Review of Scientific Instruments

Real-time nonlinear correction of back-focal-plane detection in optical tweezers

Tanuj Aggarwal and Murti Salapaka

Citation: Rev. Sci. Instrum. **81**, 123105 (2010); doi: 10.1063/1.3520463
View online: http://dx.doi.org/10.1063/1.3520463
View Table of Contents: http://rsi.aip.org/resource/1/RSINAK/v81/i12
Published by the American Institute of Physics.

REVIEW OF SCIENTIFIC INSTRUMENTS **81**, 123105 (2010)

Real-time nonlinear correction of back-focal-plane detection in optical tweezers

Tanuj Aggarwal[a)] and Murti Salapaka[b)]
University of Minnesota, 200 Union St SE, Minneapolis, Minnesota 55455, USA

(Received 24 July 2010; accepted 5 November 2010; published online 22 December 2010)

Photodiode based detection of laser trapped beads using forward scattered light is a frequently employed technique for position measurement. There is a nonlinear relationship between photodiode outputs and bead position but for small displacements linear approximation holds well. Traditionally, the nonlinearity is compensated by normalizing the photodiode's position signal with the intensity signal and then using a polynomial fit in the range where voltage to position mapping is one to one. In this article, this range is extended by using the intensity signal as an independent input along with the two position signals. A map from the input signals to the actual position values is obtained. This mapping is one-to-one for a larger range that results in an increased detection range. An artificial neural network that facilitates implementation is employed for this purpose. This scheme is implemented on a Field Programmable Gate Array based data acquisition and control hardware with closed loop bandwidth of 50 kHz. Detection of the order of 350 nm from the center of detection laser is demonstrated for a 500 nm diameter bead compared to 180 nm achieved by a polynomial fit.
 [doi:10.1063/1.3520463]

I. INTRODUCTION

Optical tweezers have played phenomenal role for over two decades in the measurement and application of forces to single biomolecules. It uses strongly focused laser light to create three dimensional potential gradient where the beads get trapped. Forward scattered light from the bead is typically collected on to a position sensitive or quadrant photodiode that provide signals that depend on the position of the bead relative to the trap center.[1] A variant of this method uses a separate laser beam of very low power that serves as a reference for all measurements. This decouples force actuation from force sensing and enables controlled experiments with position or force feedback.[2] Such a setup has been used to study various motor proteins like kinesin, a protein that walks on microtubules.[3] Many studies of motor proteins use optical tweezers under constant force mode, where the separation between trap center and bead center is regulated at a constant value. One of the limitations in such studies is the short detection range. During experiments, if the bead reaches the limits of detection, then either the entire sample is repositioned to a nominally selected initial position or the bead is forced back, disturbing the experiment midway. An increased linear detection range is therefore desirable. The relationship between photodiode signals and the actual position of the bead is nonlinear.[4] Researchers often use a linear approximation, which is typically valid for small deviations about the center. A polynomial fit of intensity normalized photodiode position signals can extend this range by approximately 30%.[3] In this article a new method to process photodiode signals is provided that gives accurate position measurement for a much larger range than the existing schemes. The method is based on artificial neural network (ANN) mapping of voltage signals to positional signals. Nonlinearity compensation techniques based on two dimensional polynomial fit[3] and ANN[5] exist but they are limited by the fact that the map from voltage to position is not one to one for a large range.[3] In this article, the domain in which voltage to position mapping is not one-to-many (feasible domain) is estimated and mapped to position signals. An order of increase in detection range was previously demonstrated by Perrone *et al.*[6] where a maximum likelihood estimator was used. However, their method is not practical for real-time implementation that is necessary in feedback based experiments. Also the results were valid for one dimension only. The method developed in this paper uses ANN instead which is easier to implement on hardware, works in two dimensions and preserves accuracy over a larger range.

II. SETUP

The experimental setup (Fig. 1) consists of a 1064 nm wavelength trapping laser source (Laser Quantum, Model Ventus IR 4W s-polarized) that passes through a two-axis acousto-optic-deflector (AOD, IntraAction Corp., DTD-274HA6). The beam is expanded and steered into the microscope objective (Nikon 100x, 1.4 NA, oil immersion). Detection laser (Point Source Inc., iFLEX 2000, 50 mW, 830 nm, p-polarized) is added collinear to the trapping laser using a polarizing beam splitter (PBS) cube. Intensity of the detection beam is reduced by placing a neutral density filter (ND) in its path. Intensity is adjusted such that it is less than required to trap a bead. After passing through the sample, the beams are collected by a 1.25NA condenser (obtained from a bench microscope). The trapping laser is blocked using a laser line filter (Thorlabs, FL830-10) and the back-focal-plane image of the detection laser is imaged onto a quadrant photodiode (Pacific Silicon Sensors, QP50-6SD2) with integrated

[a)]Electronic mail: aggar028@umn.edu.
[b)]Electronic mail: murtis@umn.edu.

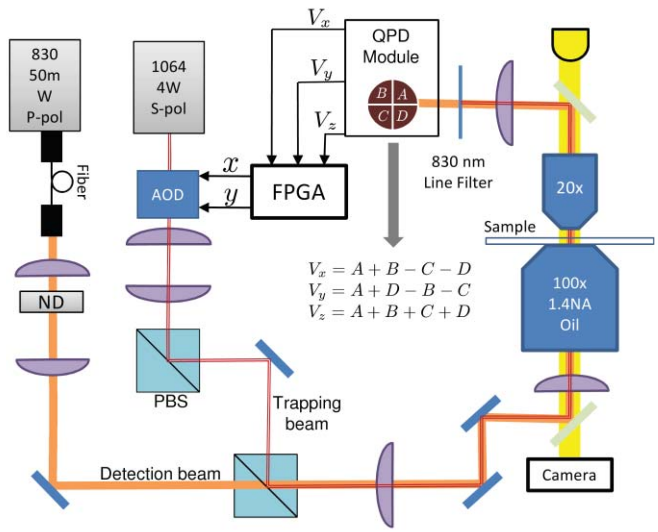

FIG. 1. (Color online) Experimental setup.

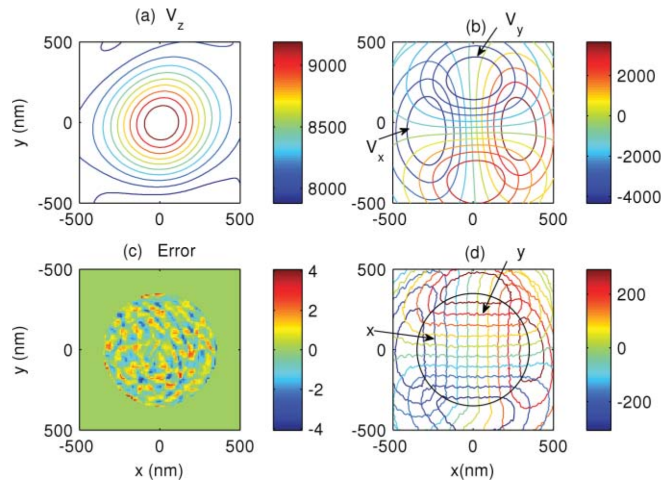

FIG. 2. (Color online) (a) Calibration data, contour plot for intensity signal. (b) Calibration data, contour plots for V_x and V_y. (c) Error plot between actual x position values and that estimated from voltage signals using a neural network. (d) Contour plots for estimated position values experimentally obtained from voltage signals. Noise in the data is partly due to thermal noise and partly due to AOD nonlinearity which is position dependent. Perpendicular intersections of the contour plots within the indicated circle of radius 350 nm indicates absence of cross-coupling.

amplifier circuit. The photodiode module provides three signals V_x, V_y, and V_z that represent asymmetry of light distribution on the photodiode along x coordinate, along y coordinate and the total intensity of light respectively (Fig. 1). The signals are captured by a Field Programmable Gate Array (FPGA) based data acquisition card (National Instruments, 7833R). Subsequent references to voltages in this article corresponds to the 16-bit number obtained after analog to digital conversion. Control logic and voltage to position mapping is programmed on this hardware using custom written code in LabVIEW for FPGA.

III. CALIBRATION AND DETECTION RANGE

Calibration data for voltage to position conversion are obtained by moving the trapped bead in a square grid using AOD centered about the detector beam. In this article, the grid size is chosen to be $1\,\mu\text{m} \times 1\,\mu\text{m}$ with 10 nm spacing between adjacent points. The scanned locations are denoted by (x, y). Photodiode voltage signals, V_x, V_y, and V_z are recorded for every grid point. The voltage signals V_x, V_y, and V_z are sensitive to the motion of the bead along x, y, and z coordinates, respectively, with strong cross-coupling as the bead moves further away from the center of the detection region. During scanning, the trap stiffness is kept sufficiently high and the scanning speed sufficiently low to ensure the bead is always equilibrated at the trap center and the effect of thermal noise is low. In traditional position inference schemes, the calibration data represent a map $g_{xv} : (x, y) \rightarrow (V_x, V_y)$ from position to voltage and the objective is to obtain the inverse map, $g_{xv}^{-1} : (V_x, V_y) \rightarrow (x, y)$ to infer the position from voltage. Inversion of g_{xv} over the entire domain (scanned region) is not possible because g_{xv} is not one-to-one in that domain. However, on a restricted domain g_{xv} is one-to-one where inversion is possible. To find the restricted domain, the following analysis is done based on the calibration data. Sample calibration data in the form of contour plots of V_x, V_y is presented in Fig. 2(b). Contour plots for V_x or V_y are overlaid for clarity that shows lines/contours joining points on x–y plane along which the voltage has a fixed value. The difference between the voltages of two adjacent contours is kept uniform. Contour lines of V_x intersects with those of V_y. If a pair of contour lines, one for V_x and the other for V_y intersect at more than one point then it implies that the voltage pair corresponds to more than one location on the x–y plane. Of all such locations that give the same voltage pair, only the one closest to the center is considered a part of the restricted domain or the detection region. By inspecting the shape of the contour plots, the boundary of detection region is inferred by the set of points where a pair contour plots are tangent to each other. This region is approximately circular with a radius of 250 nm. This limit is not useful in practice due to reduced noise robustness at the edges. A practical limit is therefore set around 200 nm. Lang *et al.*[3] use a fifth order polynomial to implement g_{xv}^{-1} in a region of radius 180 nm.

To overcome this limitation, a new methodology is proposed. A two input and three output map, $f_{xv} : (x, y) \rightarrow (V_x, V_y, V_z)$ is considered, which is shown to be invertible over a larger domain because now three contour lines (one for each voltage signal) must intersect simultaneously at multiple points to be out of the domain which happens approximately after 500 nm [by visual inspection of contour plots in Figs. 2(a) and 2(b)]. Beyond 500 nm the intensity signal does not change significantly and therefore it will be difficult for a static map to increase the detection range beyond 500 nm (for a similar setup as in this article). In this article, ANN is employed to do the inversion resulting in larger detection range. Here as well, the theoretical limit of 500 nm is not easily achieved and instead a practical limit of 350 nm is obtained. An ANN needs training sequences which is obtained from the calibration data.

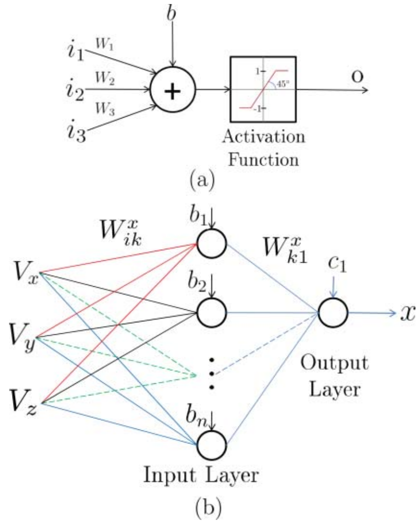

FIG. 3. (Color online) (a) Neuron model. Activation function shown is called "satlins" for symmetric saturated linear function. i_k are the inputs to the neuron. W_k are weights associated with the kth input. b is the bias associated with the neuron and O is the output of the neuron. (b) ANN model. Circles represent neurons. Inputs V_x, V_y, and V_z are inputs to the n neurons of input layer weighted by W_{ik} from ith input to the kth neuron. Biases for each neuron as shown by b_k. Likewise, output layer weight and bias is represented by W_{ko} and c_k, respectively. x is the output of the ANN. Another similar ANN is constructed with same inputs but different weights and biases that gives output y.

IV. ARTIFICIAL NEURAL NETWORK MAPPING

An ANN consists of layers of interconnected neurons as in Fig. 3(b). Figure 3(a) shows the model of a neuron that has several inputs and one output. The neuron computes the weighted sum of inputs, add a constant (bias), and passes the result to a nonlinear function (also known as activation function). The proposed network consists of two ANNs, one to estimate x position and the other to estimate y position. Each ANN consists of three input nodes, one input layer with 50 neurons and one output layer with one neuron and one output node. A simpler or smaller network resulted in higher than acceptable errors. All the input layer neurons have symmetric saturating linear activation function ["satlins," Fig. 3(a)] while the output layer neurons have "pure linear" activation function [$f(x) = x$]. Satlins is chosen for its ease of implementation. From a hardware point of view, satlins is simply a saturation filter. Other complex functions can be chosen using a look-up table but they are sensitive to discretization errors when implemented in fixed point hardware. The equations describing the action of neural network are as follows (using notation as in Fig. 3):

$$x = c_1 + \sum_{k=1}^{n} \text{satlins}\left\{ \sum_{i\in\{x,y,z\}} V_i W^x_{ik} + b_k \right\} W^x_{k1},$$

$$y = c_2 + \sum_{k=1}^{n} \text{satlins}\left\{ \sum_{i\in\{x,y,z\}} V_i W^y_{ik} + b_k \right\} W^y_{k2}.$$

The inputs to the ANNs are (V_x, V_y, V_z) and the desired outputs (targets) are the (x, y). The objective is to find the ANN weights and biases so that the neural network maps inputs to the targets. A two dimensional polynomial fit (position to voltage) is used to smooth out the noise in the calibration data due to Brownian motion of the bead and local nonlinearities in AOD. Typical nonlinearity in AOD is expected to manifest in form of small deviations ($\pm$4 nm) from the commanded position (see Valentine *et al.*[7]) which are smoothed out due to low order polynomial fitting. Smoothed version of voltage values are used for the subsequent computations. Scan locations that are within the circle with radius of about 350 nm from the center of scan are used for training ANN. Mapping a larger region requires bigger network that translates into slower computation and more integration of discretization error. Therefore, there is a trade-off between bandwidth and range. The training is done in MATLAB using the Neural Network Toolbox. The FPGA hardware does not support floating point operations whereby a fixed point implementation is written with weights and biases scaled by a large number (2^{20}) and the remaining fractional part rounded off. All computations and numbers are stored as 32 bit integers. After integer calculations, the result is scaled back. This process does not significantly affect the accuracy for the chosen architecture. The computational delay is less than 15 μs. Remaining operations, like data acquisition, control logic and AOD control are done in parallel to the above operations and take less than 5 μs. Sampling time is set to 20 μs. There is, however, a delay of a couple of sampling cycles from sampled data to the computed control signal. The number of neurons, network weights and biases are programmable during run-time and thus a recompilation of the code is not required. Polynomial computation within hardware like FPGA is not recommended because it requires evaluating powers and serial multiplications which accumulates errors if implemented in fixed point. Method suggested by Perrone *et al.*[6] requires the use of a parametric map, which is basically a form of calibration data. A data with 100 grid points will have 10 000 memory elements, therefore hardware implementation will require storing these many elements which again is quite impractical considering other tasks that the hardware has to perform and a smaller number of grid points will lead to interpolation errors where nonlinearity is severe. Second, complex algorithm like maximum likelihood estimator will have huge overheads of their own.

V. VERIFICATION AND CONCLUSIONS

The nonlinear mapping capability of ANN is demonstrated in Fig. 2(b). In the region indicated by a circle of radius 350 nm, the output of the network matches with the scanning location with excellent accuracy (approximately 1 nm rms and 4 nm peak-to-peak). Experimental data taken after calibration for a scanning experiment is shown in Fig. 2(a). Time data sample for similar experiment is shown in Fig. 4.

Tests independent of AOD were performed to verify the mapping method and to expose errors due to AOD nonlinearities that might be hidden in the previous result. For this purpose, statistics of the Brownian motion of the trapped particle were obtained postcalibration and various tests performed including power spectrum, histogram, and diffusion rate. For power spectrum, a 500 nm diameter bead was trapped and

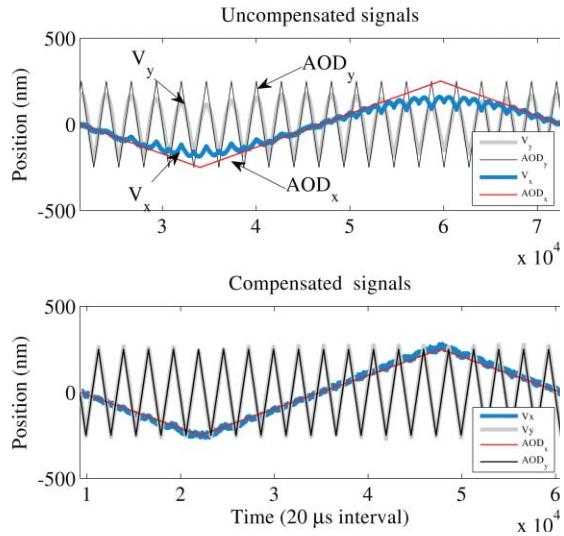

FIG. 4. (Color online) Time plot of signals during scanning of size 500 nm peak to peak obtained from hardware. AOD$_x$ and AOD$_y$ are the input signals. V_x and V_y are the corresponding outputs. Bead reaches approximately 350 nm away from the center at certain times. (a) Unfiltered signal with highly nonlinear response (voltage signals are scaled and shifted for comparison). (b) Position signals obtained after passing voltage signals through neural network with no scaling or shifting residual nonlinearity in the V_y signal is due to the AOD.

positioned at five locations within the calibrated region. One location centered at the origin of calibration region and the four others were centered diametrically opposite at a distance of 100 nm from the center. The Brownian fluctuations took the bead further to about 150 nm from the center. Within this range, the detection is linear therefore raw photodiode output can be scaled to get position estimate. Figure 5 shows that the ANN calibrated data matches the linearly scaled photodiode output.

Histograms for x and y were obtained for trap positioned at the center and near the edge of ANN detection range. The histograms (Fig. 6) are close to Gaussian curves. There is some variation in the histograms at different locations (seen in spectrum plots as well). Some of these variations are expected due to the nonlinearities in AOD that are getting exposed after calibration (smoothing process removed AOD nonlinearities during calibration). Variation due to calibration errors is not ruled out, however, the results are satisfactory for most practical experiments. Another reason for the variation is small but measurable effect of the detection beam. A particle experiences the effect of both the trapping beam and the detection beam. Moreover, the effect of detection beam is nonlinear in the calibrated range (350 nm from center).

To test the effect of detection beam, a diffusion test was performed on the particle. The main trap was set to high intensity and switched on and off at a rate of 50 Hz, i.e., the trap stayed on for 10 ms and off for 10 ms. Several records of bead trajectory after the trap went off was stored. From the records the variance of the particle position (over several records) with time is plotted (Fig. 7). Ideally, if the effect of the detection beam were truly negligible then a straight line plot is expected, but in reality a slowly saturating curve is obtained that is to expected of a Brownian particle in a potential field. Simulations were performed for the experimental conditions and stiffness of detection beam was chosen to fit the experimental diffusion curve. The obtained stiffness values (otherwise not easily obtainable) characterized the effect of detection beam on the bead.

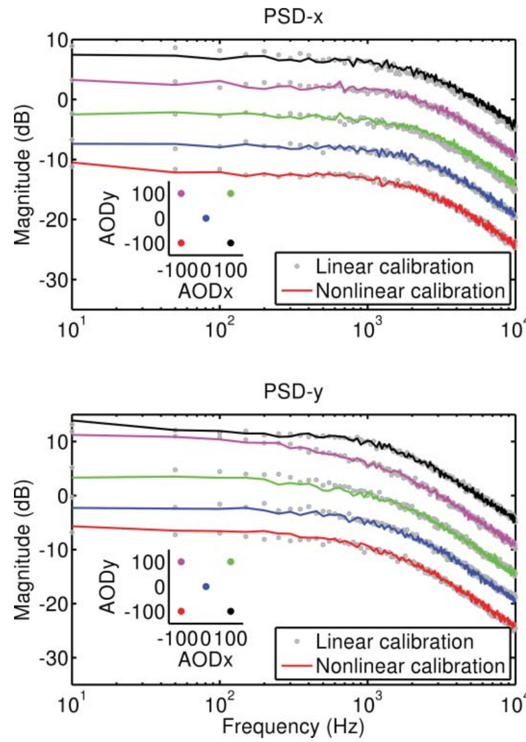

FIG. 5. (Color online) Power spectrums of x (top) and y (bottom) bead positions. Spectrums are offset to each other for clarity. Bead was centered at five locations marked by the colors in the inset with axes labeled (AOD$_x$, AOD$_y$). These locations are within the traditional linear detection range (150 nm). Dotted line shows the spectrum obtained by scaled raw photodiode data. Spectrum obtained by ANN calibration overlay those obtained by linear scaling.

In summary, we have demonstrated a convenient way to process photodiode signals that removes detection

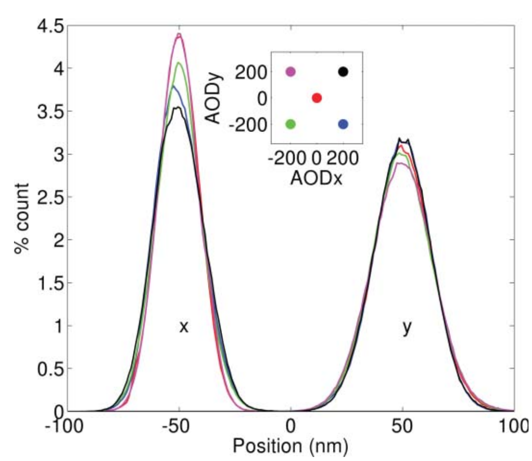

FIG. 6. (Color online) Normalized histograms of trapped bead position obtained with ANN calibration. Typical Gaussian curves are obtained as expected of harmonic traps. Variance in the x coordinate is smaller compared to y coordinate indicating higher stiffness in x direction. This is expected of traps created by linearly polarized lasers. The inset shows legend for the histograms in graphical format, with markers indicating the position of the trap (AOD$_x$, AOD$_y$) where histograms were measured. The mean position of the histograms was shifted to -50 and 50 for x and y, respectively, for clarity.

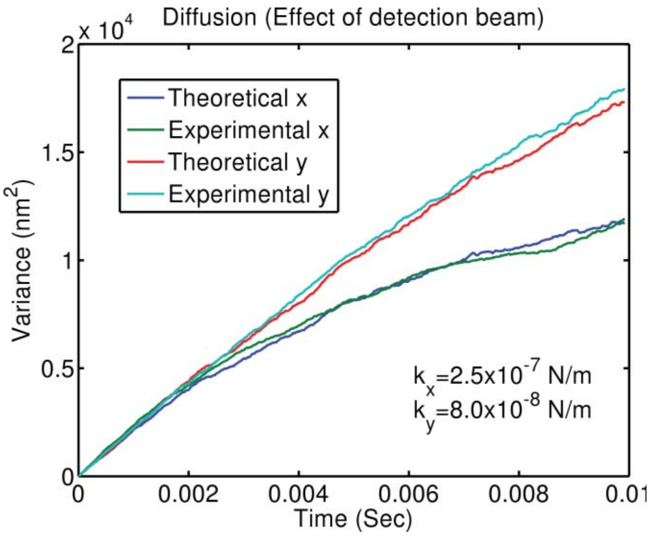

FIG. 7. (Color online) Diffusion in presence of detection beam. Trap was shut off to allow bead diffusion. Particle was positioned close to the origin of calibrated region before beginning the experiment. Stiffness of detection beam is lesser along y direction resulting in more linear plot compared to x direction. Simulated diffusion plot is overlaid for chosen stiffness k_x and k_y to match the experimental curves. A good match indicates that the detection beam has a measurable effect on the bead dynamics.

nonlinearity for a larger range and also suited for real-time implementation that is necessary for feedback based experiments. It is also an effective method to remove cross coupling between x and y position signals which occurs if alignment is imperfect. The calibration process is quick enough to be used on a regular basis. Independent tests were performed based on the Langevin dynamics of a trapped particle to test the calibration. The results were largely satisfactory with small variations not necessarily attributed to the calibration errors but to AOD nonlinearities and the effect of detection beam that were actually getting captured after the calibration process. The scheme can potentially be used for three dimensional calibration as well.

[1] A. Pralle, M. Prummer, E. Florin, E. Stelzer, and J. Hörber, Microsc. Res. Techn. **44**, 378–386 (1999).
[2] K. Visscher, S. Gross, and S. Block, IEEE J. Sel. Top. Quantum Electron. **2**, 1066–1076 (1996).
[3] M. Lang, C. Asbury, J. Shaevitz, and S. Block, Biophys. J. **83**, 491–501 (2002).
[4] F. Gittes and C. Schmidt, Opt. Lett. **23**, 7–9 (1998).
[5] W. Zi-qiang, L. Yin-mei, L. Li-ren, W. Heng-hua, and W. Zhong, Opt. Precis. Eng. **16**, 6–10 (2008).
[6] S. Perrone, G. Volpe, and D. Petrov, Rev. Sci. Instrum. **79**, 106101 (2008).
[7] M. Valentine, N. Guydosh, B. Gutiérrez-Medina, A. Fehr, J. Andreasson, and S. Block, Opt. Lett. **33**, 599 (2008).