Good afternoon. My name is Michael Culbertson, and I will be speaking to you today for myself and my colleagues Dan Burden and Josh Williams of Wheaton College on "Beyond the Boundaries: Numerical Fluorescence Correlation Spectroscopy for the Analysis of Complex Biomolecular Dynamics."

Over the last thirty years, Fluorescence Correlation Spectroscopy has grown to a place of prominence in microscopy circles as a versatile method of analyzing molecular dynamics. The technique has been applied to samples as diverse as DNA, polyelectolytes, colloids and nanoparticles, organic thin films, lubricants, emulsions, lipid membranes, ionic liquids, infectious agents, and the list goes on.

In traditional one-photon FCS, a sample with a fluorescent tracer is placed in a focused laser beam, and the fluorescent bursts are captured and recorded as the tracer moves through the detection region. This fluorescence signal is then autocorrelated;

and the autocorrelated data are fit with an analytical expression to extract a quantity of interest, such as diffusion constant, triplet-state dynamics, reaction kinetics, or rate of bulk flow, among others. In conjunction with advances in the quality of the optics in confocal microscopy, FCS has achieved the capability of extracting this information even from samples of single-molecule concentration.

Now, despite the power and general utility of the technique, traditional FCS has some limitations, like every other method. Although analytical expressions exist for a good variety of experimental conditions, some scenarios present particular challenges to deriving a mathematical expression for the interaction of tracers and the three-dimensional detection volume, especially if they involve some nanoscale geometric constraint that limits the tracers' freedom in the detection region.

Normal free-volume measurements in solution can be represented by graphic A in the upper left. Here, a molecule diffusing, say, along the red path is free to occupy any space in the yellowish detection region, and the molecule behaves homogeneously no matter its relative position within the volume. This fairly straightforward situation poses little problem for FCS, and analytical expressions abound for these conditions. However, other plausible measurement scenarios aren't quite so simple. For example, the analyte may be present in a gradient flow profile, as depicted in graphic B in the upper right. Here, molecules in the faster laminae will have considerably shorter fluorescence bursts than those in the slower, altering the autocorrelated signal in complex ways. Other examples depicted in C and D at the bottom include molecules confined inside or on the surface of nano-tubular structures or in multi-domain two-component lipid membranes. In each case, the fluorescent molecules' motion is confined or changes depending on the molecules' position.

Now, the autocorrelation curves produced under these kinds of conditions still contain valuable information, even though it can't be accessed via traditional FCS. Rather, the complex interactions between fluorescent tracers and the detection volume in many of these cases can be simulated quite easily or modeled by other numerical methods unamenable to standard curve-fitting.

We therefore propose an extension method of FCS, dubbed Numerical Fluorescence Correlation Spectroscopy that circumvents the analytical expression of traditional FCS by making use of these numerical techniques for modeling the behavior of fluorescent molecules in the detection region.

Instead of fitting the autocorrelated data with an analytical expression, our new method compares the experimental data with a synthetic curve based on a guess of the parameters of interest.

These parameters are then adjusted iteratively to minimize the differences between the experimental and synthetic curves, as represented in the sum of squared residuals. This process is related to the "simulation-based fitting" that has been applied to other analytic methods such as Fluorescence Resonance Energy Transfer (FRET).

The synthetic curves may be obtained alternatively via simulation or direct calculation based on numerical models of the detection volume and molecular motion.

In the case of simulation, there are a large number of techniques available in the literature. I'll briefly describe the method we use for single-molecule diffusion studies in our lab. Our simulator superimposes a 3D lattice onto a numerical map of the detection profile, symbolized by the grey region. Molecules walk along the lattice as shown, for example, by this black path. The step size along the lattice, dx, is adjusted per species to equal to the root-mean-squared diffusion distance given the species' diffusion constant, D, and a step time, dt, typically on the order of 500 ns. Each step, the molecule moves one lattice position in a random direction, and the fluorescence probability is calculated for a molecule in that position. The fluorescence probabilities are summed over all molecules for a given bin width, and the final simulated collected fluorescence emission is sampled from a Poisson distribution, giving a real-time record of fluorescence bursts. This record is then autocorrelated for NFCS analysis.

Here's a cartoon of the simulator operating. You can see that as molecules (the green dots) move into the laser area (the red circle), a fluorescence burst is recorded below. Our simulator doesn't actually make use of a graphical display such as this one in order to maximize computational efficiency. The core routines of our simulator are coded in C for speed, but the user interface is provided through

scripting in the Python programming language, which allows for the powerful, dynamic execution of both simulation and analysis on the fly.

Now, simulations can become computationally intensive rather quickly, especially as more features are added to the simulation. To render extensive simulation for analysis more feasible, we distribute the computational load across a bank of user workstations.

Each simulation is divided into time segments, which are processed by individual computers in parallel to reduce computation time. The results are then transfered back to a main computer. The simulations run at a low priority, allowing them to recover otherwise-wasted CPU cycles without disrupting other users. They could run equally well, of course, on a dedicated scientific computing cluster.

Now, when simulations produce a large volume of data, as for example with a real-time fluorescence trace, network data transfer rates can limit the throughput of the simulator.

We can mitigate this effect by reducing the amount of data sent back to the main computer by beginning the data processing—namely, the autocorrelation—at the remote source. To that end, we've developed a novel autocorrelation algorithm that combines the partial results from different computers as if their real-time records had been concatenated before autocorrelation.

Our distributed correlation algorithm is an extension of an existing software multi-tau correlator. The algorithm keeps track of both the beginning and end of the processed data stream, in addition to the partial correlation sums.

Only this reduced set of data—a mere 4.5 kB—is transfered back to the main computer, which then calculates the correlations that overlap the break between the segments, based on the reserved data. If the data stream processed by two correlators, A and B, is repre-

sented by this dashed line, each correlator reserves a copy of the beginning and end of its record (the clear ovals). When the partial results are returned to the main computer, the correlations for the data in between (in the greyed section) have already been calculated. The combination algorithm then merely calculates the correlation for each lag-time that crosses the line. The use of distributed programming, along with this new distributed autocorrelation algorithm, has permitted the successful and timely use of lengthy simulation in our Numerical FCS analyses.

The alternative method to simulation for generating the synthetic curves in NFCS is direct calculation from numerical models of both the detection volume profile and molecular motion.

This method is particularly useful for molecules that obey standard Brownian motion in abnormally shaped detection volumes.

Here, the autocorrelation function is written as the probability of producing a photon at two different locations (r_1 and r_2), moderated by the probability of moving from the first location to the second in the given lag time (τ), integrated over all space. The photon intensity probability, O, is given by the detection volume profile, and the movement probability by Green's Function, Ψ.

The double integral can be calculated quite efficiently by performing a Fast Fourrier Transform-based convolution of the detection profile with Green's function, multiplying the result by the detection profile, and summing in three dimensions. So, Numerical FCS can make use of either of these two methods—simulation or direct calculation—to generate synthetic autocorrelation curves based on a guess of the experimental parameters. This guess is then iteratively adjusted to minimize the differences between the synthetic curve and the experimental data.

As an example of the utility and effectiveness of NFCS, we've

applied the method to analyze diffusion data for Rhodamine B under abnormal optical conditions. Just as nano-scale geometric constraints can limit the effectiveness of traditional FCS, a highly non-standard detection profile can equally render derivation of an analytical expression for the autocorrelation function impossible. Although our example is a bit contrived, it illustrates the principles of NFCS, which translate directly into real-world situations. Anomalous detection volumes, such as the one in this image from the Molecular Probes website, can result from misaligned confocal optics. These volumes can be mapped quite easily by scanning through the detection region with a small fluorescently tagged polysterene sphere that has been non-specifically adsorbed to a glass slide. We use 170-nm diameter 540/560-nm fluorescent spheres. The acquired image can be smoothed with various filtering procedures and sharpened by deconvolution with the 170-nm sphere. The processed map can then be used in the simulation or direct calculation of synthetic autocorrelation curves for NFCS analysis of experimental data taken with the same microscope under the same optical conditions.

We employ a custom-built confocal microscope in our lab, illustrated by this diagram. Light from a laser is directed through a shutter and iris to a dichroic mirror, which reflects the incident light through an objective and onto a sample resting on a 3D piezoelectric translation stage. The same objective collects fluorescence emissions, which pass through the dichroic mirror, band-pass filter, and pinhole, and are finally focused onto an avalanche-photodetector module. Half of the collected light is diverted to a CCD camera for brightfield imaging. For the purpose of this demonstration, we modified the optical train to include two beam-splitting cubes and a couple of mirrors. These create two nearly-copropagating beam paths that are slightly misaligned from one another, just to make things

interesting. Either of the beams can be selected individually with a simple beam stop.

We collected detection volume profiles and data from free-diffusion in aqueous solution under a variety of optical configurations. Each employed the 543-line of a 5-mW Helium-Neon laser, a 150x, 1.25 NA epiplan-apochromat water-emmersion objective, and Rhodamine B.

Varied parameters included pinhole size (from 75 μm to 400 μm), whether or not the back aperture of the objective was overfilled, and the number of beams (either one or both).

These different configurations produced some widely varied and interesting detection volume shapes, such as those depicted here. The tall volume on the left was produced with an underfilled 400 micron pinhole, and the one in the center top with an overfilled 75 micron pinhole. The lower central volume was produced with an overfilled 100 micron pinhole. Although these conditions are frequent in FCS measurements, notice how the effected volume is rather peanut-shaped, as opposed to the more cylindrical approximation used to derive the analytical expressions for traditional FCS analysis. The most interesting, I think, is the detection region created by the two co-propagating beams, which has these light and dark regions near the top where the two beams overlap and constructively and destructively interfere.

The autocorrelation signatures of molecules diffusing through these volumes are equally as interesting. As would be expected, larger volumes produce longer fluorescent bursts—that is, curves shifted to the right. But, the differently shaped volumes also produce autocorrelation curves with different contours, as shown by these three example curves. For example, notice how the plus curve is more curved than the relatively flat descent of the triangle and

square curves. These differences in curve contour reflect deviations in the shape of the detection region from the idealized mathematical model used to derive the analytical expressions for traditional FCS, and could pose problems for analysis via traditional FCS if these deviations can't be modeled analytically, as in the case of the double-beam detection volume.

Both simulation and direct calculation produced similar autocorrelation curves, as seen on the left: the points represent sum-of-squared-residual comparisons of simulated curves with a given experimental trial, and the line comparisons using direct calculation. Notice how the curves match over a large range of diffusion constants. Any optimization algorithm can be used to match synthetic and experimental data. We use the Simplex algorithm to minimize the sum of squared residuals. For direct calculation, this works quite well on its own. But, since simulated autocorrelation curves are subject to stochastic noise, the merit function has a large number of local minima which makes getting a precise value difficult, as seen on the right. For NFCS via simulation, we typically use the Simplex algorithm to estimate the diffusion constant and then fit a range of values near the estimate with a quadratic to find the actual minimum. Analysis via direct calculation is generally faster than simulation, but simulation can provide more flexibility in modeling unusual or tricky conditions.

Matching autocorrelation curves, such as the red one here, fit the experimental data rather well, with sum of squared residual values frequently on the order of $10^{-}4$.

Finally, we performed collected three sets of detection volume profiles and associated diffusion data for each optical configuration and analyzed them with NFCS using both simulation and direct calculation. The results showed a high degree of consistency, with an

average of about $4.5 \cdot 10^{-6}\mathrm{cm}^2/\mathrm{s}$, which falls within the range of diffusion constants reported for Rhodamine B in the literature. High levels of consistency in the final result despite considerable differences in the input conditions indicate the capability of Numerical FCS to accommodate non-standard conditions and extract reliable diffusion information from scenarios previously inaccessible by traditional FCS, either due to aberrations in the shape of the detection volume profile, as demonstrated here, or in nano-scale geometric constraints to molecular motion. Given the flexibility of the numerical approach, we expect Numerical FCS to broaden the possibilities for molecular dynamics analysis in many new and interesting systems.

Among the many who have helped make this work possible, I would like to express our particular gratitude to the Math/Computer Science Department of Wheaton College, who graciously grant us the unused CPU cycles of their student computer lab. And, I thank you for your attention.