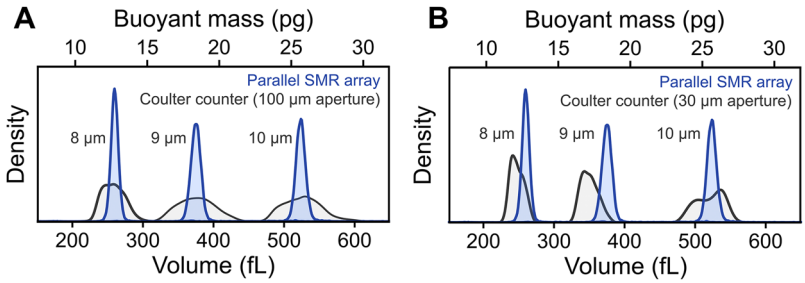
Regardless of the flow rate, the maximum allowable concentration of the particle suspension is limited by the restriction that only one particle can occupy each cantilever at a time. We modeled the probability of double-occupancy events—instances where two particles occupy a cantilever at the same time—as a function of sample concentration, modeling particle loading as a Poisson process (supplementary material, Note 1). In particular, limiting the likelihood of double-occupancy events to <10% requires that the average sample concentration be less than one particle per ~10 cantilever volumes. For the parallel SMR arrays, this requirement limits the maximum particle concentration to approximately 850 particles/μl. At this concentration, the current version of the parallel SMR array achieves a maximum throughput of ~6800 particles/min through 12 sensors, a 55-fold improvement compared to a single SMR operating with the same resonance-tracking bandwidth (Table I). Furthermore, for the future parallel SMR arrays compatible with deconvolution, this maximum concentration corresponds to an estimated maximum throughput of 24 000 particles/min.

Particle-sizing precision

We compared the precision of the parallel SMR array to that of a high-end commercial Coulter counter (Multisizer 4, Beckman Coulter). Specifically, we compared the instruments' ability to resolve the size distributions of monodisperse 8–10 μm diameter NIST-traceable polystyrene particles (Thermo Scientific 4000 Series). The Multisizer 4 was configured either with a standard 100 μm aperture [Fig. 5(a)] or a high-sensitivity 30 μm aperture [Fig. 5(b)]. The parallel SMR array was configured with second-order Butterworth dynamics with 500 Hz bandwidth. We measured at least 1000 particles of each size using each instrument and calculated the robust coefficient of variation (0.741 × interquartile range/median) as a measure of precision.

The 8–10 μm diameter particles have manufacturer-reported volume coefficients of variation (CVs) of 3.3%, 3.0%, and 2.7%, respectively, as determined by microscopy [Fig. 5(c)]. The parallel SMR array resolved narrower size distributions for all particle sizes (2.0%, 1.8%, and 1.3% for the 8–10 μm sizes, respectively), suggesting that it achieves higher relative precision than the approach used by the manufacturer. The Coulter counter measured slightly broader size distributions than the SMR when configured with the high-sensitivity 30 μm aperture (4.0%, 3.3%, and 4.1%) and even wider particle size distributions with the standard 100 μm aperture (6.5%, 7.4%, and 5.0%). Still, the Coulter counter has the advantage of achieving throughput still an order of magnitude greater than the current implementation of the parallel SMR array, on the order of

Published under license by AIP Publishing



C **Measured coefficient of variation**

Particle diameter:	8 μm	9 μm	10 μm
Parallel SMR array	**2.0%**	**1.8%**	**1.3%**
Coulter counter, 30 μm aperture	4.0%	3.3%	4.1%
Coulter counter, 100 μm aperture	6.5%	7.4%	5.0%
Manufacturer-reported (microscopy)	3.3%	3.0%	2.7%

FIG. 5. Precision of parallel SMR array vs Multisizer 4 Coulter counter. [(a) and (b)] Distribution of particle sizes for nominal 8–10 μm diameter polystyrene beads measured with the parallel SMR array (blue) vs a Multisizer 4 Coulter counter configured with either (a) a standard 100 μm aperture or (b) a high-sensitivity 30 μm aperture. $N > 1000$ particles measured for all conditions. (c) Robust coefficients of variation (0.741 × interquartile range/median) for each instrument.

tens of thousands of particles per minute. Selecting the appropriate instrument still depends on the desired trade-off between throughput and precision, and whether measuring particle mass or volume happens to be preferable for the specific application.

Although these measurements are not sufficient to determine how much of the measured size variation results from true differences in particle size as opposed to random measurement error from each instrument, they do indicate that SMRs offer improved resolution over other particle sizing techniques, which may prove useful in detecting subtle size changes or the existence of subpopulations in heterogeneous samples.

To identify whether there were significant differences in precision between individual sensors in the parallel SMR array, we measured a sample of 8 μm polystyrene beads and compared the size distributions measured in each cantilever (supplementary material, Fig. 7). The coefficients of variation for each cantilever (ranging from 2.1% to 3.7%) were comparable to the overall coefficient of variation (2.6%).

DISCUSSION

Here, we have described methods to increase the volumetric throughput of SMR-based particle mass sensors. First, using new computational methods for extracting mass measurements from faster-flowing particles, throughput was improved by a factor of 16 compared to previously reported systems. Second, using devices containing arrays of SMRs connected fluidically in parallel and operated simultaneously, throughput was increased by a factor of ~27, while still resolving the size distribution of a polystyrene bead standard with greater precision than a high-end commercial Coulter Counter. Furthermore, we estimate that future versions of the parallel SMR arrays designed with full-length fluidic channels for compatibility with the deconvolution algorithm could achieve throughput as high as 59 μl/min, a 197-fold improvement compared to existing systems, and a further 7-fold improvement over the parallel SMR arrays described here.

In the future, larger arrays of SMRs could be designed with more than the 16 sensors used here. Several factors limit the maximum number of sensors that can be operated on the same chip. One of the primary limitations is the requirement to space out the resonances in the frequency domain to avoid mechanical or electrical coupling between sensors. We have observed that with our current approach, we begin to observe coupling between SMRs when the resonance frequencies are spaced much closer than ~25 kHz. Therefore, adding more sensors requires that the cantilevers be designed with resonance frequencies distributed over a wider range, i.e., for 50 cantilevers, the resonances must be distributed over a range of >1.2 MHz to achieve a minimum 25 kHz spacing. While this can be achieved, driving the cantilevers with sufficient amplitude at these high frequencies presents some practical challenges that would need to be solved.

We envision that increased throughput will extend the range of applications for which mass-based particle sizing can be employed. While mass has traditionally been less commonly used than volume for characterizing particle size, in some applications it provides unique advantages, such as the ability to distinguish particles of the same size but different densities.[15] In other cases, mass and volume provide equivalent information about a particle's size, such as in suspensions of solid particles of uniform density; in these cases, SMRs may still provide an advantage due to their increased precision compared to other size measurements. The increased precision of SMRs relative to other methods increases further for sub-micron-scale particles: SMRs have been successfully scaled down to measuring particles with buoyant masses as small as ~10 Ag (10 nm gold nanoparticles[16]), while miniaturized resistive pulse sensing instruments have been limited to particles 50 nm or larger.[17] The methods developed here are equally useful for increasing throughput for SMRs of all size scales.

SUPPLEMENTARY MATERIAL

See the supplementary material for the following figures and text referenced in the manuscript: Supplemental Figure 1. Resonance-tracking and noise-rejection dynamics of the SMR-PLL loop. Supplemental Figure 2. Comparison of optimization procedures for fitting transit time and entrance time. Supplemental Figure 3. Signal-to-noise ratio as a function of resonance-tracking bandwidth. Supplemental Figure 4. Parallel SMR array

Published under license by AIP Publishing

characterization. Supplemental Figure 5. SMRs with fluidic channels extending to the second-mode antinode. Supplemental Figure 6. Probability of two particles occupying the same cantilever simultaneously as a function of sample concentration. Supplemental Figure 7. Comparing precision between sensors. Supplemental Note 1. Maximum particle concentration to avoid multiple occupancy events.

ACKNOWLEDGMENTS

We thank Steven Wasserman for helpful discussions. This work was funded in part by Grant No. R33 CA191143 and the Cancer Systems Biology Consortium, Grant No. U54 CA217377 from the NCI. This work was supported in part by the Koch Institute Support (core) Grant No. P30-CA14051 from the National Cancer Institute.

REFERENCES

[1] A. Khodabandehloo and D. D. Y. Chen, Bioanalysis **9**, 313 (2017).
[2] K. D. Caldwell, H. K. Jones, and J. C. Giddings, Colloids Surf. **18**, 123 (1986).
[3] S. Son, A. Tzur, Y. Weng, P. Jorgensen, J. Kim, M. W. Kirschner, and S. R. Manalis, Nat. Methods **9**, 910 (2012).
[4] M. M. Stevens, C. L. Maire, N. Chou, M. A. Murakami, D. S. Knoff, Y. Kikuchi, R. J. Kimmerling, H. Liu, S. Haidar, N. L. Calistri, N. Cermak, S. Olcum, N. A. Cordero, A. Idbaih, P. Y. Wen, D. M. Weinstock, K. L. Ligon, and S. R. Manalis, Nat. Biotechnol. **34**, 1161 (2016).
[5] J. Riordon, M. Nash, M. Calderini, and M. Godin, Microelectron. Eng. **118**, 35 (2014).
[6] T. Vaclavek, J. Prikryl, and F. Foret, J. Sep. Sci. **42**, 445 (2018).
[7] D. K. Sharma, D. King, P. Oma, and C. Merchant, AAPS J. **12**, 455 (2010).
[8] K. L. Cooper, S. Oh, Y. Sung, R. R. Dasari, M. W. Kirschner, and C. J. Tabin, Nature **495**, 375 (2013).
[9] T. P. Burg, M. Godin, S. M. Knudsen, W. Shen, G. Carlson, J. S. Foster, K. Babcock, and S. R. Manalis, Nature **446**, 1066 (2007).
[10] N. Cermak, S. Olcum, F. F. Delgado, S. C. Wasserman, K. R. Payer, M. A. Murakami, S. M. Knudsen, R. J. Kimmerling, M. M. Stevens, Y. Kikuchi, A. Sandikci, M. Ogawa, V. Agache, F. Baléras, D. M. Weinstock, and S. R. Manalis, Nat. Biotechnol. **34**, 1052 (2016).
[11] S. Olcum, N. Cermak, S. C. Wasserman, and S. R. Manalis, Nat. Commun. **6**, 1 (2015).
[12] C. J. Schuler, H. C. Burger, S. Harmeling, and B. Scholkopf, in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (IEEE, 2013), Vol. 1067.
[13] J. Lee, R. Chunara, W. Shen, K. Payer, K. Babcock, T. P. Burg, and S. R. Manalis, Lab Chip **11**, 645 (2011).
[14] J. Lee, A. K. Bryan, and S. R. Manalis, Rev. Sci. Instrum. **82**, 023704 (2011).
[15] W. H. Grover, A. K. Bryan, M. Diez-Silva, S. Suresh, J. M. Higgins, and S. R. Manalis, Proc. Natl. Acad. Sci. U. S. A. **108**, 10992 (2011).
[16] S. Olcum, N. Cermak, S. C. Wasserman, K. S. Christine, H. Atsumi, K. R. Payer, W. Shen, J. Lee, A. M. Belcher, S. N. Bhatia, and S. R. Manalis, Proc. Natl. Acad. Sci. U. S. A. **111**, 1310 (2014).
[17] J. L. Fraikin, T. Teesalu, C. M. McKenney, E. Ruoslahti, and A. N. Cleland, Nat. Nanotechnol. **6**, 308 (2011).

Published under license by AIP Publishing

Supplemental Note 1: Maximum particle concentration to avoid multiple occupancy events

Regardless of volumetric flow rate, the maximum particle concentration that can be measured is limited by the fact that at high sample concentrations, multiple particles will occupy a cantilever at the same time, and the resulting resonance frequency signal is the sum of the two individual peak shapes. While it is possible to decouple these signals[11], it is useful to limit the particle concentration to avoid the need to employ these techniques. Here we estimate the frequency of double-occupancy events – events where two particles occupy the cantilever simultaneously – as a function of particle concentration.

Modeling the loading of particles into cantilevers as a Poisson process, the distribution of times between successive particles entering the cantilever is exponential, with density

$$f(t; T_{mean}) = \left(\frac{1}{T_{mean}}\right) e^{-t/T_{mean}} \quad [5]$$

A double-occupancy event occurs when the separation between consecutive particles is less than the transit time. This occurs with probability

$$P(\text{double occupancy}) = \int_{t'=0}^{t'=T_{transit}} f(t';\, T_{mean})\, \mathrm{d}t' = 1 - e^{-T_{transit}/T_{mean}} \quad [6]$$

The mean particle spacing T_{mean} can be expressed in terms of the sample concentration (c_0) and volumetric flow rate (Q) by noting that the flux of particles through the sensor is given by (equivalently)

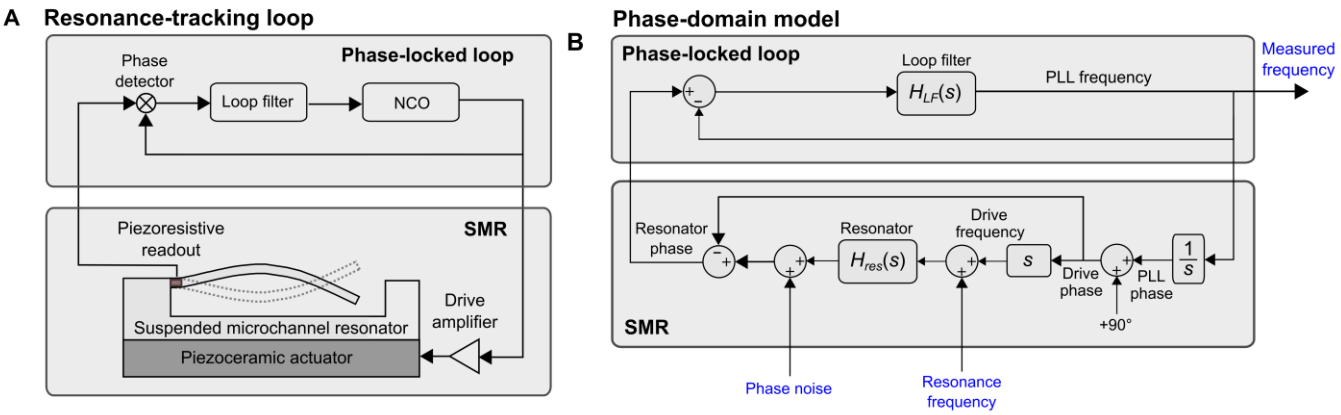
$$1/T_{mean} = c_0 Q \quad [7]$$

Finally, the volumetric flow rate is related to the sensor volume and particle transit time, $Q = V_{sensor}/T_{transit}$, under the approximation that particles travel at the mean flow velocity. Collecting these terms, the probability of a double occupancy event is

$$P(\text{double occupancy}) = 1 - e^{-c_0 V_{sensor}} \quad [8]$$

where the quantity $c_0 V_{sensor}$ is a dimensionless concentration corresponding to the average number of particles per cantilever volume (Supplemental Fig. 6). Therefore, to limit the probability of double-occupancy events to be no greater than p, the average number of particles per cantilever volume should be less than $\ln(1/(1-p))$. In our case, we chose to limit the probability of double occupancy events to $p < 0.1$, requiring that on average there be no more than ~0.1 particles per cantilever volume, or 1 particle per 10 cantilever volumes (Supplemental Fig. 6).

Supplemental Figures



C **Resonance-tracking and noise-rejection dynamics**

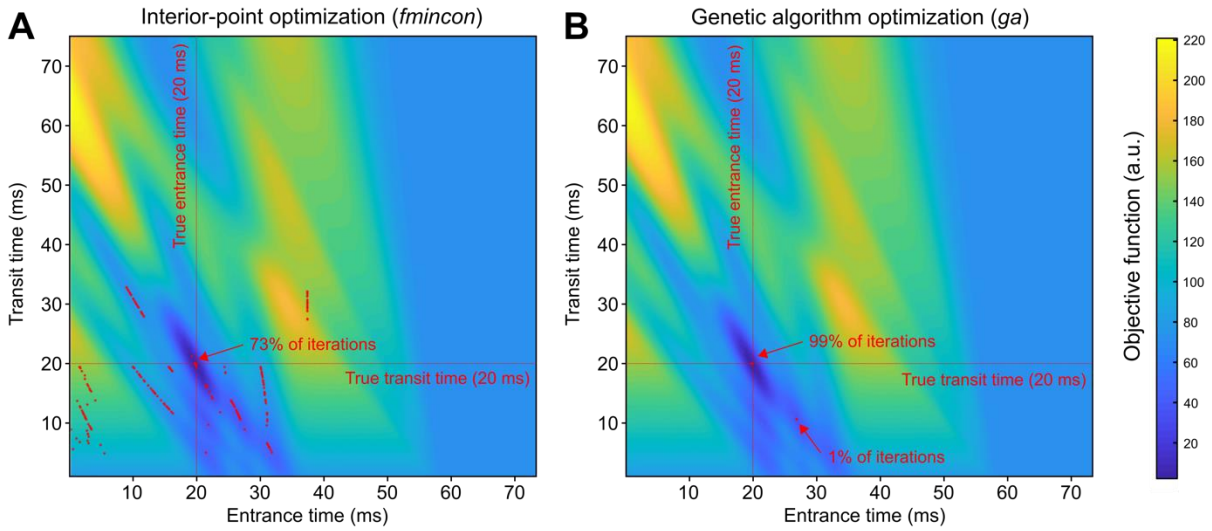
$$H_{\text{track}}(s) = \frac{(\text{measured frequency})}{(\text{resonance frequency})} = \frac{-H_{res}H_{LF}(s)}{1 - H_{res}H_{LF}(s)}$$

$$H_{\text{noise}}(s) = \frac{(\text{measured frequency})}{(\text{phase noise})} = \frac{H_{LF}(s)}{1 - H_{res}H_{LF}(s)}$$

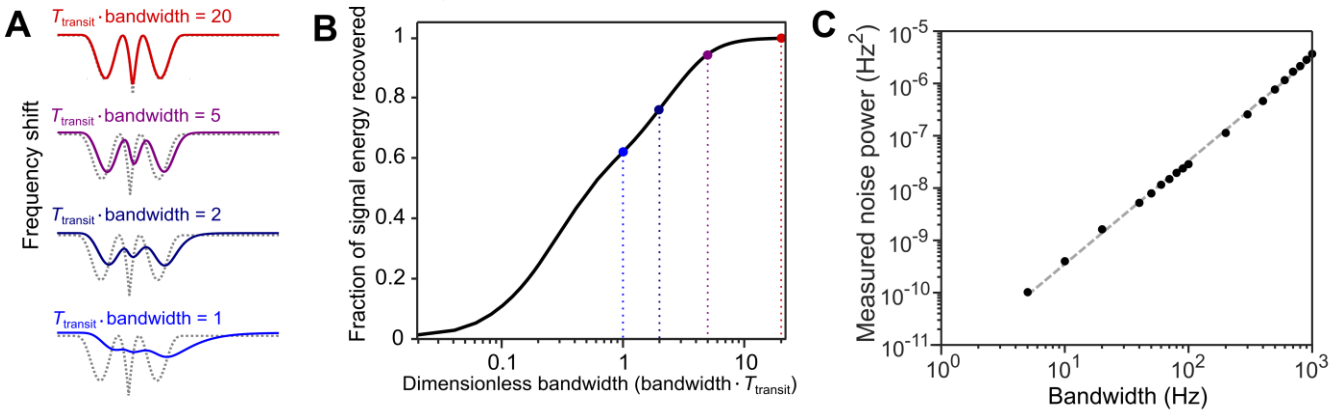
Supplemental Figure 1. Resonance-tracking and noise-rejection dynamics of the SMR-PLL loop.

(A) Schematic of the SMR-PLL loop used to continuously drive one or more cantilevers at resonance.

(B, C) Phase-domain model of the SMR-PLL feedback loop. The loop filter (transfer function $H_{LF}(s)$ is designed to shape the closed loop resonance-tracking and noise-rejection transfer functions ($H_{track}(s)$, $H_{LF}(s)$) with the desired dynamics.



Supplemental Figure 2. Comparison of optimization procedures for fitting transit time and entrance time. Color indicates the value of the objective function for a simulated peak with 20 ms true transit time and 20 ms true entrance time, with a global minimum coinciding with the true entrance time and transit time. Red points indicate the minima identified by 1000 repeated applications of the interior point (A) and genetic algorithm (B) optimization approaches. (A) The interior-point optimization procedure performed poorly, with only 73% of iterations correctly estimating the transit time within +/- 1 ms. (B) The genetic algorithm performed better, with ~99% of iterations correctly estimating the transit time within +/- 1 ms.

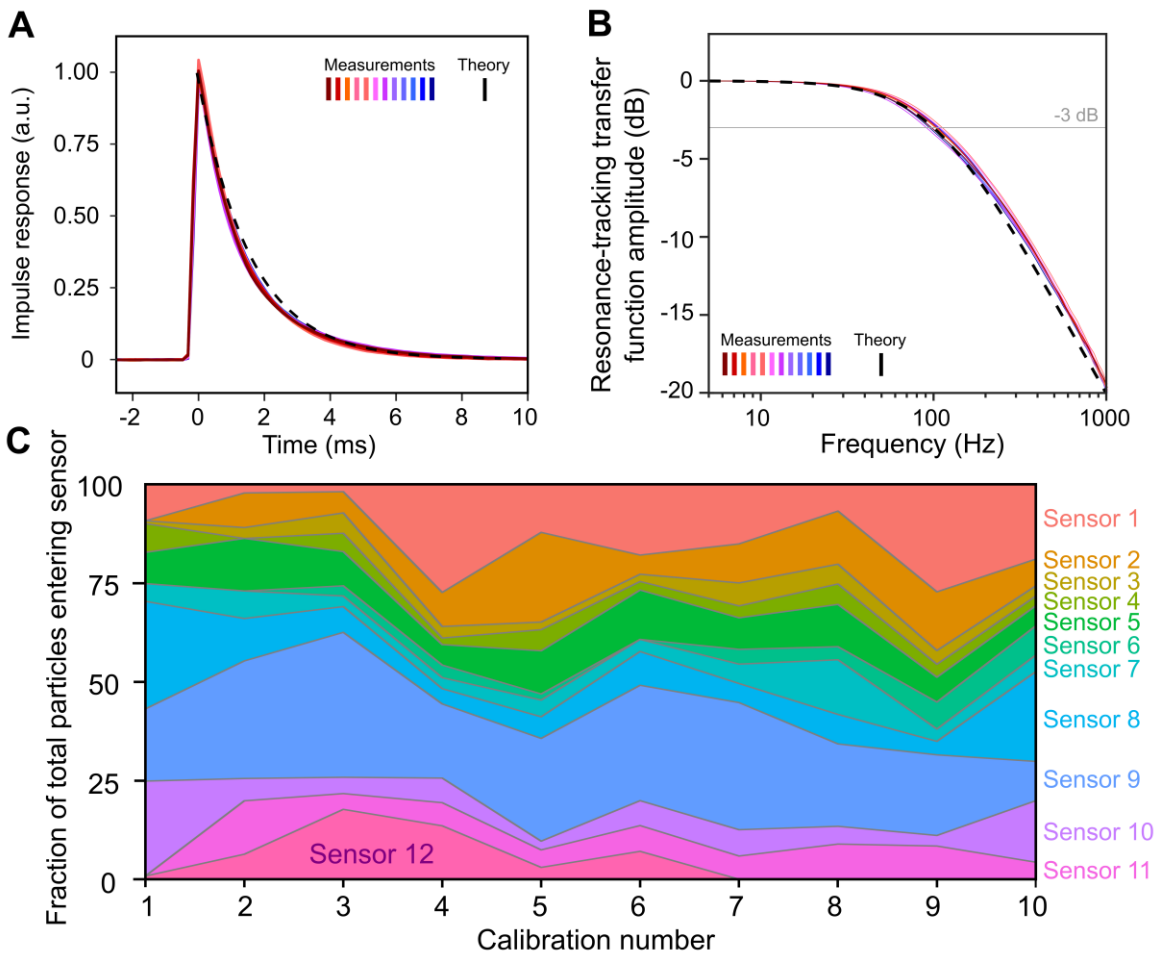


Supplemental Figure 3. Signal-to-noise ratio as a function of resonance-tracking bandwidth.

(A) Distorted peak shapes predicted by the linear model of the SMR-PLL resonance-tracking loop, as a function of the dimensionless bandwidth $T_{\mathrm{transit}} \times \mathrm{bandwidth}$. At narrower loop bandwidths, the noise power decreases, while the degree of signal distortion increases.

(B) Fraction of total signal energy recovered as a function of the resonance-tracking bandwidth of the SMR-PLL loop (shown here for a first-order resonance-tracking transfer function). The signal is fully resolved (signal energy >99.9%) for dimensionless bandwidths greater than approximately 24.

(C) Measured total noise power as a function of resonance-tracking bandwidth, determined from 1-second noise samples. The noise is approximately white (log-log slope approximately +2).

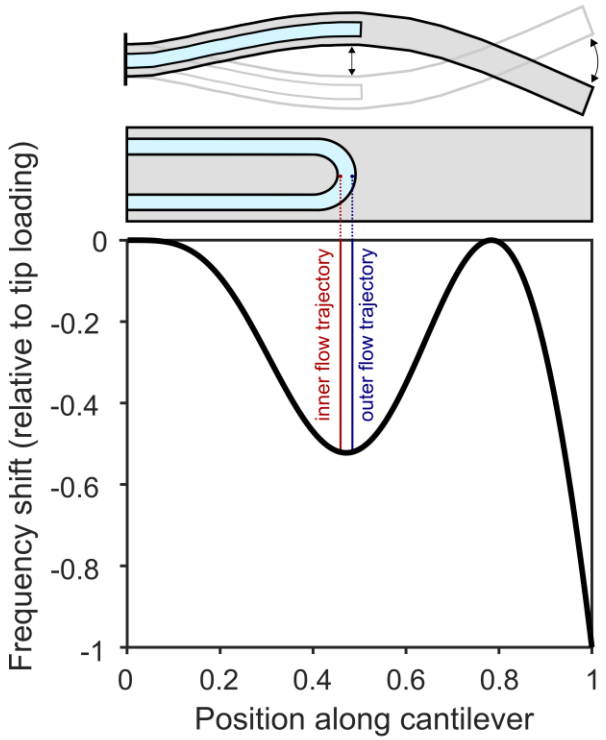


Supplemental Figure 4. Parallel SMR array characterization.

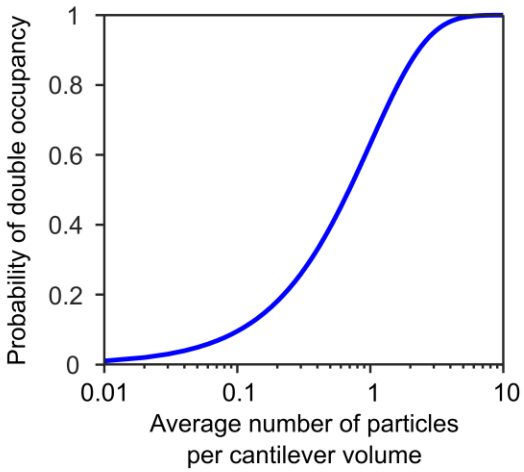
(A) Measured (red-blue) and predicted (black) impulse response for ten SMR-PLL resonance-tracking loops, each configured with a first-order resonance-tracking transfer function with 100 Hz bandwidth. Each trace is the average of 100 impulses induced by instantaneously offsetting the internal PLL phase by ten degrees.

(B) Measured (red-blue) and predicted (black) resonance-tracking transfer functions for the same ten SMR-PLL resonance-tracking loops, obtained as the Fourier transform of the measured impulse responses.

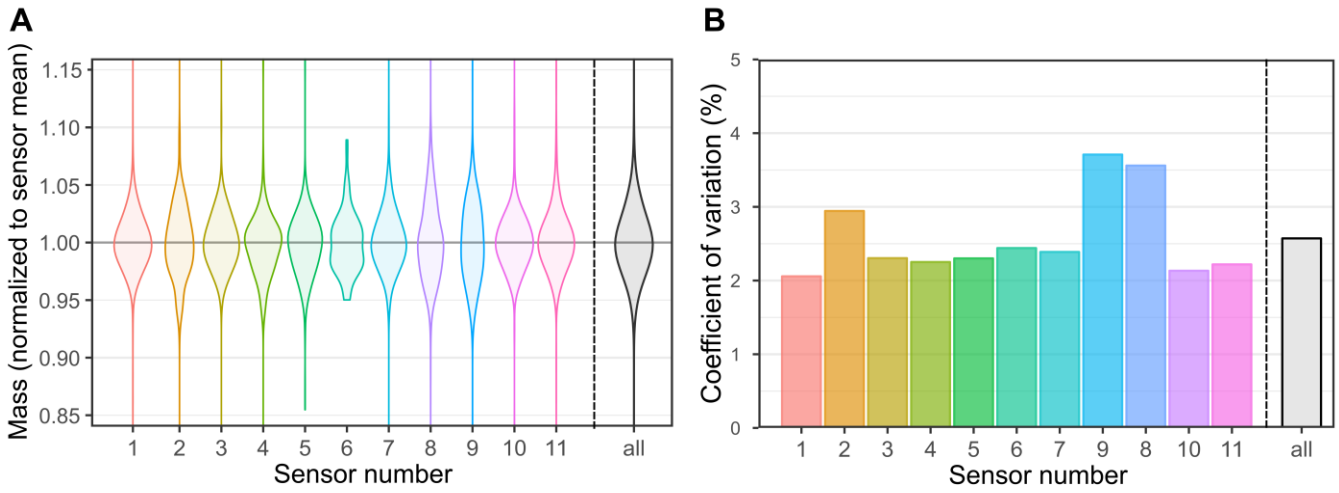
(C) Distribution of particle counts across the twelve sensors of a parallel SMR array measuring suspensions of 7 μm polystyrene particles on different days. Although the pressure control configuration remains nominally the same from day to day, we observe slight drift in the total fractions of particles passing through each sensor.



Supplemental Figure 5. SMRs with fluidic channels extending to the second-mode antinode. The cantilevers of the parallel SMR arrays were designed with fluidic channels extending only to the antinode when the cantilever is driven in the second mode (~48% of the cantilever length), resulting in a resonance frequency signal with a single peak when a particle passes through the sensor. This results in reduced position-dependent error (0.2%), since the peak resonance frequency shift is relatively insensitive to whether a particle takes the inner, outer, or center path when turning around at the tip of the cantilever.



Supplemental Figure 6. Probability of two particles occupying the same cantilever simultaneously as a function of sample concentration. As described in Supplemental Note 1.



Supplemental Figure 7. Comparing precision between sensors.

(A) Measured mass distributions for a sample of nominal 8 μm polystyrene beads measured on a parallel SMR array. An average of 874 particles were measured per sensor, with a minimum of 94 in sensor 6. The mass distribution for each sensor is normalized to mean 1.

(B) The coefficients of variation for each cantilever ranged from 2.1-3.7%, comparable to the overall coefficient of variation of 2.6%, suggesting similar performance between sensors.