ISICL: *In situ* coherent lidar for particle detection in semiconductor-processing equipment

Philip C. D. Hobbs

A scanning coherent lidar (laser radar) for detecting and mapping isolated submicrometer particles in hostile or inaccessible regions such as plasma chambers, ovens, tanks, and pipes is described. The sensor uses a noise-canceled diode laser homodyne interferometer of novel design that is insensitive to misalignment, runs at the quantum limit, and requires just one access window. At a false-count rate of 10^{-5} Hz, the sensor needs 50 photons to detect a particle. A combination of techniques makes the system immune to stray light or laser light scattered from the chamber walls, though these other light sources may be 10^6 times more intense than the desired signal.

Key words: Lidar, laser radar, semiconductor-process control, particle counting, *in situ* measurement, laser particle counters, particulate contamination.

1. Introduction

One of the most daunting challenges faced by semiconductor manufacturers is the rapid tightening of allowable limits on particulate contamination, in both concentration and size. If these limits are not met, chip yields will decrease catastrophically as features shrink and chips grow. Because of this, instruments for detecting particulate contamination are of great economic importance.

Particles come from a great many sources; the historically prominent ones of stirred-up dust and human byproducts (fluff, skin flakes, droplets) are dealt with by the use of extremely effective clean rooms and mini-environments, in which the wafers are not exposed to ambient air during handling. The dominant sources of particulate contamination in semiconductor manufacturing today are the process tools and fluids themselves. In many cases there is no good way to detect elevated concentrations of these particles before they have damaged a considerable number of wafers or to determine where and when the particles arise in the process.

Many types of semiconductor-processing tools are important sources of particulate contamination on wafers. Some of the worst offenders are quartz-sputtering chambers, reactive-ion etchers, electron cyclotron resonance ashers, ion implanters among vacuum tools, and wafer tracks (for application of photoresist) and dip tanks among liquid-based ones.

Currently, there are three main methods for single-particle counting (i.e., counting particles one at a time rather than in an ensemble as in turbidity meters, integrated scatterometers, and laser diffractometers): fluid sampling, wafer scanning, and *in situ* monitoring.

Wafer scanning further breaks down into patterned (product) wafer scanning and bare (monitor) wafer scanning. Patterned wafer scanners are in general more expensive, less sensitive, and slower than bare wafer scanners; this is because of the large background signal from the wafer itself. This background must be reduced by the use of techniques such as video-image processing, grazing-incidence dark-field detection, or spatial filtering, in order to make the particle signals detectable. On the other hand, many processes cannot be adequately tested with monitor wafers. When using monitors to test a process, a bare or blanket-film wafer is put into the volume of interest. The process under test is run, then the wafer is examined to determine how many particles were deposited during the process. An example of this approach is the use of monitor wafers to measure contamination in sputtering and other process chambers, which has been the most common approach for many years. This requires no tool modifications or special handling, and since it uses a count of particles per wafer to predict the counts on other wafers, it has a comfortable familiarity. How-

The author is with the Manufacturing Research Department, IBM Thomas J. Watson Research Center, Yorktown Heights, New York 10598.

Received 1 June 1994; revised manuscript received 25 August 1994.

0003-6935/95/091579-12$06.00/0.

© 1995 Optical Society of America.

ever, it has a number of disadvantages that are making it less popular. Each wafer costs perhaps $100, and even with rework costing around $20 per cycle, it can be used only approximately 10 times before being scrapped due to surface haze or other sorts of degradation. There is also a loss of tool availability. Altogether, it is too expensive to run such a test more than once per day, or once per shift at most, resulting in a long feedback cycle time. Many processes, such as chemical vapor deposition and reactive-ion etching, are sensitive to loading; process conditions in a silicon etch, for example, are strongly affected by the proportion of the wafer surface that has exposed silicon. Thus a bare silicon monitor wafer cannot be used with the plasma on, and the test is usually done with the plasma off. The correlations obtained in this way between particle counts on the monitor wafer and those on product wafers run in the real process are often poor, the more so as tool conditions often change too fast to be caught by such infrequent testing.[1] Testing with patterned wafer scanners is even more sporadic. Although they are very useful, these tools are slow and expensive, so that often it is only wafers from processes that are current trouble spots that get tested.

A sampling system removes some of the working fluid of the tool (gas or liquid) and causes it to flow through the sensitive volume of an external instrument. When detecting aerosols, the usual choices are a laser particle counter, generally based on dark-field light scattering, or a condensation nucleus counter.

Liquid-borne particles require entirely different equipment, either of the dark-field light-scattering type or, more recently, a Nomarski-type interferometric sensor with improved rejection of bubbles.[2] Sampling particle detectors are cost effective for those applications that allow their use. The equipment differences are largely due to the requirements of different working fluids. Air at room temperature is relatively benign, but some instruments must work in flue gases, hot oxygen, or other corrosive or explosive mixtures. Liquid-borne particle counters may require sample cells that resist attack by strong corrosives such as HF (requiring sapphire windows) or that can cope with the growth of biological films in their plumbing. These particulars lead to a large variety of specialized hardware, which tends to increase the cost and inconvenience of particle detection.

Besides practical and economic difficulties, there is an obvious fundamental flaw in a sampling scheme as applied to plasma chambers: there is no fluid to sample. Particles in these chambers may be ballistic or electromagnetically entrained[3–5] (although some of them are carried by the flowing gas) and often cannot be removed by sampling. Existing systems for mapping particle traps in plasma chambers[6] require good optical access to the chamber and cannot in general detect single submicrometer particles. Photoresist application tools and dip tanks are somewhat more accessible, but it is often difficult to put the sampler close enough to the wafer to get good correlations between sample counts and particles on the wafer.

Most *in situ* systems operate by placing part of the particle-counter apparatus inside the desired sensing region. This has been done quite extensively in the pump lines of load locks[7,8] and chambers,[9] and to a lesser extent elsewhere.[10,11] This technique works well for some purposes, often yielding information about the state of the process tool that is not apparent from any other measurement, and offers real-time process monitoring without excess tool downtime. An important advantage of this class of techniques is that the counting statistics are usually good—tens to thousands of counts per wafer.

It is often difficult to transport particles to the pump line for sensing, especially since the particles of greatest economic importance are usually large—a few micrometers or even bigger. In the pressure and velocity regime of most chambers, these particles are not strongly influenced by the gas flow and would be unlikely to survive passage through a turbo pump to get to the particle counter. Many *in situ* techniques are hampered by the environmental sensitivity of the counter heads, which usually will not tolerate elevated temperatures, corrosive gases, or condensation of material on their optical surfaces. They are also often sensitive to strong electromagnetic fields, such as those inside the shield of an ion implanter beam line. Besides the process environment interfering with the counter, the counter may interfere with the process. It is obviously infeasible to put a counter head between a sputtering target and a wafer, for example, or inside an electrostatic particle trap in a reactive-ion etch chamber, yet these are some of the most interesting applications for *in situ* monitoring.

The primary rationale for the work described in this paper was the need for a robust, affordable, noninvasive device to count small particles anywhere in a plasma chamber or fluid tank, with sensitivity good enough to be useful (<0.25 μm), and a volumetric sampling rate that (though necessarily small because of limited detection N.A.) is nonetheless large enough (30–100 mL/min) to obtain reasonable counting statistics in a short time. The device had to operate inside the chamber with the plasma on (running the real process); to sample as near the wafer as possible; to yield spatially resolved particle information, so as to ease the identification of particle sources and the finding of traps; to provide reliable counts, without false alarms and with stable sensitivity; to be reasonable in cost, so that it can be used widely; and above all to work on the widest possible variety of processing equipment.

This paper begins with a discussion of the design constraints and performance requirements for the system, then treats its principles and design, and finally presents the performance of the current development version. Mathematical Appendix A contains

a calculation of the expected signal strength, for comparison with the signals observed.

2. Backscatter Operation

The application presents a stiff list of design constraints. The sample volume and spatial resolution requirements mandate scanning the sensing volume, which means that the detector must scan too; because of this, and since most processing tools do not have two large, diametrically opposed windows, the sensor had to work in backscatter.

Commercially available particle counters work in dark-field mode, as shown in Fig. 1, either in near-forward scatter or in doubly dark-field mode, with the detector near 90° to the incident beam; none works in backscatter, and for good reason. The smallest particles to be detected (which drive the entire design) are near the Rayleigh limit, so that the differential scattering cross section is a weak function of azimuthal angle (with the polarization vector of the incoming beam being the polar axis); thus in the equatorial plane the signal strength does not depend strongly on the position of the detector. The stray light, on the other hand, is a strong function of the detector position. Light scattered from transmitting optical surfaces is worst in the near-forward directions, but light scattered from the beam dump is worst in backscatter. In a dark-field detector, any scattered light must have made at least two bounces off the walls of the enclosure in order to get to the detector, and with good beam-dump design this will greatly reduce the scattered-light intensity. In backscatter, however, only one bounce is required; the point at which the beam intersects the chamber wall is in the field of view of the detector. Besides geometry, a backscatter detector has no control over its optical environment; few processing chambers have the flat black finish of a good beam dump. Together, these reasons mean that a backscatter particle detector must deal with bright stray light,

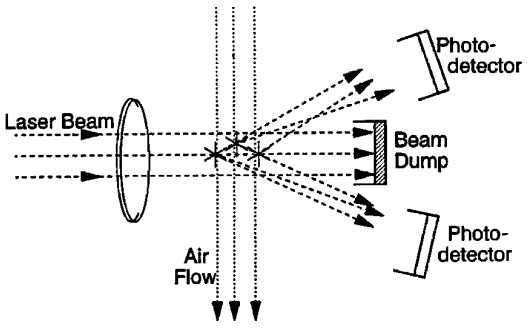

Fig. 1. Simplified optical schematic of a typical dark-field particle detector. Particles coming from the top enter a sampling volume, where they encounter a laser beam. The beam crosses the chamber at least once (many times in some designs) and finally is absorbed in a beam dump. Some of the light scattered by the particles in the sensitive volume is collected in photodetectors, which may be photomultipliers, avalanche photodiodes, CCD's, or ordinary silicon photodiodes. A particle crossing the beam thus produces a flash of light at the detector, which gives rise to an electrical pulse that can be detected by the back-end electronics.

thousands of times worse than that seen by a dark-field detector. We can estimate the stray light in the following way. Assume a photon-counting detector requiring 50 photons for reliable particle identification (see Appendix A), operating in backscatter with a 100-mW laser at 820 nm and a N.A. of 0.008 (these numbers are close to those for the actual system). Assume further (optimistically) that the back wall of the chamber is a Lambertian reflector, so that the differential scattered brightness in backscatter is $(\pi)^{-1}\,\mathrm{sr}^{-1}$. The incident beam has 4×10^{17} photons/s, so that the backscattered stray light has an average brightness of 1.3×10^{17} photons/s/sr; the detector solid angle is $\pi\mathrm{N.A.}^2 \approx 0.0002$ sr, so the total detected signal due to stray light is approximately 2.6×10^{13} photons/s. Assuming (again similarly to the actual system) that a particle takes 3 μs to yield 50 photons (1.7×10^7 photons/s), the stray light is 10^6 times more intense than the signal from a nominally detectable particle. What is worse, the signal from the back wall exhibits speckles, which move rapidly as the beam is scanned, giving rise to large (order-unity) fluctuations about the average stray-light intensity. The size of the speckles (and hence the bandwidth of the speckle noise) depends on the distance from the focus to the chamber wall and on the surface finish of the wall. Besides the back-wall scatter, in many applications the chamber interior is brightly lit by fluctuating plasma glow, as well as a continuum from ion gauge filaments and the room lights. Scatter from the chamber window is not usually a significant problem, because a window sufficiently dirty to scatter comparably to the chamber wall is too opaque to use—if the sensor can reject the wall, it can reject the window.

3. Coherent Detection

It is clearly impossible to use ordinary photon counting to find a 50-photon signal in a fluctuating background this large. Spatial filtering helps to some degree, but not enough to overcome such a factor. The most suitable (and probably the only feasible) candidate is coherent detection.[12] A coherent detection system superposes the received signal on a local oscillator (LO) beam, as in a heterodyne radio receiver, and then detects the sum of the two beams on a photodiode. Because of the interference between the two beams, signal components that are temporally and spatially coherent with the LO beam are amplified greatly, while others are not. The amplification is nearly noiseless; the signal-to-noise ratio of the beat signal is nearly the same as that of the scattered light; thus coherent detection gives the effect of a matched spatial filter plus a nearly noiseless amplifier. In the present system, the effect of coherent detection is to reduce the ratio of the photocurrent i_{stray} (due to the stray light) to i_{ac} from 10^6 to approximately 100. This helps a great deal, although problems may still be encountered if a very non-Lambertian back wall happens to reflect strong stray light directly along the axis of the LO beam.

4. *In Situ* Coherent Lidar

Figure 2 shows the optical schematic of the *in situ* coherent lidar (ISICL) system. In essence, it is a Michelson-type interferometer that uses well-known polarization techniques to separate and recombine the beams with little power loss. It also uses a good-quality corner reflector in one arm to produce a LO beam that is accurately parallel to the received light but offset from the transmit beam axis. The other arm contains a lens, which will sometimes have a particle near its focus. Since a lens with a reflector at its focus behaves optically exactly like a corner cube (its return rays come back reflected through the lens axis but exactly parallel to the incident rays), this design will guarantee that the beams from the two arms of the interferometer will interfere well when they are recombined, thus eliminating the sensitive angular alignment requirements of a plane-mirror interferometer.

Beginning with the left-hand side of Fig. 2, the diode laser assembly consists of a temperature-stabilized, single-longitudinal-mode diode laser (SDL-5420), a collimating lens, and (in a later system) an anamorphic prism pair for beam circularization. The beam emerges accurately collimated (laser collimation is the only critical adjustment in the system) and enters a polarizing beam-splitter (PBS) cube. The laser is oriented so that its beam is linearly polarized at an angle of approximately 5° from horizontal, so that it is approximately 99% *p* polarized at the cube diagonal and 1% *s* polarized. The cube reflects a small amount of the *p*-polarized light as well as the *s* component reflected from the cube diagonal, so that the reflected power is approximately 2%, and approximately 98% of the light emerges as the transmit beam. This beam passes through a quarter-wave plate (which converts it to left-circular polarization in the usual beam-separator configuration), and then through the objective lens, a 400-mm achromatic doublet. The now weakly focused beam (0.008 N.A.) emerges from the lens, is steered by a two-axis galvanometer mirror scanner, then passes through the window and into the chamber.

The 2% of the light that was reflected from the cube diagonal passes into a dielectric corner reflector and thence back into the polarizing cube; it is antiparallel to and offset from the beam leaving the cube. If its polarization were unaltered, it would be mostly reflected at the cube diagonal once more; however, since it has undergone three successive total internal reflections from a dielectric–air interface, its polarization has been changed, so that some of the beam is transmitted by the PBS, emerging orthogonally to the transmit beam (shown as downward in Fig. 2) to form the LO beam of the interferometer. The remainder, reflected by the PBS back towards the laser

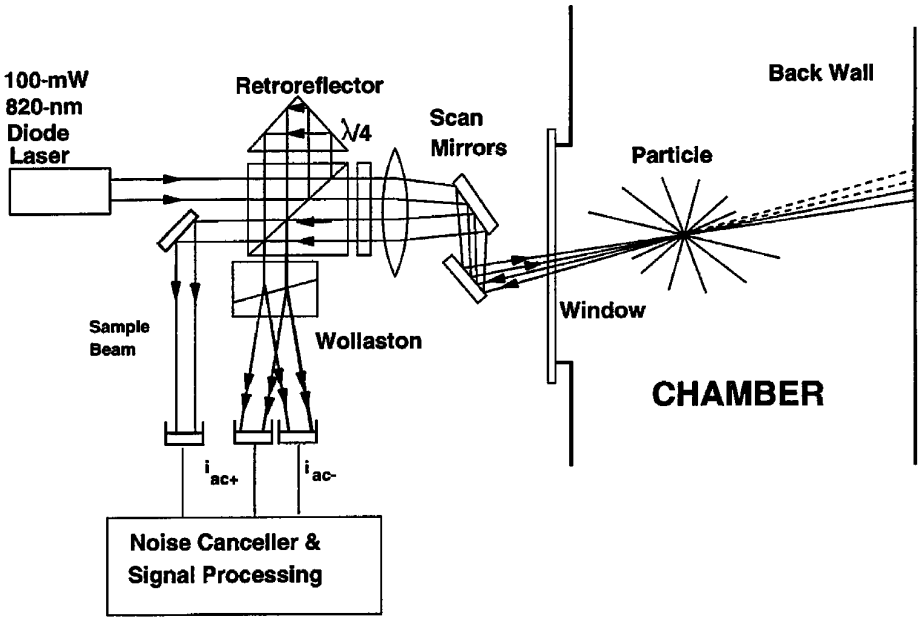

Fig. 2. Block diagram of the ISICL sensor head: light from the laser on the left-hand side passes through a polarizing beam splitter cube. Most of the light is transmitted, converted to left-circular polarization by the quarter-wave plate, focused by the large objective lens, and scanned around the chamber by the galvo mirrors. Most of the light in the transmit beam hits the back wall and is scattered in all directions. If, as shown, a particle happens to cross the focal region, it scatters some of the light. A small fraction of this scattered light makes its way back via the galvo mirrors and lens to the wave plate, where its polarization is converted to vertical, so that it is entirely reflected at the cube diagonal. There it is recombined with a LO beam derived from the light that was reflected from the beam splitter on the first pass; this beam has been folded by a corner reflector, so that it returns antiparallel to itself. Provided that the particle is at the focus of the lens, and the corner cube has been made sufficiently accurately (20 μrad), the LO beam and the scattered light interfere perfectly. They are, however in orthogonal polarization states, a problem remedied by resolving their fields about the axes of a Wollaston prism oriented near 45°; when these two components are detected and the resulting photocurrents are subtracted, the beat signal is recovered with excellent photon efficiency, while laser noise and dc are rejected.

(but offset sideways), is picked off by a small mirror and detected with a photodiode. The remainder serves as a sample beam for the laser noise canceller. The amount of polarization change, and hence the relative strengths of the two beams, may be adjusted by rotation of the corner reflector slightly about its threefold symmetry axis, although the beam should not be permitted to hit an edge of the cube.

When a particle crosses the beam near focus, it scatters light into approximately 2π sr. A small fraction of the scattered light (Fig. 2, dotted lines) will re-enter the optical system in such a way as to intersect the diagonal of the beam splitter on top of the LO beam. This light will be predominantly right-circularly polarized, because of the reflection, and so will be converted to vertical linear polarization (*s* polarized at the beam splitter) by the wave plate. Thus the scattered light from the particle is reflected from the PBS and combined with the light of the LO beam. If the particle is at the focus, the scattered beam is recollimated by the objective lens, and emerges antiparallel to the transmit beam. Since both the LO beam and the scattered beam are reflected once from the diagonal of the PBS, the identical actions of the lens–particle system and the corner cube ensure that the scattered beam is exactly aligned with the LO beam, and so (apart from being in orthogonal polarizations) the two beams will interfere well. This means that the interferometer needs no alignment except for collimation of the laser beam, which ensures that the beam waist coincides with the geometric focus of the lens. The LO and scattered beams are in the same state of focus if the particle is at the transmit beam waist.

The combined scattered and LO beams pass through a Wollaston prism oriented near 45° to their polarization vectors, producing two beams whose amplitudes are proportional to the sum and difference of the LO and scattered field amplitudes. Because of the sign difference, when these two beams are detected on photodiodes, they produce interference terms i_{ac+} and i_{ac-} of opposite sign, whereas the intensity terms i_{LO} and i_{scat} are of the same sign. These two photocurrents are subtracted, yielding a nearly zero-background measurement of the interference signal.

The combined scattered and LO beams could be made to interfere by the use of an analyzer and a single photodiode; this is wasteful, though, and prevents the most efficient use of differential detection and laser noise cancellation to eliminate excess laser noise.

The output of the differential detector is a tone burst, whose duration is the transit time of the particle through the sensitive volume, and whose carrier frequency is the Doppler shift of the scattered light.

Particles not exactly on the beam axis, or not exactly in focus, will not provide perfect interference, so the signal strength in this case will be reduced, with the reduction becoming worse the further from the axis the particles are.

Since the LO and transmit beams are of identical diameter, the transmit and receive N.A.'s are the same. This means that in the focal plane, the sensitive volume has the same width and functional form as the transmit beam intensity pattern. Because of the angular offset between the two beams, their intersection region is shorter than the depth of focus of the beams, and the sensitivity tapers off before and behind the focal plane as the angular offset increases the transverse displacement of the beams. With a N.A. of 0.008, the sensitive volume at the e^{-2} points is approximately 72 μm wide × 2 mm long.

Collimation of the beam is important since the received electrical power is proportional to the square of the Strehl ratio of the spot, measured at the geometric focus of the lens. In this case, with an 8-mm 0.5-N.A. collimating lens, a Strehl ratio of better than 0.95 (measured on a Wyko Ladite tester) can be reliably obtained without the use of astigmatism-correction optics. Although the laser has astigmatism, this can be compensated by using the lens slightly off axis, introducing an adjustable combination of coma, defocus, and tilt that can be used to compensate the astigmatism quite well. An astigmatic beam has saddle-shaped wave fronts, so that adding coma to flatten one side of the saddle converts most of the residual error into defocus and tilt. This is convenient since the usual astigmatism-correction methods, which rely on cylindrical lenses, have serious disadvantages. Anamorphic lens systems can circularize the beam as well as cancel the astigmatism but are limited by the difficulty of fabricating good-quality cylindrical lenses at low cost. Single-lens correctors, typically small cylindrical lenses of approximately 4000-mm focal length, can be adjusted only by rotating them; unless the required correction happens to fit the lens closely, correction of the required quality cannot be obtained this way.

Coherent detection is very helpful in rejecting spurious signals, but in the present case, with the clutter more than 60 dB (optical) stronger than the desired signal, more is required. The moderate working distance (30 cm) required allows the use of depth discrimination and scattering angles offset somewhat from 180°, so that the clutter is not in the same state of focus as the scatter and, more importantly, is not exactly on the axis of the LO beam. Because of the angular offset, baffles can be placed so as to block the largest stray-light contributions without obscuring the desired signal. These effects lead to greatly reduced sensitivity of the system to the clutter versus the scatter and, together with coherent detection and Doppler processing, make the system feasible.

5. Doppler Processing

One solution used in radar applications is Doppler processing; since the target is generally moving rapidly compared with the coherent background (clutter), the two can be distinguished reasonably reliably by the use of filtering. In the present system, there are two ways of looking at the origin of the Doppler

shift in the scattered light, which are of course mathematically equivalent but are heuristically different enough to be complementary.

The first is that the particle forms the mirror in one arm of an interferometer, so that the particle's radial velocity causes the phase delay in that arm to change with time, leading to a frequency shift in the detected interference signal. The second is to picture the LO beam as propagating in from infinity, passing through the sample volume (interfering with the transmit beam there), and then in through the rest of the system until it reaches the detectors. In this picture, the frequency shift is caused by the particle crossing the fringes formed by the interference of the two beams at the sensitive volume. This picture is helpful for calculations of sample volume and so on.

In either case, the Doppler shift in the detected signal due to a particle that is traveling with velocity $\boldsymbol{v}$, encountering incident light with wave vector $\boldsymbol{k}_i$ and scattering it into a wave with wave vector $\boldsymbol{k}_s$, is

$$\boldsymbol{f}_d = \frac{\tilde{\mathbf{v}}(\tilde{\mathbf{k}}_s - \mathbf{k}_i)}{2\pi}. \tag{1}$$

For a system operating in backscatter, $\mathbf{k}_s$ is approximately $-\mathbf{k}_i$.

At 830 nm, a particle moving axially at 50 cm/s will give rise to a tone burst whose carrier frequency is approximately 1.22 MHz. This is the nominal maximum particle axial velocity for the current version of the ISICL.

6. Laser Noise Cancellation

Optical particle counters must be high-sensitivity devices. In the Rayleigh limit, a particle of radius a has a differential scattering cross section proportional to a^6. Even for an amplitude-sensitive device such as the ISICL (whose signal amplitude goes as a^3), this function dies off discouragingly rapidly as a decreases. The system must work at the limit of its capacity in order to make the most of the limited number of photons available.

The leading sources of noise and interference in the system are stray light, excess laser noise, and shot noise. Given the large optical losses encountered in scattering light from a small particle, squeezed light is of little value here, and so the best we can do is the shot-noise limit. Even attaining that limit requires elimination of the excess laser noise.

Good, single-frequency diode lasers are already reasonably quiet, except for their relaxation oscillation peak at a few gigahertz; in general, however, they are still quite far from the shot noise, and are highly susceptible to mode hopping resulting from spurious optical feedback. Noise due to mode hopping and other noise sources (such as mode beats in gas lasers or power supply fluctuations) can be essentially eliminated through the use of all-electronic laser noise cancellation, as has been shown recently.[13] This earlier study on single-ended optical detectors achieved a noise floor of 3 dB above the shot-noise level and an ultimate excess-noise rejection of 60 dB over a bandwidth of more than 10 MHz. In that study feedback-controlled subtraction of a sample photocurrent was used. The 3-dB penalty arises from the shot noise of this sample current, which contains no information but must have the same strength as the signal current, and hence contributes the same amount of shot noise. In the present case, this 3-dB penalty can be eliminated, since both of the beams contain signal information as well as shot noise. The laser noise canceller used here, shown in Fig. 3, directly subtracts the photocurrents i_{ac+} and i_{ac-} (which are

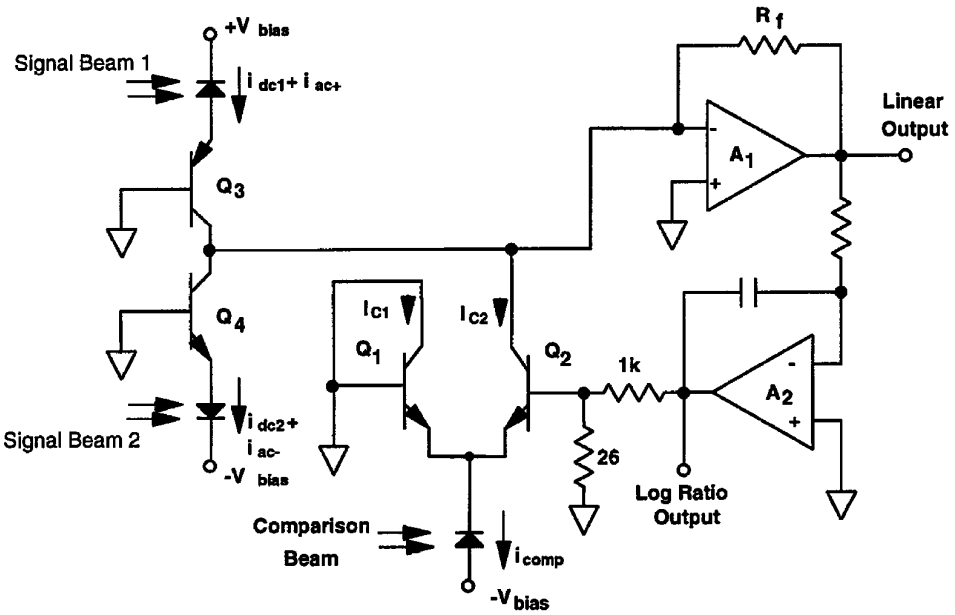

Fig. 3. Simplified schematic diagram of the differential laser noise canceller, which allows the system to be shot-noise limited with lasers that are much noisier than the laser noise canceller.

due to the two beams leaving the Wollaston prism) to do most of the cancellation, leaving only the small discrepancy between them to be eliminated by the feedback-controlled subtraction. This strategy preserves all the information from the scattered light, while achieving the same order of laser noise rejection as before. The sample-current shot noise can be made negligible, since the sample current can be much smaller than the signal currents.

Laser noise appears as a common mode signal in the differential detector, whereas true signals enter differentially. Thus spurious differential signals arising from laser mode hops (induced by back-wall reflections) and incidental étalon fringes may give rise to artifacts that the noise canceller is powerless to remove. Careful design and construction is required to eliminate these problems.

Having eliminated both stray light and excess laser noise as limiting factors, the performance of the system is set solely by the number of signal photons collected and the shot noise encountered in the detection bandwidth. As shown in section 8, the elimination of laser noise is so effective that the detected noise follows the theoretical Gaussian amplitude statistics of pure shot noise (which predict the number of false counts due to noise as a function of threshold level) within 0.10 dB out to 7.1 sigma, corresponding to an outlier rate of 1 part in 10^{11}, even with severe scatter from the back wall of the chamber.

7. Signal Processing

The basic signal-processing task is now straightforward. After the detected tone burst is amplified and filtered appropriately, the instantaneous signal amplitude is compared with a tracking threshold, and when a threshold crossing occurs, a valid particle event is declared. The remaining engineering problems center on the choice of detection bandwidths and the accurate estimation of the shot-noise level in the presence of large signals due to particles.

It is well known that the optimal receive filter for a signal with additive white Gaussian noise (in the sense of maximizing signal-to-noise ratio) is the complex conjugate of the received pulse spectrum.[14] In a pulsed radar application, where the Doppler shift is small compared with the reciprocal of the pulse width, the spectrum of the received pulse differs little from that of the transmit pulse, so that the optimal receive filter is just the complex conjugate of the transmit pulse spectrum. In the present case, where the Doppler shift may be large compared with the transit time bandwidth, the optimal signal-processing system depends on the range of particle velocities expected. In quiet plasma chambers, where most particles orbit slowly within well-defined traps, the maximum expected velocity may be as low as 5 cm/s, whereas in an environment such as the roughing line of a vacuum pump or a rapidly stirred fluid tank, the velocity range may be much greater. The scanning of the beam focus through the inspected volume is much more rapid (approximately 20 m/s) than the transverse particle speed in most cases, so that the width of the tone burst is set by the scanners rather than by the tangential velocity of the particle. On the other hand, since the system operates in cw, and the transmit and LO beams are practically completely temporally coherent with each other, the beat note has no intrinsic minimum bandwidth to speak of; therefore, the required detection bandwidth is set by the combination of the particle transit time (the tone-burst width) and the range of Doppler shifts expected.

With carrier frequencies ranging from 0 to 1.22 MHz, the Doppler bandwidth is much larger than the transit time bandwidth of the present system [150 kHz for a 3-μs burst (FWHM)], so that it is inefficient to perform the thresholding operation in a single band. In the present system, four bands are used to cover the Doppler bandwidth, as shown in Fig. 4.

8. False-Alarm-Rate Statistics

The design goal of accurately measuring low particle concentrations mandates a very low, controlled false-alarm rate (FAR). For Gaussian noise with bandwidth B, the level-crossing rate R at a threshold set at α times the rms noise is[15]

$$R(\alpha) = \frac{B}{2^*} \exp\left(-\frac{\alpha^2}{2}\right), \tag{2}$$

where the constant 2^* is close to 2 but depends on the filter used to set the bandwidth; 2^* is equal to $\sqrt{3}$ for a rectangular (brick-wall) low-pass filter and is equal to 2 for a 2-pole Butterworth low pass. In the present sensor, the design point is 1 false count per day (1.2×10^{-5} Hz), which in a bandwidth of 1.22 MHz and with a symmetrical bipolar threshold requires a threshold ratio α of approximately 7.1, so that approximately 7.1^2, or 50, photons are required for detection of a particle, assuming a matched filter (see Appendix A). Figure 5 shows the close agreement of the experiment and theory, out to 7σ. The semilog plot in Fig. 5(a) has too many vertical log cycles for experiment and theory to be distinguished well, so Fig. 5(b) shows the ratio of the actual to the predicted counts, converted to an imputed amplitude error in decibels, together with the 2σ (90%) confidence limits (the imputed amplitude error is the amount that α has to be adjusted to make Eq. (2), predict the exact measured count rate). The good agreement between the actual and predicted counts indicates that the shot noise of the scattered light itself is the dominant noise source in the system. I was unable to make the test chamber clean enough to measure false counts out to quite 7.1σ, which with the single threshold used for the test, corresponds to one count every 2 days, so the last data point (2 counts in 69 h) was taken with the beam shining into the chamber but with the scan turned off. False counts at high σ (with the scan on or off) are quite distinctive looking, being a single peak shaped roughly

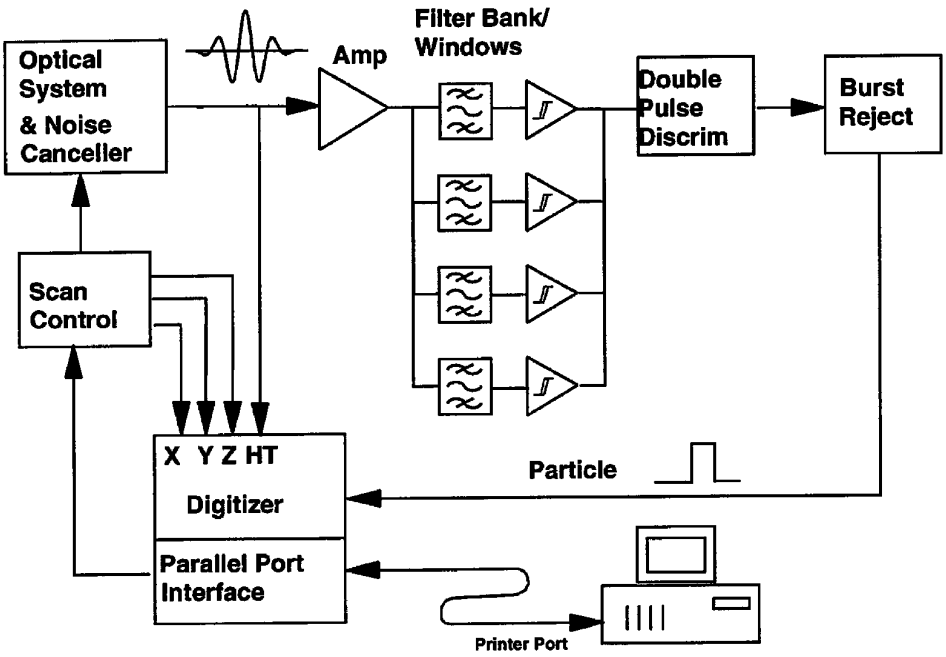

Fig. 4. Block diagram of the signal-processing system, showing the noise canceller, filter banks, FAR trackers, and burst-elimination circuits.

like the impulse response of the filter used; this shape is quite different looking from a real particle signal during scanning, which is typically a few cycles of tone burst, so the discrimination can be made with good confidence by eye. The distinctive shape is due to the fact that a wider or taller pulse, or two large excursions separated by more than the correlation time of the filter, are even more unlikely than a single high-sigma event. The ease of distinguishing these events is due to the false-count measurement having been done in a band from nearly dc to 1 MHz, rather than with a filter bank as in the actual signal processor, where the impulse responses are much more like true tone bursts in shape. (This matched filter response is necessary for best sensitivity.)

9. False-Alarm Rate Tracker System

Most lasers exhibit some power variation over their useful lives. This is due to gradual degradation of their optical performance, drifts in their power supplies, contamination of the optics, and other factors. In addition, the gains of amplifiers and the quantum efficiencies of photodiodes are not constant. In a thresholding operation, it is essential to set a high enough threshold that the sensor does not report erroneously high particle counts, possibly resulting in needless downtime for the processing tool being monitored. At the same time, it is economically important to use the available laser power as efficiently as possible; the laser used in this sensor costs over $1200, so that (loosely speaking) too-high thresholds cost $400 per decibel. The signal-processing strategy is to set separate bipolar thresholds for each band, using an automatic thresholding circuit. This circuit exploits the accurately known noise amplitude statistics to servo on the false counts themselves and ignore signal pulses, however large they may be. In radar applications, this is known as a constant FAR(CFAR) servo; the technologies employed are quite different, however, since a radar system can look at a given target many times, and its noise is very non-Gaussian. In the present system, the FAR tracker can accurately maintain the FAR at a level well below the true count rate in most applications.

A simplified block diagram of the FAR tracker[16] appears as Fig. 6. Two comparators are used to compare the instantaneous signal voltage to each of two thresholds, which are held in a fixed ratio by a voltage divider. In the absence of true signals, the noise is Gaussian to high accuracy, as shown above, so that if the FAR at one threshold and the bandwidth are known, the FAR at the other threshold can be predicted with good confidence. The lower threshold V_L is chosen to have a FAR that is small compared with the bandwidth but large compared with the expected rate of threshold crossings due to real particles; 5 kHz is used here. At this rate, true signals can be ignored compared with the false counts, so a good estimate of the rms shot-noise voltage can be made; the false alarm pulses from the lower comparator, C_2, are fed to a frequency-to-voltage converter, whose output is compared with a reference voltage. The error signal drives an integrating servo amplifier, which adjusts the thresholds to keep the FAR at C_2 at 5 kHz. Provided that the bandwidth does not change, this guarantees that the FAR at comparator C_1 is 10^{-5} Hz, even though this is likely to be much smaller than the rate of true counts at this output. Thus no laser power is wasted by a needlessly high threshold, and no imaginary disasters are reported as a result of a dangerously low one. Many measurements require fixed thresholds, so the sys-

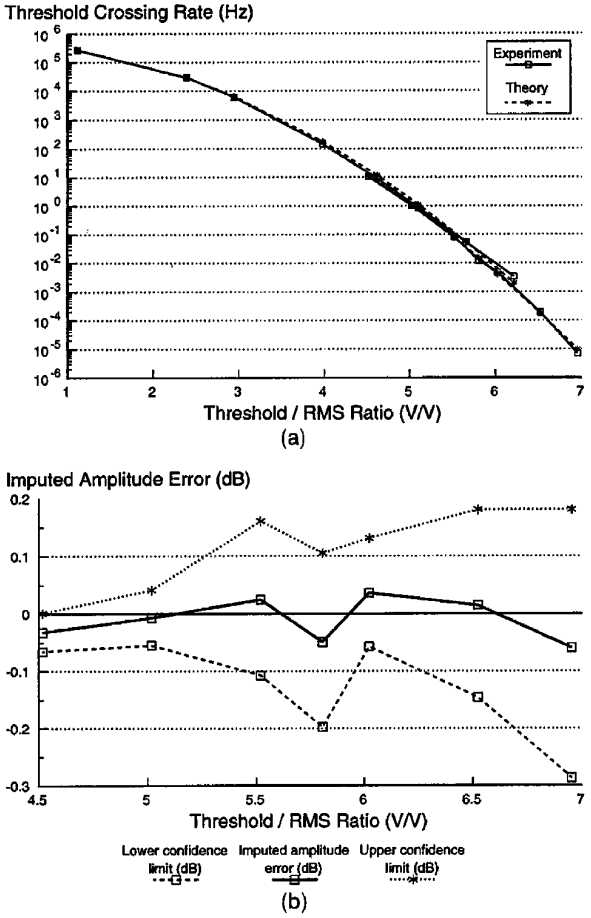

Fig. 5. (a) FAR of the particle counter, compared with the theoretical prediction. Here the beam is scanning a clean chamber, whose back wall is made from sandblasted aluminum, which is the most difficult wall material found to date. Without care being taken in the optical and laser system design, the false-count rate often increases greatly when the scan is turned on (depending on the distance to the back wall and on the material the wall is made of), but this effect has been substantially eliminated in the current system. Two data sets are superposed here, taken with two different thresholding devices (a digital oscilloscope and a frequency counter) that appear to have had different offset voltages. Nevertheless, the curves agree well over 11 orders of magnitude in FAR. I was unable to make the chamber clean enough to allow the last data point (at $\alpha = 7.0$) to be taken with the scan on, so for that point alone, the beam was stationary. The theoretical curve here was a one-parameter fit; the fitted parameter was $B/2^*$ in Eq. (2). (b) Portion of data of (a), converted to an imputed amplitude error by solution of Eq. (2) (using the same fitted value of $B/2^*$) for the values of α needed to reproduce the measured FAR exactly, then division of this result by the true α. The measured data are plotted with the upper and lower 2σ confidence limits.

tem measures the pulse height as well, allowing the computer to bin the counts repeatably. Adjustable thresholding is nonetheless necessary to allow bins near the noise limit to be used at all.

The outputs of the comparators for each band are ORed together, and the result is a logic signal that becomes active when a particle is detected. In some versions of the signal processing, additional qualification of particle events is performed, using the known differences between signal and noise pulses.

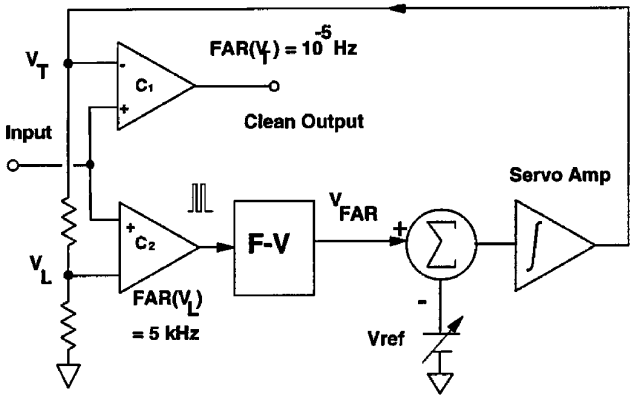

Fig. 6. Block diagram of the FAR tracker circuit. Comparators C_1 and C_2 produce positive pulses whenever the instantaneous signal voltage exceeds their thresholds, V_L and V_H, respectively, which are held in a fixed ratio by a resistive voltage divider. The desired ratio is a function of bandwidth and the amplitude statistics of the noise and is calculated so that when the FAR(V_H) at C2 is at the desired level (approximately 10^{-5} Hz), the FAR(V_L) of C_1 is approximately 5 kHz. The value of FAR(V_L) is chosen to exceed the largest expected true count rate, while being much less than the false-count rate at a threshold of 0 (which is approximately equal to half the bandwidth); this allows the assumption that all counts at C_2 are false counts. A servo system consisting of a frequency-to-voltage converter and an integrating servo amplifier keeps the average value of FAR(V_L) at 5 kHz, regardless of changes in signal level, while ignoring the occasional true counts (even though they may be very much larger in amplitude). Providing that the amplitude statistics are accurately known, this circuit allows precise control of the FAR in C_2 at a level far below the expected true count rate.

10. Experimental Results

The theoretical sensitivity of the ISICL counter is shown in Fig. 7. The horizontal dashed lines show the ultimate sensitivity of the two versions of the counter, and the three curves are the differential scattering cross sections $\partial\sigma/\partial\Omega$ for spheres of three different materials. Polystyrene latex (PSL) has a refractive index of 1.52, close to those of SiO_2 and glass. Figure 8 shows a typical tone burst from a 0.8-μm-diameter PSL sphere, together with cursors showing the calculated peak-to-peak voltage for a particle passing through the sensing volume. The two agree to within less than 10%, indicating that the theory correctly predicts the signal size and signal-to-noise ratio. These data were taken by the use of a system with an elliptical laser beam, so that the N.A. along the major axis of the ellipse was 0.0082, while the N.A. along the minor axis was 0.0025. This results in a sensitivity reduction of approximately tenfold in $\partial\sigma/\partial\Omega$, with an increase of the same order in the instantaneous sample volume. Because of mistuning of optical coatings on the window, scan mirrors, and photodiodes, the total photon efficiency of the detection side was 0.64.

In a somewhat later system, currently under development, these technical factors have been improved, yielding an improvement of several decibels. The improvements are primarily circularizing the beam, substantially eliminating the remaining traces of

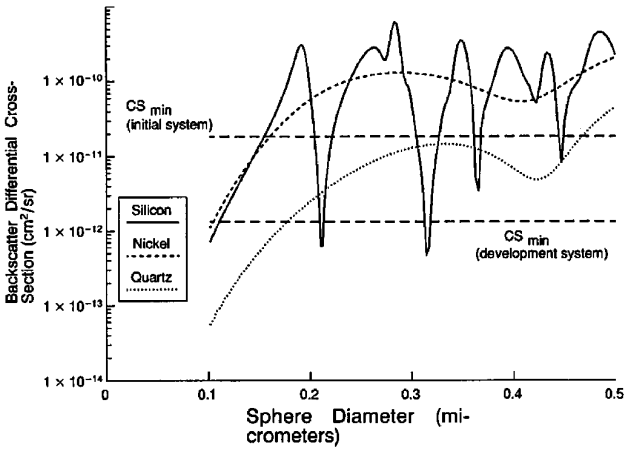

Fig. 7. Theoretical ultimate sensitivity of the two versions of the ISICL, 2×10^{-11} cm^2/sr for the elliptical beam version discussed here (with N.A.'s of 0.008 and 0.0025 along the major and minor axes, respectively), and 10^{-12} cm^2/sr for the improved version (N.A. 0.008) that is currently under testing. The ultimate sensitivity is that for which the volumetric sampling rate goes to zero. These sensitivities are plotted together with the differential scattering cross sections of variously sized spheres of silicon ($n = 3.5 + i0.07$), nickel ($n = 1.5 + i3.26$), and glass ($n = 1.52$). The calculations were done with Mie theory.

noise due to mode hopping, and increasing the photon efficiency through better coatings and photodiodes. Preliminary tests indicate that the system sensitivity has improved to 1×10^{-12} cm$^{2/\mathrm{sr}}$, yielding a particle size cutoff (where the volumetric sampling rate goes to zero) of approximately 0.12 μm for silicon or metal, or 0.18 μm for low dielectrics such as glass, and a practical volumetric sampling rate of 65 ml/min for 0.34-μm-diameter PSL spheres, or 150 ml/min for 0.62-μm PSL spheres. Work is underway to quantify the performance of the new system in sensitivity,

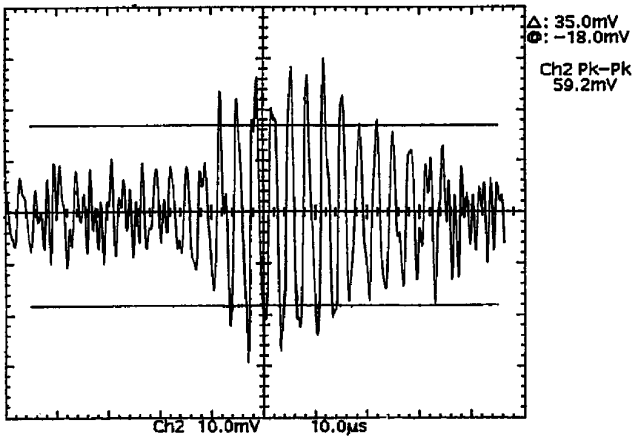

Fig. 8. Typical tone burst, resulting from an 0.8-μm-diameter polystyrene latex sphere, with cursors showing the peak-to-peak signal amplitude predicted by Eq. (A9). In order to show clearly the shape and amplitude of the tone burst, this record was taken with a much wider bandwidth than is used in the actual signal processor. In addition, the system had an elliptical beam and poor (0.64) photon efficiency. Both of these contribute to an apparent signal-to-noise ratio that is unrealistically poor.

volumetric sampling rate, and tolerance of dirty windows, and in application to counting particles in liquids. These tests are preparatory to integrating the ISICL into a semiconductor-processing line in for field testing. Further enhancements are planned, such as multiple detectors to increase the number of photons collected and thereby increase the sensitivity and sampling rate.

11. Discussion

The ISICL sensor has the unique capability of counting and mapping particles directly above the wafer during processing, while remaining totally noninvasive. It makes minimal demands on the geometry of and optical access to the chamber; a single, ordinary O-ring type side window as supplied with the tool is satisfactory. The current prototype version of the tool achieves a volume sampling rate of 65 mL/min for 0.2-micrometer spheres of silicon or metal, and 0.34-micrometer spheres of low dielectrics such as polystyrene latex or quartz, rising to 150 mL/min for 0.62-micrometer PSL spheres, and higher still for larger particles. Although extensive tests in liquids have not been performed, the system should work there as well, with performance degraded only by the presence of bubbles in the liquid and the astigmatism caused by a thick region of dielectric material at sharp scan angles. The system as it stands shows excellent immunity to back-wall reflections and window scatter, and it does not suffer from corrosion or deposition of material during processing (provided the window remains reasonably clean). It is robust and trouble free, needing no fine alignment, and can be calibrated easily in the field.

For semiconductor applications, it is the particles that actually land on the wafer that cause yield loss. In order to establish the economic value of this sensor, a predictive statistical relation between the measured particle counts in the processing chamber and the number of particles added to the wafer during the process must be found. Prototype hardware and software development are being completed, in preparation for an extended manufacturing test of the ISICL sensor, which is planned for later in 1994.

Appendix A: Signal Level Calculation

The amplitude of the signals received by the ISICL sensor depends on the laser power, N.A., particle position and velocity, detection bandwidth, and other factors. The ISICL's sensitivity is determined by the minimum number of photons it requires in order to detect a particle. For a particle in the center of the sensitive region, the received power is the product of the transmit beam intensity times the differential scattering cross section $\partial\sigma/\partial\Omega$ of the particle times the detector solid angle Ω_d. Since the shot-noise level is equivalent to one coherently detected noise photon in the measurement time (1 s^{-1} Hz^{-1}), it is convenient to work in units of photons, since that yields the signal-to-noise ratio directly, independent

of the signal and filter bandwidth. Initially we ignore the losses imposed by the matched filter.

If the scattered field is Ψ_{scat}, and the LO field is Ψ_{LO}, then the total photocurrent i_p is given by

$$i_p = K \iint_{\text{det}} \mathrm{d}^2x |\Psi_{\text{scat}} + \Psi_{\text{LO}}|^2 = K(i_{\text{LO}} + i_{\text{scat}} + i_{\text{ac}}), \tag{A1}$$

where K is a constant and, i_{LO}, i_{scat}, and i_{ac} are the photocurrents due to the LO beam alone, the scattered beam alone, and the interference of the two, respectively; these photocurrents are given by

$$i_{\text{LO}} = K \iint_{\text{det}} \mathrm{d}^2x \Psi_{\text{LO}} \Psi_{\text{LO}}^*, \tag{A2}$$

$$i_{\text{scat}} = K \iint_{\text{det}} \mathrm{d}^2x \Psi_{\text{scat}} \Psi_{\text{scat}}^*, \tag{A3}$$

$$i_{\text{ac}} = 2K \,\text{Re}\left\{ \iint_{\text{det}} \mathrm{d}^2x \Psi_{\text{LO}} \Psi_{\text{scat}}^* \right\}. \tag{A4}$$

For a Gaussian transmit beam of power P at wavelength λ, focused at a numerical aperture NA, the photon flux at the beam waist is

$$F(P, \lambda, NA) = \frac{2\pi NA^2 P}{\lambda hc}. \tag{A5}$$

Assuming that the scattered field is constant over the detector aperture, Eq. (A4) predicts that the effective detector solid angle is

$$\Omega_{\text{d}} = \pi NA^2, \tag{A6}$$

and so the expected number of photons detected per second is

$$\langle n_0 \rangle = \frac{2\pi^2 NA^4 P}{\lambda hc} \frac{\partial \sigma}{\partial \Omega}. \tag{A7}$$

The sample volume is defined by the Gaussian beam shape. For a circular beam with numerical aperture NA and crossing angle ϕ, the beam waist is $w_0 = \lambda/(\pi \text{NA})$, and [assuming that $\phi \ll 1$ and ignoring a fixed phase factor $\exp(i\phi)$], the ac photocurrent due to a particle at the center of the beam is given by

$$i_{\text{ac}} \approx 2(i_{\text{LO}} q Q n_0)^{1/2} \times \exp\left\{ i2\pi f_d t - \left[\frac{2(x^2 + y^2)}{{w_0}^2} + \frac{z^2\phi^2}{2{w_0}^2} \right] \right\}, \tag{A8}$$

where Q is the quantum efficiency and f_d is the Doppler frequency.

The beam scans in the x direction at velocity v_x, so a shift in x is equivalent to a time shift; without loss of generality, for a particle at $(0, y, z)$, the complex detected photocurrent is

$$i_{\text{ac}}(t) = 2(i_{\text{LO}} e Q N_0)^{1/2} \exp\left[i2\pi f_d t - \left(\frac{2{v_x}^2 t^2}{{w_0}^2} \right) \right] \times \exp\left[-\left(\frac{2y^2}{{w_0}^2} + \frac{z^2\phi^2}{2{w_0}^2} \right) \right]. \tag{A9}$$

which is (for particle velocities much less than v_x) a Gaussian envelope tone burst whose carrier is the Doppler frequency and whose amplitude is set by the interference between the scattered and LO light.

The number of photons that are required for a burst big enough to detect is equal to the square of the required α, because the ISICL measures the amplitude rather than the intensity of the scattered light. A matched filter imposes a 3-dB signal loss on a Gaussian pulse,[14] but since the measurement detects threshold crossings, this 3-dB is made up by the factor of $\sqrt{2}$ gain from the peak-to-average ratio, so that the minimum detectable number of photons (in the deterministic approximation) is α^2.

The author wishes to thank J. S. Batchelder, D. M. De Cain, D. Guidotti, K. L. Haller, and M. A. Taubenblatt for fruitful discussions and practical aid, and P. De Maria, W. E. Gaudelli, and W. B. Hildenbrand for creative model making, printed-circuit design, and mechanical design, respectively.

References

1. J. E. Stern and K. B. Albaugh, "Correlation issues with *in-situ* particle monitoring," in *Microcontamination '92 Proceedings* (Canon Communications, Santa Monica, Calif., 1992), pp. 401–409.
2. J. S. Batchelder and M. A. Taubenblatt, "Interferometric detection of forward scattered light from small particles," Appl. Phys. Lett. **55,** 3 (1989).
3. R. N. Carlile, S. Geha, J. F. O'Hanlon, and J. C. Stewart, "Electrostatic trapping of contamination particles in a process plasma environment," Appl. Phys. Lett. **59,** 1167–1169 (1991).
4. R. M. Roth, K. G. Spears, G. D. Stein, and G. Wong, "Spatial dependence of particle light scattering in an rf silane discharge," Appl. Phys. Lett. **46,** 253–255 (1985).
5. M. J. McCaughey and M. J. Kushner, "A model for particulate contaminated glow discharge," J. Appl. Phys. **69,** 6752–6761 (1991).
6. G. S. Selwyn, J. S. McKillop, K. L. Haller, and J. J. Wu, "*In situ* plasma contamination measurement by He–Ne laser light scattering: a case study," J. Vac. Sci. Technol. A **8,** 1726 (1990).
7. P. G. Borden and L. A. Larson, "Benefits of real time *in situ* particle monitoring in production medium current ion implantation," IEEE Trans. Semicond. Manuf. **2,** 141–145 (1989).
8. P. Geraghty, J. McInerny, and P. Borden, "Strategies for real-time particle control in a cvd system," in *1990 Proceedings of the Institute of Environmental Sciences* (Institute of Environmental Sciences, Mt. Prospect, Ill., 1990), pp. 390–392.
9. P. Borden, "Surface defects result largely from pumpdown in hot-wall cvd systems," Microcontamination **9,** 11 (1991).
10. J. R. Monkowski and D. W. Freeman, "Real-time contamination monitoring of process equipment: the challenge of the

90s," *1990 Proceedings of the Annual Technical Meeting of the Institute of Environmental Sciences* (Institute of Environmental Sciences, Mt. Prospect, Ill., 1990), pp. 393–395.
11. U. Koch, H. Kovach, and C. Michael, "Particle measurements in gas system components," *1990 Proceedings of the Annual Technical Meeting of the Institute of Environmental Sciences* (Institute of Environmental Sciences, Mt. Prospect Ill., 1990), pp. 320–327.
12. A. V. Jelalian, *Laser Radar Systems* (Artech, Boston, Mass., 1992), pp. 33–41.
13. P. C. D. Hobbs, "Reaching the shot noise limit for $10," Opt. Photon. News **2**(4), 17–23 (1991).
14. J. W. Taylor and M. Skolnik, ed., "Receivers," in *Radar Handbook*, 2nd ed. (McGraw-Hill, New York, 1990), pp. 3.21–3.22.
15. A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 2nd ed. (McGraw-Hill, New York, 1984), pp. 345–348.
16. P. C. D. Hobbs, "Noise cancelling circuitry for optical systems with signal dividing and combining means," U.S. patent 5,204,631 (20 April 1993).