

# OPTIMAL SIGNAL PROCESSING IN CAVITY RING-DOWN SPECTROSCOPY

Kevin K. Lehmann *and* Haifeng Huang

---

## Contents

1. Introduction	624
2. The Model	625
3. Generalized Least Squares Fit with Correlated Data	628
4. Weight Matrix for Model of Cavity Ring-Down Data	630
5. Detector Noise Limited Cavity Ring-Down Data	633
5.1. To average and then fit, or fit each decay and average the fit results?	638
6. Linearization of the Fit in Cavity Ring-Down Spectroscopy	638
7. Determination of Ring-Down Rate by Fourier Transform Method	641
8. The Successive Integration Method for Exponential Fitting	643
9. Analog-Detected Cavity Ring-Down	646
9.1. Phase shift method	646
9.2. Gated integrator method	649
9.3. Logarithm-differentiator method	650
10. Shot Noise Limited Cavity Ring-Down Data	651
11. Effect of Residual Mode Beating in the Ring-Down Decay	655
12. Conclusions	657
Acknowledgments	657
References	657

## Abstract

In this chapter, the authors present a systematic statistical analysis of cavity ring-down signal extraction. The traditional uncorrelated least squares fit can be generalized to the situation with data correlation (e.g. caused by data filtering, which is essential to minimize noise). If the data is sufficiently highly sampled, the effect of the data correlation can be included by introducing an effective variance of the data. This correction has substantial influence on the estimation of the standard error of the fit parameters for correlated data, especially for the fitted decay rate  $k'$ , because this determines the final sensitivity. For both the white noise dominated and the shot noise dominated situations, the sensitivity limit is given. The authors found that the bias of  $k$  in the white noise situation is normally very small and can be neglected. The authors also compared several commonly used alternative algorithms in cavity ring-down community. These methods include linearized weighted least squares fit, determining the decay rate by Fourier transform, corrected successive integration (CSI)

method, and several analog methods in extracting  $k$  from decay signal. The bias, dispersion prediction (including optimum results), and speed of these methods are discussed in detail. Among all these methods, the least square fit gives the smallest dispersion estimation of fit parameters. Considering the special properties of exponential function, the authors found that the least squares fit has lower computation cost than the Fourier transform and CSI methods. Lastly, the effect of residual mode beating on the extracted  $k$  is also analyzed.

**Keywords:** cavity ring-down spectroscopy; signal processing; data correlation; optimal signal processing; cavity-enhanced spectroscopy; generalized least squares fit; white noise; fourier transform method; successive integration method; analog detected cavity ring-down; shot noise; residual mode beating



## 1. INTRODUCTION

Since the seminal work of O'Keefe and Deacon [1], cavity ring-down spectroscopy (CRDS) has become an important method for obtaining highly sensitive absolute absorption spectra of weak transitions or rarefied species. Twenty years after its introduction, the popularity of this and related methods that use high-finesse optical cavities continues to grow. Our bibliographic database now contains over 900 entries, the majority of entries from the last 4 years alone. The work in the field up to 1999 was well represented in a collective volume edited by Busch and Busch [2]. In 2008, another collective volume is due to be published, edited by Berden and Engel [3]. There have been a number of excellent reviews published over the years, including Refs [4,5] and most recently Ref. [6].

In standard CRDS, a spectrum is obtained by observing changes in the optical decay rate of a high-finesse optical cavity as a function of excitation frequency. It therefore follows that the sensitivity of CRDS is directly proportional to the stability and signal-to-noise ratio of the cavity decay rate, which can be extracted in a number of different ways from the ring-down transient of cavity output power as a function of time. Many different methods have been described for determining the cavity decay rate, which is not surprising given the rich variety of lasers that have been used and samples that have been studied.

Starting with Romanini and Lehmann [7], several papers have reported the analysis of the expected signal-to-noise ratio of the cavity decay rate with different assumptions about the data analysis method and the statistical character of the decay. However, this chapter presents what the authors believe to be the most systematic investigation to date into the statistical analysis of cavity ring-down signal extraction. We review and derive sensitivity limits for all the methods commonly used to extract cavity decay rates in the CRDS community. In Section 2, we lay out the mathematical model of the data, including the data correlation that arises from a low pass filtering of the data. Such filtering is required to obtain an optimized signal-to-noise ratio and thus minimize statistical fluctuations in the extracted cavity decay rate, which determines the sensitivity of a CRDS experiment. Such data correlations are ignored in standard least squares fit treatments. Section 3 sets out the form of the least squares fit

problem with data correlation [8]. Section 4 derives the tridiagonal fit weight matrix that arises for exponential data correlation as is created by a low-pass filter. For CRDS data under cases where the data are sufficiently highly sampled, the least squares fit equations only change by the introduction of an effective variance of the data. This correction of the data variance is essential to correctly predict the standard error of the fit parameters, including the cavity decay rate. Section 5 presents an analysis of least squares fitting of CRDS data in the case where detector and other forms of intensity independent noise dominate, including derivation of the sensitivity limit that can be expected. Section 6 examines the weighted least squares fit that results from linearization of the problem by calculation of the log of the data points. In high signal-to-noise situations, this gives the same sensitivity as the direct least squares fit to the data, but the non-linear transformation introduces some bias that become important at modest signal-to-noise ratios. At low signal to noise, the method introduces substantial bias [9]. Sections 7 and 8 examine the noise and bias expected when the cavity decay rate is determined using the Fourier transform [10] and the successive integration (SI) [11] methods, respectively. These are two noniterative methods that do not require prior knowledge of the baseline and are being used in the CRDS community as alternatives to the least squares fitting of the data. In both cases, the predicted noise in the decay rate is moderately higher than for the least squares fit to the data. We do not find that these methods offer computational advantages over an optimized least squares fitting program. Section 9 examines three different analog processing approaches for the determination of the ring-down rate. The first is the phase shift method [12,13], which uses lock-in demodulation of the cavity transmission; the second calculates the decay rate from the ratio of signal sampled by two gated integrator windows [7]; and the last uses a log-amplifier followed by a differentiator [14]. These methods are useful for cases of very high cavity decay event rates, where all the individual decays can no longer be digitized and fit, though the last appears to have a poorer sensitivity limit. Section 10 presents an analysis of the standard least squares fit in the case where shot noise in the ring-down signal dominates. This proves to be a particularly simple case where the solution of the least squares fit equations proves to be equivalent to taking the normalized first moment of the signal transient. Section 11 considers the effect of residual mode beating in the ring-down decay upon the expected sensitivity. Such beating is usually unavoidable when using a pulsed laser to excite the ring-down cavity. Section 12 presents some closing remarks.

## 2. THE MODEL

We will begin with a model for the signal observed in a cavity ring-down experiment. Let  $\gamma(t)$  be the time dependent optical power on the detector. We expect the signal, expressed in watts, to be of the following form:

$$\gamma(t) = F(t) + \epsilon(t) = Ae^{-kt} + B + \epsilon(t), \quad (1)$$

where  $\epsilon(t)$  is the noise term. We will start by assuming that we have white noise with spectral density  $\sigma'(t)$  (units of watts per  $\sqrt{\text{Hz}}$ ), which implies the following statistical properties:

$$\begin{aligned} \langle \epsilon(t) \rangle &= 0 \\ \langle \epsilon(t) \cdot \epsilon(t') \rangle &= \sigma'^2(t) \delta(t - t'), \end{aligned} \quad (2)$$

where  $\langle \rangle$  denotes an ensemble averaging over a large number of macroscopically identical experimental runs.  $A$  is the amplitude of the ring-down signal.  $k$  is the ring-down rate and is typically in the range  $(0.001 - 1) \cdot 10^6 \text{ s}^{-1}$  in most CRDS experiments.  $\tau = k^{-1}$  is the decay time constant for the cavity.  $B$  represents a baseline signal after the light inside the cavity has decayed to a negligible value. In cases where  $B$  is stable with time, it will not be treated as a fit parameter, and its value can be subtracted from the data prior to fitting. In any real experiment, we will not have an infinite bandwidth for our detector. Furthermore, to optimize our signal-to-noise ratio, we may want to further limit the detection bandwidth. We assume that we have a single pole low-pass filter with a time constant  $t_f = 1/k_f$ . This leads to the following redefinition of the time dependent signal:

$$\bar{y}(t) = \int_0^\infty k_f y(t - \tau) \exp(-k_f \tau) d\tau = \bar{F}(t) + \bar{\epsilon}(t) \quad (3)$$

with analogous definitions for  $\bar{F}(t)$  and  $\bar{\epsilon}(t)$ . If we further assume that  $\sigma'(t)$  is effectively constant over time intervals  $\sim t_f$ , then it can be shown from these definitions that our averaged noise has the following correlations:

$$\langle \bar{\epsilon}(t) \cdot \bar{\epsilon}(t') \rangle = \frac{k_f}{2} \sigma'^2 \left( \frac{t + t'}{2} \right) \cdot \exp(-k_f |t - t'|). \quad (4)$$

Thus, the time constant leads to a “memory” of  $t_f$  for the noise. The form for  $\bar{F}(t)$  depends on what we assume for the signal for negative time. Two choices are  $F(t < 0) = B$ , which would be a model for excitation of the ring-down cavity by a short pulse at  $t = 0$ , and  $F(t < 0) = A + B$ , which is a model for quasi-cw excitation of the ring-down cavity. In these two cases we have:

$$\bar{F}(t > 0) = \begin{cases} \frac{k_f}{k_f - k} A (e^{-kt} - e^{-k_f t}) + B & \text{if } F(t < 0) = B \\ \frac{k_f}{k_f - k} A \left( e^{-kt} - \frac{k}{k_f} e^{-k_f t} \right) + B & \text{if } F(t < 0) = A + B. \end{cases} \quad (5)$$

In most CRDS experiments, the very beginning of the decay is not fit because it is distorted by various effects, including the finite response time of the detector and electronics. Regardless, as long as  $k_f \gg k$  (i.e.  $\tau \gg t_f$ ), which will typically be the case, then the extra term in  $\bar{F}(t)$  will quickly decay (compared to the ring-down time), and thus we can fit most of the decay without needing to include the corrections to a single exponential decay. For pulsed excitation in CRDS, one has

100% amplitude modulation of the signal on the 1–10 nsec time scale and the use of a higher order filter to drastically reduce this modulation without cutting into the detection bandwidth is advised. However, as long as  $t_f$  is the shortest filter time constant, the present treatment should give a good approximation.

We will now consider that the ring-down transient is digitized by an analog-to-digital (A/D) convertor with  $\bar{y}(t)$  sampled at equally spaced times  $t_i = i \cdot \Delta t$  with  $i = 0 \dots N - 1$ . Spence et al. [14] discussed some of the effects that A/D nonidealities can introduce, but we will not consider them further as our experience is that these are not important if the root-mean-squared (RMS) noise input of the A/D is greater than the least significant bit. To most faithfully reproduce the ring-down signal, we want  $k\Delta t \ll 1$ .  $\Delta t$  is usually determined by the hardware limitations. An important practical trade-off is between the dynamic range (how many bits of resolution) versus speed of the digitizer. Even without the hardware constraints, the memory and computational cost to determine the ring-down rate from a decay scales inversely with  $\Delta t$ , and thus oversampling may limit the number of decay per second that can be processed. Below, we will examine the sensitivity of the determination of the ring-down rate to the value of  $k\Delta t$ . We will denote the set of measured values as  $y_i$ , or collectively as  $\vec{y}$ . Likewise, we define  $F_i = \bar{F}(t_i)$  and  $\epsilon_i = \bar{\epsilon}(t_i)$ , with  $\vec{F}$  and  $\vec{\epsilon}$  denoting the vector of values.

We also need a model for the noise density. The two intrinsic noise sources in CRDS are (i) shot noise and (ii) detector noise. If  $\bar{F}(t)$  is expressed in watts, we have light frequency of  $\nu$ , our detector has quantum efficiency  $Q$ , and noise equivalent power  $P_N$ , then we can write for the noise density:

$$\sigma'^2(t) = \frac{h\nu}{Q} A e^{-kt} + P_N^2, \quad (6)$$

where  $h$  is Planck's constant.  $P_N$  has contributions from the shot noise of the detector dark current and also Johnston noise in the resistor that is used to convert the photocurrent into a voltage. Thus,  $P_N$  is minimized by having that resistance large enough that the dark current noise dominates, but this will often restrict the detector bandwidth.

The measurements will have variance  $\sigma_i^2 = (k_f/2)\sigma'^2(t_i)$ . Under the assumption of slowly changing  $\sigma'^2(t)$ , we can write the covariance matrix,  $\sigma^2$ , for the data as

$$(\sigma^2)_{i,j} = \langle \epsilon_i \epsilon_j \rangle = \frac{\sigma_i^2 + \sigma_j^2}{2} e^{-|i-j|k_f \Delta t}. \quad (7)$$

In this chapter, we will consider the two limiting cases, namely shot noise limited CRDS, where the first term dominates and the baseline offset,  $B$ , from detector dark current can be neglected and detector noise limited CRDS, where the second term dominates. Of course, because of the exponential decay, the second term will dominate the first sufficiently far into the decay, but the tail of

the decay makes only a modest contribution to the fitted parameters, so no instability will be presented by the assumption of shot noise limited decay. As will be shown below, the assumption of shot noise limited decay leads to simpler expressions, including the fact that the exact least squares solutions can be written down directly.

We will now consider the statistical properties of a fit of the ring-down curve. This fit will give parameters  $A'$ ,  $k'$ , and possibly  $B'$ , which are estimates for unknown ring-down amplitude, rate, and signal baseline, of which  $k'$  is the one used in extracting the spectrum or determining analyte concentration. The obvious goal is to use estimates that are unbiased (i.e. on average will give the correct values) and have the smallest possible dispersion. If we assume that the noise has a Gaussian distribution, then these criteria are met by a properly weighted least squares fit of the data [8]. An important point, however, is that the assumption that is typically made, that the experimental data points have uncorrelated errors, is not valid unless  $k_f \Delta t \gg 1$ . However, it is possible to properly account for correlation by a minor modification of the least squares equations.

### 3. GENERALIZED LEAST SQUARES FIT WITH CORRELATED DATA

The mathematical treatments of linear and linearized nonlinear least squares fits to uncorrelated experimental data are well known and described in detail in many texts [15,16]. However, the treatment of the more general problem with correlated data is much less well known, at least in the spectroscopy community, though the proof of the general solution of the linear model has been presented by Albritton et al. [8]. In this section, the solutions for this more general least squares problem will be given in the context of fitting CRDS data, which is a nonlinear model.

If we assume that the individual  $\epsilon_i$  are Gaussian distributed, then by going to the basis set that diagonalizes the  $\tilde{\sigma}^2$  matrix [8] (in this basis set the errors are uncorrelated), it is easily shown that we can write the probability density for any given set of experimental errors as follows:

$$P(\vec{\epsilon}) = \frac{1}{\sqrt{(2\pi)^N \det(\tilde{\sigma}^2)}} \exp\left(-\frac{1}{2} \vec{\epsilon} \cdot (\tilde{\sigma}^2)^{-1} \cdot \vec{\epsilon}\right). \quad (8)$$

We will denote the fitted function by  $F'(t) = A' e^{-k't} + B'$ , and its value at the measured times as  $F'_i$ . The principle of maximum likelihood states that the fit parameters  $\{A' B' k'\}$  that best estimate the unknown decay parameters  $\{A B k\}$  are those that maximize the probability that the observed set of data will occur.

For any given parameters, we can write  $\vec{\epsilon} = \vec{y} - \vec{F}'$ . We also introduce a generalized weight matrix:

$$\tilde{W} = \left( \tilde{\sigma}^2 \right)^{-1}. \quad (9)$$

Thus, we have a generalized  $\chi^2$  function:

$$\chi^2(A', B', k') = \left( \vec{y} - \vec{F}' \right) \cdot \tilde{W} \cdot \left( \vec{y} - \vec{F}' \right), \quad (10)$$

for which we will find the minimum. Note that under the above assumption of Gaussian errors, the  $\chi^2(A, B, k)$  should follow a chi-squared distribution with  $N$  degrees of freedom, whereas  $\chi^2(A', B', k')$  should follow a chi-squared distribution with  $N - 3$  degrees of freedom. The generalized least squares equations, found by minimizing  $\chi^2$  with respect to the parameters  $p_j = \{A', B', k'\}$ , are as follows:

$$\left( \vec{y} - \vec{F}'(A', B', k') \right)^\dagger \cdot \tilde{W} \cdot \widetilde{\nabla F} = \vec{0} \quad (11)$$

$$\nabla_{F_{ij}} = \left( \partial \frac{F'_i}{\partial p_j} \right). \quad (12)$$

Note that we do not have derivatives of the  $\tilde{W}$  matrix with respect to fit parameters even if, because of shot noise, the noise of the signal is intensity dependent. In the limit of a diagonal  $\tilde{W}$ , this reduces to the standard least squares equations found in numerous texts [15,16]. Given the solution to the above equations and assuming linear behavior near the solution (i.e.  $\frac{\partial \vec{F}'}{\partial p_j}$  are constant), we can write the covariance matrix,  $\tilde{\tau}$ , for the fit parameters in terms of the curvature matrix,  $\tilde{\alpha}$ , by

$$\tilde{\alpha} = \widetilde{\nabla F}^\dagger \cdot \tilde{W} \cdot \widetilde{\nabla F} \quad (13)$$

$$\tilde{\tau} = \tilde{\alpha}^{-1}. \quad (14)$$

The diagonal elements of  $\tilde{\tau}$  give the variances of the fit parameters and the off-diagonal elements of the covariances  $c_{ij} / \sqrt{c_{ii}c_{jj}}$  (which is between  $-1$  and  $1$ ) gives the correlation coefficient between the  $i^{\text{th}}$  and  $j^{\text{th}}$  parameters. The least squares equations can be solved iteratively, given an initial estimate for the parameters and by calculating the change that will solve the equations under the assumption of linearity:

$$\vec{\delta p} = \tilde{\tau} \cdot \vec{\beta} \quad \text{with} \quad \vec{\beta} = \widetilde{\nabla F}^\dagger \cdot \tilde{W} \cdot \left( \vec{y} - \vec{F}' \right). \quad (15)$$

This equation can also be used, with  $\vec{y} - \vec{F}'$  replaced by  $\vec{\epsilon}$ , to give the sensitivity of the parameters to a perturbation of the input data.

Because of parameter correlation, the standard error of  $k$  will be higher for a fit that adjusts the baseline parameter,  $B$ . Often, the detector and electronics offsets are more stable with time than the accuracy to which  $B$  can be determined in a fit to a single decay. In these cases, lower fluctuations in the fitted  $k'$  values are expected if only the two parameters  $A'$ ,  $k'$  are floated in the fit. This is done simply by subtracting the fixed value  $B$  from the experimental data and then removing the  $B$  row and column from the curvature matrix,  $\tilde{\alpha}$ .

If we have an incorrect model for the variances and covariances of the noise, the predicted weight matrix,  $\tilde{W}'$ , will be in error. How do we estimate the effect that this will have on the fit results? The mean value of the fit parameters will still be the true parameters, but the covariance matrix calculated from Equation (14) using  $\tilde{W}'$  (let us call it  $\tilde{c}'$ ) will not give correct estimates for the true uncertainties in the fit parameters. If we neglect data correlation (set off-diagonal elements of the weight matrix to zero), the resulting  $\tilde{c}'$  gives smaller estimates for the parameter standard errors than predicted by  $\tilde{c}$ . However, in general, we expect the true covariance matrix for an improperly weighted fit, which we will denote by  $\tilde{c}''$ , to give larger uncertainties than those of the appropriately weighted fit,  $\tilde{c}$ , as it is known that the properly weighted least squares fit is the lowest dispersion linear estimator of the unknown parameters [8]. By using the standard linearized equation for error propagation,

$$c''_{ab} = \langle \delta p_a \delta p_b \rangle = \sum_{ij} \left( \frac{\partial p_a}{\partial y_i} \right) \sigma_{ij}^2 \left( \frac{\partial p_b}{\partial y_j} \right) \quad (16)$$

combined with the modified version of Equation (15) ( $\tilde{W}$  replaced by  $\tilde{W}'$ ), we have

$$\begin{aligned} \tilde{c}'' &= \tilde{U}^\dagger \cdot \tilde{\sigma}^2 \cdot \tilde{U} \\ \tilde{U} &= \tilde{W}' \cdot \tilde{\nabla} F \cdot \tilde{c}' \\ \tilde{c}' &= \left( \tilde{\nabla} F^\dagger \cdot \tilde{W}' \cdot \tilde{\nabla} F \right)^{-1}. \end{aligned} \quad (17)$$

The matrix  $\tilde{c}'$  is the covariance matrix we would calculate assuming  $\tilde{W}'$  is the weight matrix. It can easily be shown that  $\tilde{c}'' \rightarrow \tilde{c}$  if  $\tilde{W}' = \text{constant} \times \tilde{W}$ . Equation (17) will be used below to calculate the statistical dispersion of the fit parameters for approximate fit models that neglect parameter correlation.

#### 4. WEIGHT MATRIX FOR MODEL OF CAVITY RING-DOWN DATA

Returning to our specific problem, we need to evaluate the various terms in the above generalized least squares expressions. For the equally time spaced points and again assuming slowly changing noise variance, we can analytically invert the covariance matrix to give the weight matrix. Let us define  $a_f = \exp(-k_f \Delta t)$ , then we have

$$W_{i,j} = \frac{2}{(\sigma_i^2 + \sigma_j^2)} \times \begin{cases} \frac{1 + a_f^2}{1 - a_f^2} & \text{if } i=j, i \neq 0, N-1 \\ \frac{1}{1 - a_f^2} & \text{if } i=j, i=0, N-1 \\ -\frac{a_f}{1 - a_f^2} & \text{if } i=j \pm 1 \\ 0 & \text{otherwise.} \end{cases} \quad (18)$$

If we have a function,  $f(t)$ , that changes slowly over time interval  $\Delta t$ , we can write

$$\left(\tilde{W}\vec{f}\right)_i \approx \frac{1}{\sigma_i^2} \left[ \left(\frac{1 - a_f}{1 + a_f}\right) f(t_i) - \left(\frac{a_f \Delta t^2}{1 - a_f^2}\right) \left(\frac{d^2 f}{dt^2}(t_i)\right) \right]. \quad (19)$$

The definition of  $\vec{f}$  is the same as  $\vec{y}$ . Thus, we can see that in the limit of sufficiently small sampling time,

$$|f(t)| \gg \frac{a_f}{(1 - a_f)^2} \left| \left(\frac{d^2 f}{dt^2}(t_i)\right) \right| \Delta t^2, \quad (20)$$

the effect of the data correlation can be approximated by multiplying a diagonal weight matrix by a scale factor of  $\left(\frac{1 - a_f}{1 + a_f}\right)$ , or equivalently, the variances of the data points should be multiplied by a uniform factor of

$$G = \left( \frac{1 + \exp(-k_f \Delta t)}{1 - \exp(-k_f \Delta t)} \right), \quad (21)$$

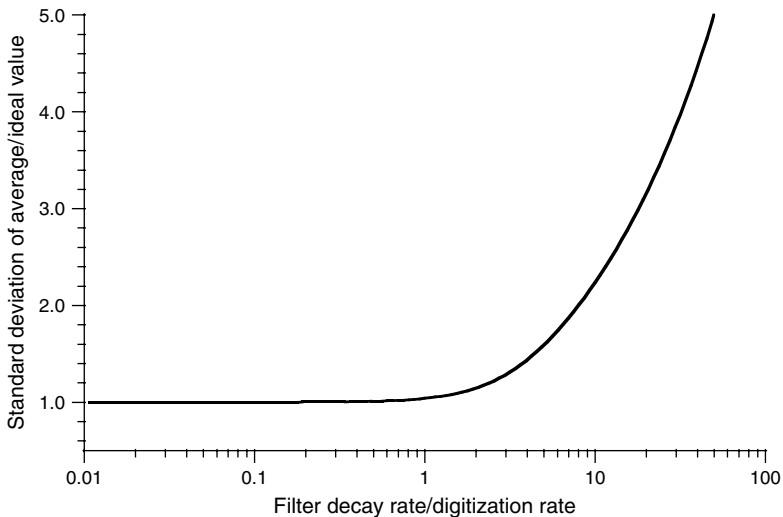
$G > 1$ . This same factor will need to be applied to the covariance matrix calculated for a least squares fit that ignores data correlation. For CRDS data,  $\left(\frac{d^2 f}{dt^2}(t)\right) \Delta t^2 = (k \Delta t)^2 f(t)$  (the baseline can be neglected for this purpose) and thus the Equation (20) holds to high accuracy for well-sampled CRDS data, for which  $k \Delta t \sim 0.01$  is typical.

An obvious question is how does the variance in our parameters changes with  $k_f$ ? We will first consider the simpler problem of averaging  $N$  data points, each separated by  $\Delta t$  for a constant signal. In this case, we can treat the noise density  $\sigma'(t)$  as constant. The best variance we can expect for the average over this interval is  $\sigma'^2/(N \Delta t)$ . It is the authors' experience that many people believe, at least intuitively, that the signal to noise will improve as the time constant,  $t_f$ , is increased to values much beyond  $\Delta t$ . However, by using the above generalized least squares equations, it can be found that the variance of the mean,  $\mu$ , is given by

$$\sigma^2(\mu) = \xi^2 \frac{\sigma'^2}{N \Delta t} = \left(\frac{k_f}{2N}\right) \sigma'^2 \left(\frac{1 + \exp(-k_f \Delta t)}{1 - \exp(-k_f \Delta t)}\right), \quad (22)$$

$\xi$  is the fractional “excess” noise in the mean over the ideal limit. In the limit that  $k_f \Delta t \ll 1$ ,  $\xi \rightarrow 1$ , which means that we have sufficiently sampled the signal to extract the maximum information. Figure 1 shows a plot of  $\xi$  as a function of  $k_f \Delta t$ . Specific values are  $\xi = 1.04, 1.15,$  and  $1.44$  for  $k_f \Delta t = 1, 2,$  and  $4$ . One way to view the increase in noise for larger values of  $k_f \Delta t$  is that the digitization process will alias the noise at all frequency components that are harmonics of  $1/\Delta t$  to zero frequency, where they will survive the averaging. Thus, we want to pick the 3-dB roll off frequency of the low-pass filter (which equals  $k_f/(2\pi)$ ) low enough that the noise contributions from these harmonics are small compared to the near DC noise density. Thus, there is simply no signal-to-noise advantage to making the time constant of the filter much longer than sampling time interval. Longer time constants will reduce the RMS noise in the data but will make the data highly correlated and thus largely redundant. The gain in the reduced noise will be almost exactly compensated for by a reduction in the effective number of independent data points. Fits that ignore the correlation of the data will predict a variance of the average that will underestimate the true variance by a factor of  $\approx k_f \Delta t/2$  when  $k_f \Delta t \ll 1$ . From the above considerations, it should be clear that for optimal signal processing in CRDS, one wants to select values of  $k_f$  and  $\Delta t$  such that  $k_f \Delta t \sim 1$  and  $k \Delta t \ll 1$ .

The authors have constructed an ensemble of 1000 fits to decays with parameters  $\sigma' = 1$ ,  $A = 100$ ,  $\Delta t = 1$ ,  $k = 0.05$ , and  $k_f = 1$ , with the data fit with and without including the correlation of the data. It was found that statistical properties of the two sets of decay rates agreed well within their uncertainty. Further, the standard deviation of the difference in decay rate for the two fits was an order of magnitude below the standard deviation of either set of fits. Thus, in this case, there is no practical difference



**Figure 1** Plot of excess noise created when digitizing and then averaging a DC signal with white noise that has been passed through a low-pass filter with a filter time constant  $1/k_f$ . The ordinate is equal to  $k_f \Delta t$  in the notation used in the text.

in the two set of fits. However, the variance in the ensemble of decay rates is only correctly predicted by the correlated least squares fit calculation. The uncorrelated fit underestimates the ensemble variance by a factor of  $\approx 2.2$ . For an ensemble with the same parameters except  $k_f = 0.2$ , the variance of the ring-down decay rates for the uncorrelated fits was found to be  $\approx 16\%$  higher than for the correlated fits. In this case, the variance predicted by the uncorrelated fits is a factor of 10 below that found for the ensemble. These simulation results were confirmed by the general calculations of the expected variance of both correlated and uncorrelated fits, which will be given below.

## 5. DETECTOR NOISE LIMITED CAVITY RING-DOWN DATA

We will now look at the predicted dispersion of the ring-down rate for different assumptions about the fit. First, we consider a fit with  $\sigma'$  constant. Ideally, this noise density arises from the intrinsic detector noise,  $P_N$ , though excess noise from the amplifier and/or digitizer (including bit quantization noise) would likely be of this form.

As demonstrated above, as long as  $k_f \gg k$ , we can replace the general least squares problem with a nondiagonal weight matrix,  $\tilde{W}$  (Equation (18)), with a simple diagonal matrix, but with the variance scaled by  $G$  (Equation (21)). This gives for the effective variance per point  $\sigma^2 = \xi^2 P_N^2 / \Delta t$ , where  $\xi$  is as defined in Equation (22) and is plotted in Figure 1. It is nearly unity for  $k_f \Delta t \leq 1$ .

In this case, we can use geometric series expressions  $\sum_{i=0}^{N-1} r^i a^i = [a(d/da)]^n [(1 - a^N)/(1 - a)]$  to write explicit closed form expressions for the elements of the curvature matrix,  $\tilde{\alpha}$ . In terms of  $a = \exp(-k\Delta t)$ , the matrix elements of the symmetric curvature matrix,  $\tilde{\alpha}$ , and vector  $\beta$  are

$$\begin{aligned}
 \alpha_{B,B} &= \frac{N}{\sigma^2} \\
 \alpha_{B,A} &= \left( \frac{1 - a^N}{1 - a} \right) \frac{1}{\sigma^2} \\
 \alpha_{B,k} &= -A\Delta t \left( \frac{a(1 - a^N)}{(1 - a)^2} - \frac{Na^N}{1 - a} \right) \frac{1}{\sigma^2} \\
 \alpha_{A,A} &= \left( \frac{1 - a^{2N}}{1 - a^2} \right) \frac{1}{\sigma^2} \\
 \alpha_{A,k} &= -A\Delta t \left( \frac{a^2(1 - a^{2N})}{(1 - a^2)^2} - \frac{Na^{2N}}{1 - a^2} \right) \frac{1}{\sigma^2} \\
 \alpha_{k,k} &= (A\Delta t)^2 \left( \frac{2a^4(1 - a^{2N})}{(1 - a^2)^3} + \frac{a^2 - (2N + 1)a^{2N+2}}{(1 - a^2)^2} - \frac{N^2 a^{2N}}{1 - a^2} \right) \frac{1}{\sigma^2}
 \end{aligned} \tag{23}$$

$$\begin{aligned}
\beta_B &= \left( \sum_i y_i - A \left( \frac{1-a^N}{1-a} \right) - BN \right) \frac{1}{\sigma^2} \\
\beta_A &= \left( \sum_i y_i a^i - A \left( \frac{1-a^{2N}}{1-a^2} \right) - B \left( \frac{1-a^N}{1-a} \right) \right) \frac{1}{\sigma^2} \\
\beta_k &= -A\Delta t \left( \sum_i y_i i a^i - A \left( \frac{a^2(1-a^{2N})}{(1-a^2)^2} - \frac{Na^{2N}}{1-a^2} \right) - B \left( \frac{a(1-a^N)}{(1-a)^2} - \frac{Na^N}{1-a} \right) \right) \frac{1}{\sigma^2}.
\end{aligned} \tag{24}$$

Note that for each cycle of the least squares fit, we have to do only three sums over the data points that can be calculated with only three floating point multiplications and three additions per data point. We need to invert  $\tilde{\alpha}$  to determine  $\tilde{c}$ . For the  $2 \times 2$  case (fixed baseline),  $c_{ii} = \alpha_{jj}/\det(\alpha)$  and  $c_{ij} = -\alpha_{ij}/\det(\alpha)$ , where  $\det(\alpha)$  is the determinant of  $\alpha$ . For the  $3 \times 3$  case,  $c_{ii} = (\alpha_{jj}\alpha_{kk} - \alpha_{jk}^2)/\det(\alpha)$  and  $c_{ij} = (\alpha_{ik}\alpha_{jk} - \alpha_{ij}\alpha_{kk})/\det(\alpha)$ , where  $ijk$  is a cyclic permutation of  $A, B, k$ . The diagonal elements of  $\tilde{c}$  give the variances (standard errors) in the corresponding parameters and the off-diagonal elements the corresponding covariances.

If we take  $N$  larger enough that the terms  $a^N$  and  $a^{2N}$  can be neglected ( $Nk\Delta t \gg 1$ ),  $\tilde{\alpha}$  is simple enough that one can write a compact, closed form expression for the covariance,  $\tilde{c}$ . This gives for the  $kk$  element

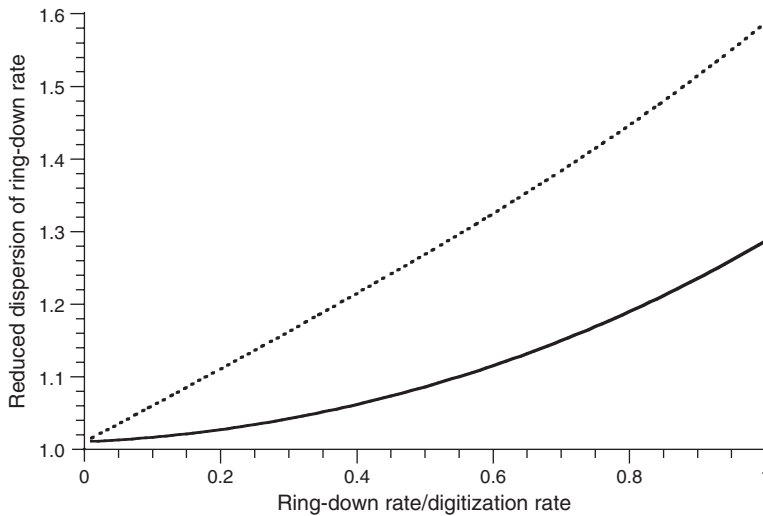
$$\sigma^2(k') = c_{kk} = \left( \frac{\xi P_N}{A} \right)^2 \frac{(1-a^2)^3}{\Delta t^3 a^2} \left[ \frac{N(1-a) - (1+a)}{N(1-a) - 2(1+a)} \right]. \tag{25}$$

In the limit that  $k\Delta t \ll 1$  and that  $\xi \rightarrow 1$  ( $k_f \Delta t < 1$ ), we approach the “ideal” or minimum value for the variance in the fitted value of the decay rate,  $k'$ ,

$$\sigma^2(k')_{\text{ideal}} = 8k^3 \left( \frac{P_N}{A} \right)^2. \tag{26}$$

Below, we will examine the convergence of  $\sigma^2(k')$  for finite experimental parameters. We will denote by the reduced variance the ratio  $\sigma^2(k')/\sigma^2(k')_{\text{ideal}}$ . The reduced dispersion is just the square root of the reduced variance.

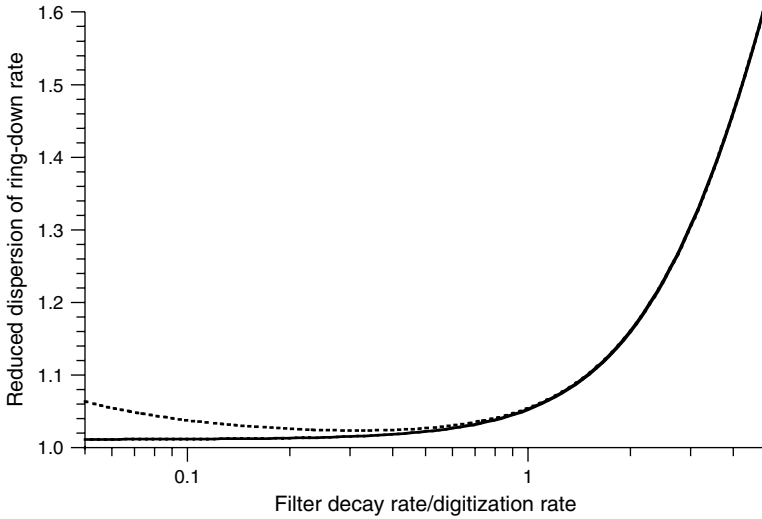
We will first assume that we are fitting well into the tail of the ring-down transient, so that  $Nk\Delta t \gg 1$ , but we have a finite value for  $k\Delta t$ . In this case, we will get the same variance for  $k'$  regardless of whether we fit a baseline in our model, because the baseline will be essentially determined by the tail of the transient. From the above curvature matrix, it can be shown that the correlation coefficient between  $k$  and  $B$  is  $\sqrt{2/(Nk\Delta t + 2)}$  and between  $k$  and  $A$  is  $1/\sqrt{2}$  in the  $k\Delta t \ll 1$  and  $Nk\Delta t \gg 1$  limit. Figure 2 shows the predicted reduced dispersion as a function of  $k\Delta t$  for both correlated and uncorrelated fits. It is important to note that here and for the rest of this chapter, these are variances calculated using Equation (17) and thus reflect influence of the data correlation, even though the fit itself is assumed to neglect data correlation. As discussed above, if we had



**Figure 2** Plot of the reduced dispersion (fractional increase in the standard deviation over the “ideal” case) of the ring-down rate as a function of the  $k\Delta t$ , which is the inverse of the number of data points per ring-down time constant. A constant value of  $k_f\Delta t = 0.5$  was used to construct this plot, and data were taken for at least eight ring-down time constants. The upper curve is for ring-down rates extracted from a fit that ignores data correlation, the lower for fits that properly account for data correlation. The ring down is assumed to be detector noise limited, that is, has constant noise.

calculated the variance using Equations (13) and (14), but with a diagonal weight matrix, we would predict variances that are too small by a factor in the order of  $1/G$ , at least for  $k \ll k_f$ . The abscissa is the inverse of the number of data points per ring-down time constant. A value of  $k_f\Delta t = 0.5$  was used to make this plot. We see from the figure that in the limit that  $k\Delta t \rightarrow 0$ , both correlated and uncorrelated fits are predicted to have dispersions close to ideal. For the correlated fit, the dispersion in the ring-down rate increases rather slowly with  $k\Delta t$ , whereas for the uncorrelated fit, the “cost” of under sampling the decay is much higher. This is in large part because of the need, for both fits, to decrease the bandwidth of the filtering to keep from under sampling the noise.

Figure 3 shows the reduced dispersion as a function of  $k_f\Delta t$  for both correlated and uncorrelated fits. Values of  $k\Delta t = 0.01$  and  $Nk\Delta t \gg 1$  are used in this calculation. This plot, which can be compared with Figure 1, demonstrates that for white noise, there is no significant improvement in the predicted sensitivity of CRDS if  $k_f\Delta t < 0.5$ . For the correlated fit, the dispersion continues to decrease, though very slowly, as  $k_f\Delta t$  decreases below 0.5. For the uncorrelated fit, however, the dispersion slowly rises. This is because of the increasing importance of data correlations for smaller  $k_f\Delta t$ , which is of course neglected in such fits. It is interesting to note that if  $F(t)$  instead of  $\bar{F}(t)$  is used in the least squares equations, then predicted dispersion continues to get smaller with decreasing  $k_f\Delta t$  and dips below the “ideal” values for fixed  $A$ . This is an artifact of the fact that to get the single exponential

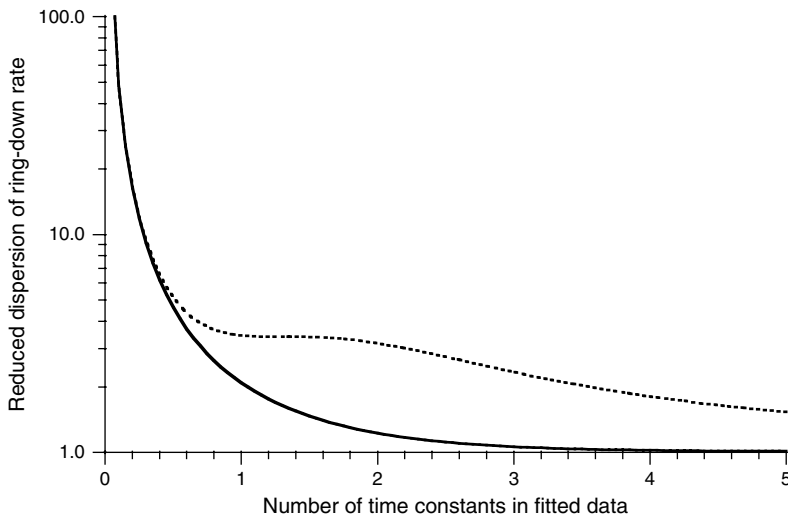


**Figure 3** Plot of the reduced dispersion of the ring-down decay rate as a function of  $k_f \Delta t$ , which equals the number of low-pass filter time constants between each data point. A value of  $k \Delta t = 0.01$  was used in calculating these curves. Again, the upper curve is for fits that ignore data correlation, whereas the lower is for fits that properly account for it. The ring down is assumed to be detector noise limited.

decay, one must delay the fit until  $k_f t \gg 1$ , which means that the appropriate amplitude of the ring down should be decreased.

We will now examine the convergence of the reduced dispersion as a function of the fraction of the ring-down transient that is sampled and fit. For a finite interval, we will get different results depending upon whether we include a variable or fixed baseline in the fit. Figure 4 shows the reduced dispersion as a function of  $Nk \Delta t$  calculated by the correlated fit with or without a variable baseline. Values of  $k \Delta t = 0.01$  and  $k_f \Delta t = 0.5$  were used in this calculation. Fits that assume uncorrelated data give almost indistinguishable results for this plot with these parameters. It can be seen that fits with fixed baseline converge quickly, and essentially no improvement in the dispersion of the ring-down rate is expected once three time constants of data are included. In contrast, fits including an adjustable baseline parameter converge rather slowly with increasing size of the data set. When fitting up to three time constants, the dispersion in  $k$  for variable baseline fits is more than twice as large as for fits that fix the baseline. Basically, this reflects the effect of correlation between the baseline and the decay rate parameters. This correlation goes to zero if an infinite number of data points are sampled, but slowly.

Because the least squares fit model is nonlinear in  $k$ , we will in general expect some bias, that is, the ensemble average of the cavity decay rate,  $\langle k' \rangle$  will differ from  $k$  by an amount that is proportional to  $\sigma_k^2$  for an “unbiased” method. Starting with the least squares fit solution equations ( $\beta = 0$ ), we have derived



**Figure 4** Plot of the reduced dispersion of the ring-down decay rate as a function of  $Nk\Delta t$ , which equals the number of cavity decay times sampled by the data. The lower curve is for fits that treat the baseline as a fixed number, whereas the upper curve is for fits that include a baseline offset as a fit parameter. Values of  $k\Delta t = 0.01$  and  $k_f\Delta t = 0.5$  were used. These curves were calculated including data correlation, but corresponding curves for fits ignoring data correlation are almost identical for these parameters.

$$\langle k' \rangle - k = \frac{1}{2} \sum_i \left( \frac{\partial^2 k'}{\partial y_i^2} \right) \sigma^2(y_i) \xrightarrow[Nk\Delta t \gg 1]{k\Delta t \ll 1} 4k^2 \left( \frac{P_N}{A} \right)^2. \quad (27)$$

Note that the bias in  $k'$  scales as the inverse square of the signal-to-noise ratio unlike the standard error of  $k'$  which scales as the inverse of the signal to noise. The bias scales as  $k^2$  compared to  $k^{3/2}$  for the standard error. For typical CRDS experiments, the bias in  $k'$  is so much smaller than its fluctuations that it can be neglected.

Since the least squares solution is in general iterative, it is useful to characterize the rate of convergence. We ran numerical calculations for an ensemble of 1000 decays with  $k\Delta t = 0.01$ ,  $A = 100\sigma$ , and  $N = 2000$  and iterated to convergence with different initial values for the decay rate. It was found numerically that the mean error in each fit (compared to the converged value for a given simulated decay) is approximately quadratic in the initial error in  $k'$ , with a mean and standard deviation of the error of  $-0.91(6)\%$  for a 10% initial error. This can be compared to a 0.30% fluctuation of  $k'$  for the ensemble after full convergence of the fits. A second iteration reduced the convergence error by more than two orders of magnitude. Thus, even with a crude,  $\sim 10\%$ , initial estimate of the decay rate, two cycles of the least squares fit will converge to well below the noise. With a moderately accurate initial guess,  $\sim 1\%$ , a single cycle of the fit should produce convergence well below the noise.

### 5.1. To average and then fit, or fit each decay and average the fit results?

If one truly has Gaussian distributed noise, there should be no difference in the standard error of the determined cavity decay rate whether one sums the transients and fits the resulting curve or if one fits each individual curve and calculates the weighted mean (based on predicted fit standard error) of the individual decay rates. Given the computational cost of fitting up to several thousand data points, one may expect that it is better to average and then fit.

It is our experience that even in a well running CRDS instrument, there is a small fraction of “bad decays” that give cavity decay rates that are outliers from the Gaussian distribution of  $k'$  values that is expected if the detector has Gaussian noise characteristics. Recently, we published a paper that presents an analysis of one source of such outliers – resonances with high order transverse modes that are coupled to the TEM<sub>00</sub> mode by mirror scattering [17]. But even with such resonances killed by control of the cavity Fresnel number [17], occasional outliers continue to exist. One likely source of such outliers is the drifting of a small particle through the mode of the cavity. Given a typical TEM<sub>00</sub> mode diameter of 1 mm, a particle with scattering cross section in the order of  $1\ \mu\text{m}^2$  will introduce  $\sim 1$  ppm per pass loss, which is orders of magnitude above the noise level we achieve. Luckily, usually the  $\chi^2$  of the fit to the cavity decay is typically significantly higher than expected for our observed bad shots. If one examines the residuals of the fit, a pronounced “ringing” is often evident. The modulation in the residuals suggests that part of the loss is modulated during the ring-down transient. Motion of a particle along the optic axis of the cavity could produce such modulation since one expects the scattering loss to be larger for the center of the particle at the antinodes of the standing wave of the excited cavity mode than at the nodes. Miller and Orr-Ewing [18] recently demonstrated that significant spatial loss modulation is expected even for particles whose diameters are many times the wavelength of the scattered light. We note that for  $\lambda = 1\ \mu\text{m}$ , a particle velocity of 1 cm/s will produce a loss modulation of 20 kHz, which will strongly distort a ring-down transient with  $k \sim 10^4\ \text{s}^{-1}$  as is typical in our experiments. When processing real CRDS data, we have the software reject data points that have excessive  $\chi^2$  values. Such filtering to reject outliers is not possible if a large number of transients are averaged and then fit. In that case, a single bad decay can cause significant deviation of the cavity decay rate. It is our suggestion, whenever possible without significantly reducing the rate of cavity decays observed, that each transient be fit and a  $\chi^2$  test be performed on the fit. In cases of very high data rates that often occurs when using mirrors of only modest reflectivity, fitting each decay will no longer be possible, but we recommend that the smallest practical sets of averaged data be fit.



## 6. LINEARIZATION OF THE FIT IN CAVITY RING-DOWN SPECTROSCOPY

In many applications of CRDS, requiring rapid fitting of a large number of observed ring-down transients, the ring-down signal is converted  $y_i \rightarrow \ln(y_i)$ , which

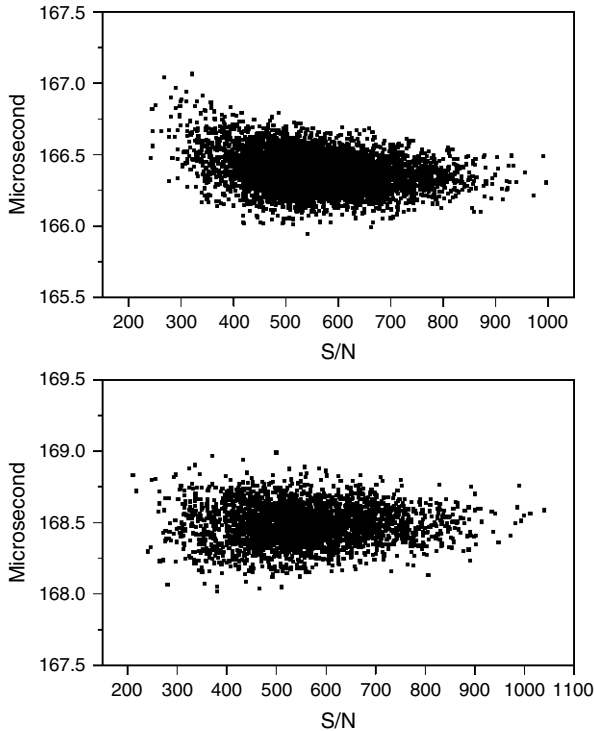
converts  $F(t) \rightarrow \ln(A) - kt$ , that is, we now have a linear least squares fit problem. The advantage of a linear least squares fit is that it can be solved exactly in a single step, instead of requiring an iterative solution, as most nonlinear least squares fits require. In the limit that the fractional error in the data is small over the entire region fit, this “linearization” will not change the least square solution or variances, provided that the uncertainties of each data point are appropriately transformed as well,  $\sigma_i \rightarrow \sigma_i/F(t_i)$ . A problem with this approach is that errors in the wing of the decay are not properly estimated. In fact, one will attempt to calculate the log of a negative number should the noise ever dip below the zero value, which it clearly must at long time. As a result, if this procedure is used, one must be careful to restrict the fit to only the high signal-to-noise portion of the decay, say requiring that  $y_i \geq m\sigma_i$ , where  $m$  is some predetermined multiplier. If this point occurs sufficiently far out in the decay, then this truncation results in a negligible decrease in precision. Another issue that should be kept in mind is that the nonlinear transformation of the experimental data will generate a bias in the distribution of errors

$$\ln(\gamma(t)) = \ln(F(t)) + \frac{\epsilon(t)}{F(t)} - \frac{1}{2} \frac{\epsilon(t)^2}{F(t)^2} + \dots \quad (28)$$

with the ensemble average of  $\epsilon(t)^2 = \sigma_i^2$ . For the case of detector noise limited CRDS (i.e. constant  $\sigma^2$ ),  $k\Delta t \ll 1$  but  $\exp(-Nk\Delta t) \ll 1$ , it can be shown by propagation of error that the second-order term in the above Taylor expansion for the transformation introduces a bias in the fitted decay rates of

$$\langle k' \rangle - k = 2k \left( \frac{\sigma}{A} \right)^2 (Nk\Delta t)^2 (1 - (2k\Delta tN)^{-1}). \quad (29)$$

Another source of bias occurs if one uses  $\sigma_i \rightarrow \sigma_i/\gamma_i$  for the uncertainty and thus weight the data proportional to  $\gamma_i^2$  instead of  $F_i^2$ , which are not known. This has the effect of giving more weight to points with positive noise and less to those with negative noise. Numerical simulations indicate that this bias is of the opposite sign as the bias due to the nonlinear transformation and about twice as large (at least for the parameter range that was explored). Figure 5 displays two scattering plots of fitted  $\tau$  (equals  $1/k'$ ) versus signal-to-noise ratio of decay transients. For the upper part, this linearization process is used in the fit, without corrections to log transformation and with the weight proportional to  $\gamma_i^2$ . One can see clearly from it that the fit is biased and the bias increases with the decrease of the signal-to-noise ratio. For an initial signal-to-noise ratio of 100:1, the noise in  $k$  was close to that predicted by Equation (26), but the mean value of  $k$  was shifted by  $-0.5\%$  from the value used to generate the data ( $k\Delta t = 0.01$ ,  $m = 1.5$ ), which was close to twice the ensemble standard deviation of  $k$ . With an initial signal-to-noise ratio of 10:1, the standard deviation of  $k$  was nearly twice the prediction of Equation (26), and the bias was  $-8.5\%$ . Even with a signal-to-noise ratio of 1000:1, the bias ( $-0.015\%$ ) is still 58% of the ensemble fluctuations in  $k$  and thus will be significant with only modest signal averaging to the decay rates. The bias in the log-transformed fit at low signal-to-noise ratios was previously discussed by von Lerber and Sigris [9].



**Figure 5** Scattering plots of the fitted  $\tau$  versus the signal-to-noise ratio of experimental decay transients. For the upper part, linearized least squares fit is used, without corrections to the log transformation and with the weight proportional to  $y_i^2$ . The fit is significantly biased for transients with signal-to-noise ratio less than 500. For the lower part, a single pass of the nonlinear least squares fit process is added to the fitting program and it is unbiased.

Numerical simulations have demonstrated that the effects of the bias can be greatly reduced by two simple changes in the routine. To correct for the leading order bias term,  $\ln y_i + (1/2)\sigma_i^2/y_i^2$  is used for the “data” in the least squares expressions. Also, the weight is calculated from an estimate of the value of  $k$  (the assumed values of  $A$  does not effect the results of the least fit results). Using an incorrect value (even with 20% deviation from  $k$ ) for the decay rate to calculate the weights,  $k_w$ , will not introduce any significant bias, but the lowest fluctuations in the fitted decay rate,  $k'$ , will be obtained if  $k_w = k$ . However, the dependence is relatively weak. The variance of  $k'$  increases from the optimal value by a factor of  $\kappa^4(2\kappa^2 - 2\kappa + 1)/(2\kappa - 1)^3$ , where  $\kappa = k_w/k$ . The variance increases by only 10% for  $\kappa = 0.87$  or 1.18. Thus, one could use a rough estimate for  $k$  for calculation of the weight with little penalty in precision. With this change in the fitting procedure ( $k\Delta t = 0.01$ ,  $m = 1.5$ ), the bias in  $k'$  with signal to noise of 100:1 is reduced to  $-0.027\%$ . With a signal to noise of 10:1, the bias is reduced to  $-2.1\%$ , but the fluctuations in  $k'$  were still about 66% higher than predicted. For such low signal to noise, the nonlinear least squares fit to the

data is needed to reduce the standard error in the fitted decay rate close to the theoretical limit.

Our suggestion for an optimal strategy is a two-step fit. First, linearization is used to provide a good estimate of the parameters  $A'$  and  $k'$ . The detector value corresponding to zero light ( $B'$ ) is either assumed (best if it is stable with time) or else estimated from the tail of the decay. Speed can be improved by storing a table of possible  $\ln(y_i)$  values rather than recomputing these for each data point. Then a single step of refinement to the solution of the original least squares fit problem be carried out. As discussed earlier, for an initial estimate of  $k$  that is in error by 1%, this single step will reduce the convergence error to the order of 0.01%. Only two or three sums over the data points (depending upon whether the baseline is constrained for fit) need to be calculated, so the single round of the nonlinear fit in fact requires less floating point operations than the linearized log fit. The lower part of the Figure 5 shows that the fitted  $\tau$  is unbiased after the nonlinear fit process is added to the fitting program. An alternative strategy is to use  $k'$  from the previous decay as the initial value for the least squares fit and dispense with the log transformation entirely. This should be an ideal approach when decays are being observed very rapidly, because in that case the cavity loss (and thus  $k$ ) cannot change much decay to decay.

## 7. DETERMINATION OF RING-DOWN RATE BY FOURIER TRANSFORM METHOD

One method widely used to extract one or more decay rates from a transient is based on taking a Fourier transform of the decay [19], which for points evenly spaced in time can be done with the fast Fourier transform method [16]. This has been proposed for the rapid analysis of CRDS data by Mazurenka et al. [10].

Let  $f_j$  be the  $\omega_j = 2\pi j/N\Delta t$  frequency component of the discrete Fourier transform of the time series  $F(t_i) = A \exp(-k\Delta t i) + B$  ( $i = 0 \dots N - 1$ ). This is easily calculated to be

$$f_j = \left( A \frac{1 - \exp(-k\Delta t N)}{1 - \exp(-(k + i\omega_j)\Delta t)} + BN\delta_{j0} \right) \Delta t. \quad (30)$$

$\delta_{j0}$  is 1 if  $j = 0$  and 0 if  $j \neq 0$ . For  $\omega_j \neq 0$ , we can take the ratio

$$\omega_j \frac{\Re(f_j)}{\Im(f_j)} = \omega_j \frac{1 - e^{-k\Delta t} \cos(\omega_j \Delta t)}{e^{-k\Delta t} \sin(\omega_j \Delta t)} \rightarrow k, \quad (31)$$

where the right arrow applies if both  $k, \omega_j \ll (\Delta t)^{-1}$ . This is the relationship we get from a continuous Fourier transform. This suggests that we can evaluate  $k$

from any frequency, except the zero frequency for which the baseline makes a contribution. It has been traditional to use the first nonzero frequency point in the computation of  $k'$  from the FFT. For finite values of  $k$  and  $\omega_j$ , bias is introduced by the above procedure. However, the first half of Equation (31) can be solved for  $\exp(-k\Delta t)$ , and this was used to evaluate  $k'$  from the ratio without bias (without noise)

$$k' = \frac{1}{\Delta t} \ln \left[ \sin(\omega_j \Delta t) \frac{\Re e(f_j)}{\Im m(f_j)} + \cos(\omega_j \Delta t) \right]. \quad (32)$$

Error propagation for  $k, \omega_j \ll (\Delta t)^{-1}$  gives the following expression for the variance in  $k'$  extracted from the  $\omega_j$  frequency component of the FFT

$$\sigma^2(k') = \frac{\Delta t^2 (k^2 + \omega_j^2)^3 N}{2\omega_j^2} \left( \frac{\sigma}{A} \right)^2. \quad (33)$$

Note that the variance grows with  $N$ , so one does not want to take too much baseline. This is particularly true if one takes the first nonzero frequency, as  $\omega_1 = 2\pi/(\Delta t N)$  and thus the variance grows as  $N^3$  because typically  $\omega_1 \ll k$ . Treating  $\omega_j$  as a continuous variable, it is easily shown that the noise will be minimized by taking  $\omega_j = k/\sqrt{2}$ , for which  $\sigma^2(k') \rightarrow (3/2)^3 k^4 \Delta t^2 N (\frac{\sigma}{A})^2$ . Because of the nonlinear relationship between  $k'$  and  $f_j$  (which is linearly related to the data), noise will create a bias because of the second derivatives of the  $k'$  with respect to the data points used to calculate it, as we had for the linearized least squares fit model above. For  $k, \omega_j \ll (\Delta t)^{-1}$  and  $Nk\Delta t \gg 1$ , this bias can be calculated analytically, giving

$$\langle k' \rangle - k = -\frac{kN}{2} \left( \frac{\Delta t (k^2 + \omega_j^2)}{\omega_j} \frac{\sigma}{A} \right)^2. \quad (34)$$

The optimal estimate of  $k$ , at least in cases of small noise where the first order error propagation is valid, would be to average the value of  $k'$  extracted from each frequency component, but with a weight inversely proportional to the variance, as given by Equation (33). Because the number of values of  $\omega_j$  that will contribute significantly to the sum will be proportional to  $N$ , this extra step should produce a value for  $k'$  with a variance that converges for large  $N$ , like for the least squares solution. However, if  $N$  is too large, the noise in the values of  $f_j$  becomes comparable to their magnitude, and the first-order treatment of the noise will no longer be accurate. Using only the first point, we have found by numerical simulation with parameters  $k\Delta t = 0.01$  and  $N = 512$  that the extracted values of  $k'$  showed fluctuations that are about 84% larger than for the ideal least squares limit. Because the number of numerical operations to evaluate an FFT scales as  $N \ln N$ , it appears that this method should be more computationally expensive than a direct least squares fit to the data.

## 8. THE SUCCESSIVE INTEGRATION METHOD FOR EXPONENTIAL FITTING

The SI method for exponential fitting was introduced by Matheson [20]. Using the fact that the integration of an exponential function is still an exponential function, the exponential fitting problem is transformed into a linear regression problem, which is a noniterative method. Halmer et al. [11] found the method can be improved by introducing a correction factor to account for errors introduced by the use of Simpson's rule to evaluate the integration from the discrete experimental points and called the improved estimate the corrected successive integration (CSI) method. Both the SI and the CSI methods give nearly the same results as the least squares algorithm. However, the dispersion estimation for fitted parameters given in both papers is incorrect. The authors utilized the curvature matrix as in the normal least squares fit. However, in the SI and CSI methods, the form of  $\chi^2$  has changed. The new curvature matrix is a function of the dependent data,  $y_i$ , not just the independent data,  $t_i$ , and parameters. As a consequence, the curvature matrix in the SI and CSI methods contains fluctuations caused by the noise in the data. These fluctuations contribute to the variance of fitted parameters. Below, we present an evaluation of standard error of the fit parameters that properly accounts for this effect.

The fitting model can be changed into a different form after one integrates the model equation [11].

$$y(t) = Ae^{-kt} + B \quad (35)$$

$$\int_0^{i\Delta t} y(t') dt' = \frac{A+B}{k} - \frac{y(i)}{k} + Bi\Delta t \quad (36)$$

$$y(i) = A + B - k \int_0^{i\Delta t} y(t') dt' + kB i \Delta t. \quad (37)$$

The correction factor  $CT(k)$  was introduced by the Halmer et al. [11].

$$\int_{n\Delta t}^{(n+1)\Delta t} y(t') dt' = B\Delta t + CT(k) \left( \frac{y(n) + y(n+1)}{2} - B \right) \Delta t \quad (38)$$

$$CT(k) = \frac{2}{k\Delta t} \cdot \frac{1 - e^{-k\Delta t}}{1 + e^{-k\Delta t}}. \quad (39)$$

With this correction, one can calculate the integral in Equation (37).

$$\int_0^{i\Delta t} y(t') dt' = CT(k) X_i \Delta t + B[1 - CT(k)] i \Delta t, \quad (40)$$

with  $X_i$  as defined by Halmer et al. [11].  $X_0 = 0$  and for  $i > 0$ ,

$$X_i = \sum_{j=0}^{i-1} \frac{y_j + y_{j+1}}{2}. \quad (41)$$

The last expression on the right results from using Simpson’s rule to evaluate the integral in terms of the discrete data points,  $y_i$ . Substituting Equation (40) into Equation (37), we have a new equation:

$$y_i = A + B - k\Delta tCT(k)X_i + k\Delta tBCT(k)i. \tag{42}$$

Following normal least squares fit procedures [11], one can have the matrix equation for the best fit parameters  $A$ ,  $B$ , and  $k$ .

$$M \cdot \beta = v$$

$$M = \begin{pmatrix} N & SY & t \\ SY & SY.SY & t.SY \\ t & t.SY & t.t \end{pmatrix}, \beta = \begin{pmatrix} A + B \\ -k\Delta tCT(k) \\ k\Delta tCT(k)B \end{pmatrix}, v = \begin{pmatrix} Y \\ Y.SY \\ Y.t \end{pmatrix} \tag{43}$$

with definitions

$$Y = \sum_{i=0}^{N-1} y_i \qquad SY = \sum_{i=0}^{N-1} X_i \qquad Y.SY = \sum_{i=0}^{N-1} y_i X_i$$

$$SY.SY = \sum_{i=0}^{N-1} X_i^2 \qquad Y.t = \sum_{i=0}^{N-1} iy_i \qquad t.SY = \sum_{i=0}^{N-1} iX_i$$

$$t = N(N - 1)/2 \quad t.t = N(N - 1)(2N - 1)/6$$

The definitions of the elements in the matrix and  $v$  vector are the same as those in the Ref. [11] though we have used sums  $i = [0 \dots N - 1]$  to be consistent with the notation of this chapter. Using Equation (39),  $k'$  is easily calculated in terms of  $\beta_1$ . These elements are functions of data points  $y_i$  and contain the noise fluctuations. This makes it much more complex to estimate the variance of parameters in the fitting than the normal curvature situation. Derivation of a fully analytic expression for the standard errors of the parameters is quite involved, and so we used a numerical method. We start with the standard expression for propagation of errors, neglecting correlation of the data points:

$$\delta\beta_i^2 = \sigma^2 \sum_{j=0}^{N-1} \left( \frac{\partial\beta_i}{\partial y_j} \right)^2 \quad i \in 0, 1, 2, \tag{44}$$

where we have assumed that detector noise dominates with a constant noise per data point. As above, data correlation can be included by taking for  $\sigma^2 = \xi^2 P_N^2 / \Delta t$ , which is a factor of  $G$  larger than the variance of the data points without correlation. Thus, we need to evaluate the partial derivatives of the fit parameters in terms of the input data. Differentiation of Equation (43) gives

$$\left( \frac{\partial\beta}{\partial y_k} \right) = -M^{-1} \cdot \left( \frac{\partial M}{\partial y_k} \right) \cdot \beta + M^{-1} \left( \frac{\partial v}{\partial y_k} \right). \tag{45}$$

By defining the matrix  $J$ ,

$$J = \begin{pmatrix} 1 & 0 & 0 \\ 1 & X_1 & 1 \\ \vdots & \vdots & \vdots \\ 1 & X_{N-1} & N-1 \end{pmatrix}, \quad (46)$$

we can express matrix  $M = J^\dagger J$  and vector  $v = J^\dagger \vec{y}$  and thus express their derivatives in terms of the derivatives of  $J$ .

$$\left( \frac{\partial M}{\partial y_k} \right) = \left( \frac{\partial J}{\partial y_k} \right)^\dagger \cdot J + J^\dagger \cdot \left( \frac{\partial J}{\partial y_k} \right) \quad (47)$$

$$\left( \frac{\partial v}{\partial y_k} \right) = \left( \frac{\partial J}{\partial y_k} \right)^\dagger \cdot \vec{y} + J^\dagger \left( \frac{\partial \vec{y}}{\partial y_k} \right) \quad (48)$$

$\left( \frac{\partial \vec{y}}{\partial y_k} \right) = \delta_{ki}$ . Using the definition of  $X_i$  (Equation (41)), we can write the nonzero elements of  $\left( \frac{\partial J}{\partial y_k} \right)$  as follows (the column and row numbering start at zero):  $\left( \frac{\partial J_{i1}}{\partial y_0} \right) = 1/2$  ( $i > 0$ );  $\left( \frac{\partial J_{i1}}{\partial y_k} \right) = 1$  ( $0 < k < i$ );  $\left( \frac{\partial J_{k1}}{\partial y_k} \right) = 1/2$  ( $k > 0$ ).

With the same values of parameters as in Ref. [11], that is,  $A = 1200$ ,  $k = 0.02$ ,  $\Delta t = 1$ ,  $B = -13$ , and  $\sigma = 20$ , we generated 5000 decay transients and fit them with different methods. We found with the usual nonlinear fit method, the averages of fitted parameters and their standard deviations,  $A' = 1200.08 \pm 5.54$ ,  $k' = 0.02000 \pm 0.00014$ , and  $B' = -13.00 \pm 0.55$ . The predicted standard deviations are 5.55, 0.00014, and 0.55, respectively, by using the curvature matrix. With the CSI method fitting the same data, the averages of fitted parameters and their standard deviations are  $A' = 1199.36 \pm 7.21$ ,  $k' = 0.01999 \pm 0.00021$ , and  $B' = -13.01 \pm 0.59$ . The predicted standard deviations by Equation (44) are 7.22, 0.00021, and 0.54, respectively, with Equation (44), that is, the standard deviation of  $k'$  in the CSI fit is about 50% higher than for a direct least squares fit to the same data. We note that Table II of the Ref. [11] gives parameter standard deviations of 2.3, 0.00048, and 0.98 matching neither of our results. Although the SI and CSI methods are noniterative, one disadvantage is that the  $M$  matrix is badly conditioned, increasingly so with increasing  $N$ , which can generate numerical instability in the calculation of its inverse.

The principal advantage that Halmer et al. [11] ascribe to the CSI method is its computational speed; their Table II shows fitting times two orders of magnitude faster than for the least squares fit using the Levenberg – Marquardt algorithm. This appears, however, to be an artifact of their using a general purpose fitting package for the least squares fit. The SI and CSI fits each requires the calculation of seven sums over data points, requiring eight floating point additions and five floating point multiplications. This can be compared to 3, 3, and 3 for the same quantities in the least squares fit using the expressions given above. Thus, one cycle of the least

squares fit should take about half the number of floating point operations if an efficient code is used that exploits the closed form expressions for the curvature matrix. Because the least squares fit converges very rapidly and does not require more than two cycles, we see no computational speed advantage to the use of the CSI method. Certainly, the substantial advantage reported in Ref. [11] is not correct.



## 9. ANALOG-DETECTED CAVITY RING-DOWN

### 9.1. Phase shift method

There are several analog electronic methods to determine the ring-down decay rate of an optical cavity. The oldest is the phase shift CRDS method (ps-CRDS) [12,13]. In this method, the light injected into an optical cavity is modulated at an angular frequency  $\Omega$ , and the light transmitted by the cavity is demodulated by a vector lock-in amplifier that determines both the in-phase ( $S_c$ ) and out-of-phase ( $S_s$ ) components of the signal and the phase shift of the signal determined by the equation  $\tan(\theta) = S_s/S_c$ . The cavity decay time can be determined by  $\tan(\theta) = \Omega\tau$ . We will now examine the expected noise in this signal extraction method.

Let us begin by assuming that the cavity is excited by a pulsed laser with pulse length much shorter than the cavity decay time, such that one can treat the excitation as impulses at times  $t_n = n2\pi/\Omega$ . In each time interval  $(t_n, t_{n+1})$ , the detector signal will be  $\bar{F}(t) = A_n \exp(-kt)$ . Demodulation gives the average of the following two signals:

$$F_c = \frac{\Omega}{2\pi} \int_{t_n}^{t_{n+1}} \bar{F}(t') \cos(t') dt' = A_n \frac{\Omega k [1 - e^{-k2\pi/\Omega}]}{2\pi(k^2 + \Omega^2)} \quad (49)$$

$$F_s = \frac{\Omega}{2\pi} \int_{t_n}^{t_{n+1}} \bar{F}(t') \sin(t') dt' = A_n \frac{\Omega^2 [1 - e^{-k2\pi/\Omega}]}{2\pi(k^2 + \Omega^2)} \quad (50)$$

$$k' = \frac{\Omega}{\tan(\theta)} = \Omega \frac{F_c}{F_s}. \quad (51)$$

If we have detector-limited CRDS, then the time averaged of both  $F_c$  and  $F_s$  will have noise with variance given by  $\sigma^2 = P_N^2 \text{BW}/2$ , where BW is the detection bandwidth on the lock-in amplifier and (as before)  $P_N$  is the noise-equivalent power of the detector. Adding these noise terms to the above ratio and doing standard error propagation, we find that the variance in the calculated cavity decay rate is given by

$$\sigma^2(k') = \frac{4\pi^2(k^2 + \Omega^2)^3}{\Omega^4(1 - e^{-k2\pi/\Omega})^2} \left( \frac{\sigma}{\langle A_n \rangle} \right)^2. \quad (52)$$

If we assume that the light field of different laser pulses adds incoherently, then it can be shown that  $\langle A_n \rangle = J(\Omega) T^2 t_r^{-1} (1 - e^{-k2\pi/\Omega})^{-1}$ , where  $J(\Omega)$  is the energy

per pulse (which is a function of the laser repetition rate  $= \Omega/2\pi$ ),  $T$  is the power transmission of the mirrors (assumed to be the same), and  $t_r$  is the cavity round trip time. The last term corrects for the energy left in the cavity from previous pulses. This gives for the noise in the estimated decay rate

$$\sigma^2(k') = \frac{4\pi^2(k^2 + \Omega^2)^3}{\Omega^4} \left( \frac{\sigma t_r}{J(\Omega)T^2} \right)^2. \quad (53)$$

We consider two limiting cases for the laser pulse energy  $J(\Omega)$ . One is that it is a constant,  $J$ , independent of repetition rate, such as when a high repetition rate laser is pulse picked or a continuous wave optical source is chopped. In that case, the lowest variance in  $k'$  occurs for  $\Omega = \sqrt{2}k$ , and for this modulation frequency,  $\sigma^2(k') = 125\pi^2 k^2 \left( \frac{\sigma t_r}{JT^2} \right)^2$ . The other limit is where the average power of the pulsed laser is fixed (such as for many Q-switched lasers at high repetition rate), in which case  $J(\Omega) = 2\pi I_{\text{avg}}/\Omega$ . In this case, the optimal modulation angular frequency is  $\Omega = k/\sqrt{2}$ , and at this modulation frequency,  $\sigma^2(k') = (27/4)k^4 \left( \frac{\sigma t_r}{I_{\text{avg}}T^2} \right)^2$ . It is worth noting that shot-to-shot fluctuations in the amplitude of the light will cause correlated noise in  $F_s$  and  $F_c$  and will not degrade the ability to extract the cavity decay rate using the phase shift method.

Because one must take the ratio of the in-phase to out-of-phase lock-in outputs to compute the decay rate, noise will generate a bias in  $k'$  equal to  $(1/2)(d^2 k'/dF_s^2)\sigma^2(F_s)$ . Evaluation using the above expressions gives the following prediction for the bias

$$\langle k' \rangle - k = \frac{4\pi^2 k(k^2 + \Omega^2)^2}{\Omega^4} \left( \frac{\sigma t_r}{J(\Omega)T^2} \right)^2 \quad (54)$$

For the two limiting cases of  $J(\Omega)$  described above, we find for the optimal modulation frequency  $\langle k' \rangle - k = 9\pi^2 k \left( \frac{\sigma t_r}{JT^2} \right)^2$  for the constant pulse energy case and  $\langle k' \rangle - k = (9/2)k \left( \frac{\sigma t_r}{I_{\text{avg}}T^2} \right)^2$  for the constant average power case.

If one is exciting the cavity with a continuous wave source that is chopped, one will want to leave the excitation source for a significant fraction of the modulation cycle. We will now consider the case of 50% duty cycle excitation of the cavity. We will also assume that the coherence time of the excitation source is much less than the cavity decay time, so that we will have a rate equation for the cavity intensity (instead of the cavity field). We will define the modulation cycle such that the input light intensity drops to zero at the beginning of the cycle and turns on to a constant value for the second half of the cycle. In that case, the mean output intensity of the cavity will be  $F(t) = A \exp(-kt)$  for  $0 < t < \pi/\Omega$  and  $F(t) = A[1 + \exp(-k\pi/\Omega) - \exp(t - \pi/\Omega)]$  for  $\pi/\Omega < t < 2\pi/\Omega$ . If we did not modulate the input light, we would have a mean cavity transmission of  $I_{\text{avg}} = A[1 + \exp(-k\pi/\Omega)]$ . If the bandwidth of the incident radiation is greater than the free spectral range of the cavity [21] or one is exciting many transverse modes to “fill in” the spectrum of

the cavity [22], then  $I_{\text{avg}} = kt_r T^2 I_{\text{inc}}$ , where  $I_{\text{inc}}$  is the intensity incident on the cavity. Using the above defined definitions of  $F_c$  and  $F_s$ , we find for this  $F(t)$  that

$$F_c = \frac{\Omega k}{\pi(k^2 + \Omega^2)} I_{\text{avg}} \quad (55)$$

$$F_s = \frac{k^2}{\pi(k^2 + \Omega^2)} I_{\text{avg}} \quad (56)$$

$$k' = -\Omega \tan(\theta) = \Omega \frac{F_s}{F_c}. \quad (57)$$

As above, if we have detector-limited CRDS, then the time averaged of both  $F_c$  and  $F_s$  will have noise with variance given by  $\sigma^2 = P_N^2 \text{BW}/2$ . Error propagation gives for the variance of the fitted cavity decay rate

$$\sigma^2(k') = \frac{\pi^2(k^2 + \Omega^2)^3}{k^2 \Omega^2} \left( \frac{\sigma}{I_{\text{avg}}} \right)^2. \quad (58)$$

The optimal modulation angular frequency is given by  $\Omega = k/\sqrt{2}$ , which gives  $\sigma^2(k') = (27/4)\pi^2 k^2 (\sigma/I_{\text{avg}})^2$ . As above, we can use the second derivative of  $k'$  with respect to  $F_c$  to predict the bias

$$\langle k' \rangle - k = -\frac{\pi^2(k^2 + \Omega^2)^2}{\Omega^2 k} \left( \frac{\sigma}{I_{\text{avg}}} \right)^2 = -\frac{9}{2} \pi^2 k \left( \frac{\sigma}{I_{\text{avg}}} \right)^2 \left( \text{For } \Omega = k/\sqrt{2} \right), \quad (59)$$

where the second equality holds for optimal modulation.

Unlike the case of pulsed excitation, fluctuations in the cavity excitation intensity will lead to noise in the values of  $F_c$  and  $F_s$  that are not perfectly correlated, that is, this form of CRDS is not immune to source noise. Most important, the field build up of each mode will, when the source coherence time is short compared to the build up, suffer “temporal speckle” with the light amplitude adding incoherently at different times [23,24]. As previously discussed, this leads to the intracavity intensity for each mode to fluctuate with a  $\chi^2$  in two degree of freedom,  $P(I) = \exp(-I/\langle I \rangle)/\langle I \rangle$ , where  $\langle I \rangle$  is the mean intensity predicted from the intensity rate equation and  $P(I)$  is the probability that the intensity will be  $I$ . In the case of one or few mode excitation of the cavity, this noise will likely be dominant over the detector noise, the effects of which we have calculated above. With many mode excitation, either because the input light is broad band and excites many longitudinal modes of the cavity or because light is injected off-axis to excite many transverse modes, the fluctuations will be greatly decreased. This type of excitation is widely used in cavity-enhanced spectroscopy [25] (also called cw integrated cavity output spectroscopy [26]) where one uses changes in the time averaged transmission of the cavity to determine absorption. In this type of spectroscopy, one must determine the empty cavity loss to convert the observed change in transmission into absolute loss [25]. In this case, the use of phase shift detected CRDS

offers an attractive approach, particularly if one's excitation source is a broad bandwidth source, such as an incoherent light source like a lamp or light-emitting diode.

## 9.2. Gated integrator method

An alternative analog detection method was introduced by Romanini and Lehmann [7] and used a pair of integrator gates to extract the cavity decay rate. After the excitation pulse of the cavity has ended, the output light intensity will decay as  $A \exp(-kt)$ . We will assume that any offset in the detector has been subtracted off. Let gated integrator 1 integrate the signal from time  $t = [0, \Delta t]$  and integrator 2 from  $t = [\delta t, \delta t + \Delta t]$ . (Note: the meaning of  $\Delta t$  has changed from previous sections.) As shown previously, the cavity decay rate can be evaluated as  $k' = -\ln(F_2/F_1)/\delta t$ , where  $F_{1,2}$  is the output of gated integrator 1 or 2. Romanini and Lehmann [7] presented an analysis of the expected shot noise limited sensitivity. We will now present the analysis of the predicted sensitivity and bias for the case where detector noise dominates.

Integration of the signals over the detection windows gives the following values for  $F_{1,2}$ :

$$F_1 = \frac{A}{k} (1 - e^{-k\Delta t}) \quad (60)$$

$$F_2 = \frac{A}{k} e^{-k\delta t} (1 - e^{-k\Delta t}). \quad (61)$$

Error propagation easily demonstrates that  $\sigma^2(F_{1,2}) = \Delta t P_N^2$ .

$$\begin{aligned} \sigma^2(k') &= \frac{1}{\delta t^2} \left( \frac{\sigma^2(F_2)}{F_2^2} + \frac{\sigma^2(F_1)}{F_1^2} \right) \\ \sigma^2(k') &= \frac{k^2}{\delta t^2} (1 + e^{2k\delta t}) (1 - e^{-k\Delta t})^2 \Delta t \left( \frac{P_N}{A} \right)^2 \end{aligned} \quad (62)$$

Minimization of  $\sigma^2(k')$  gives optimal values  $\delta t = 1.11k^{-1}$  and  $\Delta t = 1.255k^{-1}$  and leads to a predicted decay rate variance  $\sigma^2(k') = 20.3k^3(P_N/A)^2$ , which is only about 2.5 times larger than the ideal weighted least squares fit prediction given above. The above analysis ignores the correlation of the noise between  $F_1$  and  $F_2$ , but this will be small as the optimized sample windows hardly overlap. In making comparison with the phase shift method, note that the present variance is computed for a single transient, and the variance will be reduced by the  $\Omega/2\pi$ , which is the repetition rate that ring-down transients are detected. If there are no constraints on the repetition rate of the light source, then the optimal gates will be reduced somewhat to allow for a higher value of  $\Omega$ , which must be less than  $2\pi(\delta t + \Delta t)^{-1}$ .

We can also calculate the expected bias from the second derivatives of  $k'$  with respect to  $F_1$  and  $F_2$ . This leads to the result

$$\langle k' \rangle - k = -\frac{1}{2} k^2 \left( \Delta t k (2 - e^{-k\Delta t})^{-2} \right) \left( \frac{e^{2k\delta t} - 1}{k\delta t} \right) \left( \frac{P_N}{A} \right)^2. \quad (63)$$

Using the optimized values for  $\Delta t$  and  $\delta t$  given above, we find predicted bias  $\langle k' \rangle - k = -9.06k^2 \left(\frac{P_N}{A}\right)^2$ .

We have assumed that the detector baseline is known and subtracted from the signal before integration. Alternatively, one could use a third integration gate placed after the ring-down signal has decayed to negligible level and the average over this time interval could be used to subtract the baseline from signals  $F_{1,2}$ . Ideally, the time interval over which the baseline should be integrated should significantly exceed  $\Delta t$  so that the noise in this baseline estimate is small compared to the noise in the signals  $F_{1,2}$ , but that may limit the repetition rate of cavity decays that could be sampled.

### 9.3. Logarithm-differentiator method

Spence et al. [14] presented an analysis and experimental results on a third analog detection method. Here, one passes the detector output through a log amplifier and then a differentiator. During the ring-down, the voltage output of the differentiator will be proportional to the cavity decay rate. If we assume that this analog signal is averaged over the time interval  $[0, \Delta t]$ , then the extracted cavity decay rate,  $k'$ , is given by

$$k' = -\frac{1}{\Delta t} \int_0^{\Delta t} \frac{d \ln(y(t'))}{dt'} dt' = -\frac{1}{\Delta t} (\ln(y(\Delta t)) - \ln(y(0))). \quad (64)$$

If we assume that the sample interval is long enough that we can treat the noise as uncorrelated at  $t=0$  and  $t=\Delta t$ , and taking the first derivative terms, it is easy to show that the variance in the extracted decay rate is

$$\sigma^2(k') = \frac{1}{\Delta t^2} \left(\frac{\sigma}{A}\right)^2 [1 + e^{2k\Delta t}]. \quad (65)$$

This is optimized with  $\Delta t = 1.109k^{-1}$  for which, including that  $\sigma^2 = (k_f/2)P_N^2$ ,

$$\sigma^2(k')_{\text{optimized}} = 4.14k^2k_f \left(\frac{P_N}{A}\right)^2. \quad (66)$$

Comparing with the least squares fit results, we see that even with an optimal sample window, the use of a log amplifier and differentiator produces an estimate for the cavity decay rate that has a variance  $\sim k_f/2k$  times larger than for optimal fitting. Note that this result is derived under the assumption that  $k_f \gg k$  (since we started the fit at zero time instead of waiting for the initial decay transient due to detector time constant ( $1/k_f$ ) to die away). For typical CRDS conditions  $k_f \sim 10^2k$  or greater to avoid significant distortion of the cavity decay. This increased noise is a natural consequence of the differentiation of the signal, as this enhances the high frequency noise. Gaussian noise in the detector will produce a bias due to the log transformation, and this is easily seen to be

$$\langle k' \rangle - k = \frac{k_f}{4\Delta t} \left(\frac{P_N}{A}\right)^2 [1 + e^{2k\Delta t}] = 2.29kk_f \left(\frac{P_N}{A}\right)^2, \quad (67)$$

where the second expression is for the case that the lowest noise time window is used. This bias is in addition to any that arises from nonideality in the log and differential amplifiers.

The advantage of using analog over digitization methods is that one does not have to process and fit large quantities of data. A negative feature is that one extracts a cavity decay rate but no direct information on how well the cavity decay transients fit to a single exponential decay. As discussed earlier, CRDS experiments, at least in our experience, have outliers, a small fraction of individual decays that are far from the mean value. In the phase shift detection method, it does not appear that one eliminates these outliers. Even if the analog-determined decay rate is determined after each decay transient, which is possible with the dual gate and log amplifier approaches, one loses the ability to selectively reject decays that do not quantitatively fit a single exponential decay, unless further analog detection channels are added.

## 10. SHOT NOISE LIMITED CAVITY RING-DOWN DATA

Let us now consider the least squares fit in the limit that shot noise dominates over detector noise and that we have a negligible background term  $B$  (i.e., dark current). This should be applicable if the ring-down decay is observed using a multichannel scaler to count photon in time bins after the excitation pulse as was recently reported by Thompson and Myers [27], as long as dark counts (which determine  $B$ ) are sufficiently rare. As before, we can derive a lowest possible variance in  $k'$  as

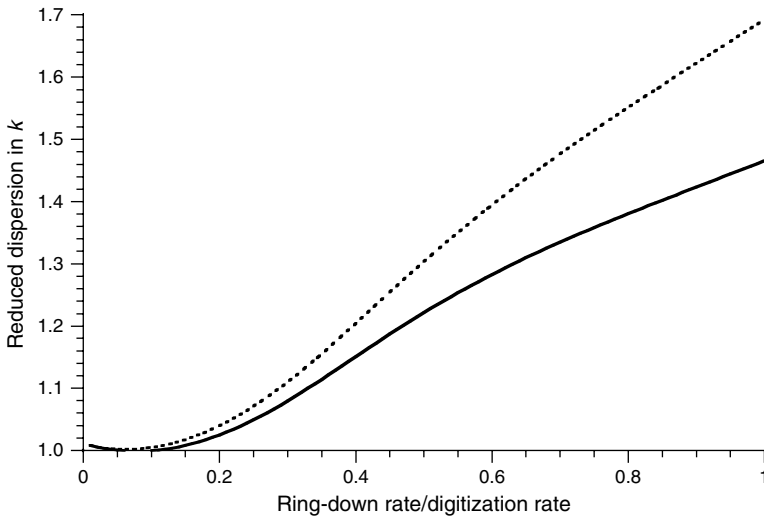
$$\sigma^2(k')_{\text{SNL}} = k^3 \left( \frac{h\nu}{QA} \right), \quad (68)$$

SNL refers to the “shot noise limit” and represents the lowest possible noise in the ring-down rate (and thus highest possible sensitivity for CRDS) for a given mirror reflectivity and transmitted optical power on the detector. This result was previously derived by Romanini and Lehmann [7].

As a point in passing, we note that we are neglecting the possibility of filling the ring-down cavity with amplitude “squeezed” light, and thereby reducing the noise in the ring-down below the shot noise limit. Even if this were done, linear loss destroys squeezing of the light [28], and thus we would expect the cavity output to rapidly evolve shot noise even if it was not present at the beginning of the decay.

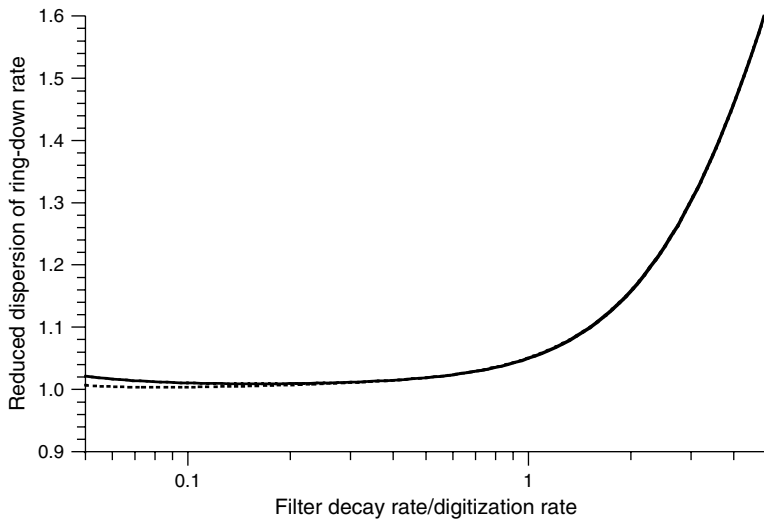
Comparison of Equations (26) and (68) shows that for identical initial signal to noise at the beginning of the ring down, the limiting variance in the case of detector noise will be eight times larger than for the shot noise-limit. This is a consequence of the fact that for the shot noise limited case, the signal to noise of the data points falls only half as fast as for the detector noise limit and thus it is as if the  $k$  value was half as large. Thus, the factor of 8 is a direct consequence of the  $k^3$  dependence of the variance.

Figure 6 shows the reduced dispersion of the ring-down rate (compared to the shot noise limiting value) as a function of  $k\Delta t$  for both correlated and uncorrelated fits

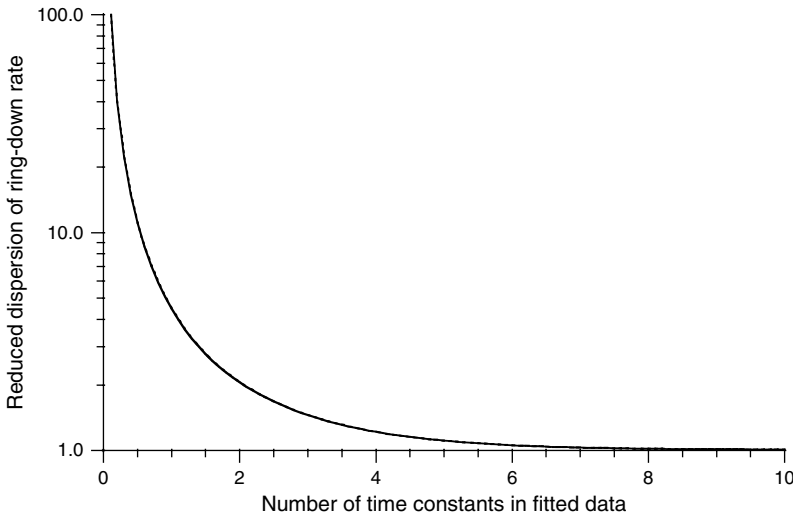


**Figure 6** Plot of the reduced dispersion (fractional increase in the standard deviation over the “ideal” case) of the ring-down rate as a function of the  $k\Delta t$ . Calculations are identical as given in Figure 2 except that the ring down is assumed to be shot noise limited, that is, the noise variance at any time is proportional to the expected light intensity.

to shot noise limited ring-down decays. The decays were fit for in excess of 12 time constants, with  $k_f\Delta t = 0.5$ . Compared to the data presented in Figure 2, it is seen that for modest values of  $k\Delta t \sim 0.1$ , the shot noise-limited case is much closer to the limiting value, particularly for the uncorrelated fit to the data. Figure 7 shows the



**Figure 7** Plot of the reduced dispersion of the ring-down decay rate as a function of  $k_f\Delta t$  for shot noise-limited ring down. Otherwise, calculations are the same as in Figure 4.



**Figure 8** Plot of the reduced dispersion of the ring-down decay rate as a function of  $Nk\Delta t$  for shot noise-limited cavity ring down. Only fits with a fixed baseline are presented for reasons given in text.

reduced dispersion as a function of  $k_f\Delta t$ , calculated with  $k\Delta t = 0.01$  and  $Nk\Delta t = 8$ . These curves look qualitatively similar to the equivalent curves for the detector noise limited case (Figure 3), but again the shot noise-limited case is more forgiving, particularly with regard to using uncorrelated fits. Figure 8 shows the reduced dispersion as a function of  $Nk\Delta t$ , calculated with  $k\Delta t = 0.01$  and  $k_f\Delta t = 0.5$ . Both correlated and uncorrelated fits are plotted but are essentially indistinguishable for these parameters. Here we see that the shot noise limiting case requires that we go about twice as far into the tail of the decay to recover the limiting dispersion. This is as expected given that the signal-to-noise decays half as fast.

If we do a fit to parameters  $A'$ ,  $k'$ , and also assume  $Nk\Delta t \gg 1$ , we can analytically solve the uncorrelated least squares equations (or the correlated equations neglecting the second derivative term in Equation (19)) to give

$$k' = \frac{1}{\Delta t} \ln \left( 1 + \frac{\sum y_i}{\sum iy_i} \right) \tag{69}$$

$$A' = [1 - \exp(-k'\Delta t)] \sum y_i = \frac{(\sum y_i)^2}{\sum (i+1)y_i} \tag{70}$$

This is just the finite step size analog of the well-known integral for an exponential

$$k = \frac{\int_0^\infty A \exp(-kt) dt}{\int_0^\infty t A \exp(-kt) dt},$$

which is applicable in the continuum limit,  $k\Delta t \ll 1$ . Romanini and Lehmann [7] had previously derived that the predicted noise of the integral estimate was the same as for a weighted least squares fit but failed to observe that it is the same estimate. Well into the tail of the exponential decay, the number of photons detected will be small and the fluctuations should follow a Poisson, not Gaussian distribution. However, making the assumption of Poisson distribution for number of detected phonons and selecting  $A'$ ,  $k'$  by the maximum likelihood principle leads to precisely Equations (69) and (70).

The covariance matrix is also easily calculated analytically and is given by

$$\tilde{c} = \frac{Gk_f hv}{2QA} \begin{pmatrix} (1-a^2)A^2 & \frac{(1-a)^2 A}{\Delta t} \\ \frac{(1-a)^2 A}{\Delta t} & \frac{(1-a)^3}{\Delta t^2 a} \end{pmatrix} \quad (71)$$

$$\xrightarrow{k\Delta t \ll 1} \frac{Gk_f hv}{2QA} (k\Delta t) \begin{pmatrix} 2A^2 & 1Ak \\ 1Ak & 1k^2 \end{pmatrix}.$$

In the shot noise limited case, because of the simple analytical solution, there is no reason to introduce the linearization of the decay. Romanini and Lehmann [7] presented the expected noise for the gated integrator approach for the case of shot noise limited detection.

For a finite value of  $Nk\Delta t$ , the least squares equations have terms proportional to  $\exp(-Nk\Delta t)$  and cannot be solved in closed form. If the fit includes a variable baseline correction, one must be careful to specify how the shot noise is estimated. If it is assumed that the baseline error arises from an offset in the electronics, then  $\sigma'^2(t)$  is proportional to  $A \exp(-kt)$ . In this case, the least squares fit equations and the curvature matrix become singular for large  $Nk\Delta t$ , because the model predicts that the signal to noise of the fit to the baseline becomes infinite at large time. However, if the baseline correction,  $B$ , represents dark counts on the photodetector, then  $\sigma'^2(t)$  is proportional to  $A \exp(-kt) + B$ . In this case, the least squares fit equations and the curvature matrix remain stable even as  $Nk\Delta t \rightarrow \infty$ .

Realistic modeling of real data will often require including both shot and detector noise or dark count contributions to  $\sigma'^2(t)$  (see Equation (6)). In this case, one cannot write down closed form expressions for the curvature matrix, because the sums are no longer simple geometric series. Converting the sums to integrals over  $t$  (which should be accurate for  $k\Delta t \ll 1$  and the errors of which can be corrected by using a Euler – MacLaurin summation formula), closed form expressions for the  $\alpha_{AA}$ ,  $\alpha_{AB}$ , and  $\alpha_{BB}$  can be expressed in terms of  $\ln$  and  $\exp$  functions.  $\alpha_{Ak}$  and  $\alpha_{Bk}$  can be expressed in terms of the dilogarithm function while  $\alpha_{kk}$  can be expressed in terms of the trilogarithm function [29, 30]. Alternatively, numerical evaluation of the sums is possible and can be used to calculate the covariance matrix for particular model though at an increased computational cost. We can expect the results to vary smoothly, with covariances to increase monotonically as the contribution of detector noise is introduced into the model and is increased to the point that it dominates the shot noise throughout the decay, which

takes us between the two limiting cases considered explicitly above. One can do a two or three parameter least square fit that treats the noise as constant during the decay, using the closed form expressions for  $\tilde{\alpha}$  and  $\tilde{\beta}$  given in Section 5. In the standard limit ( $k\Delta t \ll 1$ ,  $Nk\Delta t \gg 1$ ), Equation (17) predicts that, for  $B=0$ , the variance of the determined  $k'$  value will be  $(80/27)$ , approximately three times, that given by Equation (68), which applies the shot noise weight to each point. For  $|B| \gg |A|$ , the noise changes little between points, so use of constant weights in the fit should be close to optimal and will never be worse than for the  $B=0$  case just discussed.

## 11. EFFECT OF RESIDUAL MODE BEATING IN THE RING-DOWN DECAY

The model used so far in this chapter has assumed a single exponential decay of the ring-down transient. In many CRDS experiments, one has multimode excitation of the ring-down cavity, and thus the decay contains residual beating terms that arise from the interference of different longitudinal and transverse modes excited in the cavity. The longitudinal mode beating is relatively easy to filter from the detection bandwidth, because it typically occurs on a time scale three to five orders of magnitude shorter than the ring-down time. The effects of excitation of transverse modes can be more problematic, because for a general cavity, the transverse and longitudinal mode spacings are irrationally related [31]; this means that at least some of the potential beating terms will be within the detection bandwidth of the transient that is fit in the experiment. These effects can be mitigated by careful control of the optical alignment and mode matching of the excitation source into the cavity, by an intracavity aperture, by insuring that there are no external limiting apertures, and that the entire transverse profile of the output radiation is focused onto a detector with uniform quantum efficiency [17]. Further, the cavity length can be adjusted to ensure that at least no low-order transverse modes are nearly degenerate with the TEM<sub>00</sub> longitudinal modes of the cavity and also that the digitization of the decay transient does not “alias” one of the stronger beating frequencies so that it appears close to DC. One of the authors has observed, using a digital oscilloscope, a case where the primary longitudinal mode beating near 150 MHz was “aliased” down to  $\sim 100$  kHz, giving the appearance of a strongly modulated ring down curve!

Even with careful attention to these factors, residual mode beating in the detection bandwidth can often not be completely eliminated, particularly if using an excitation source of poor spatial and temporal coherence, as is produced by many multimode pulsed dye lasers. For this reason, it is useful to have an estimate of how much the residual beating will affect the fitted ring down rate. We will model the beating by assuming a ring down of the following form:

$$F(t) = A(1 + c_{mb}\cos(\omega t) + s_{mb}\sin(\omega t))\exp(-kt). \quad (72)$$

If we go the continuum limit,  $k\Delta t$ ,  $\omega\Delta t \ll 1$ , the formulae for error propagation (Equation (15)) can be solved in closed form for the detector and shot noise limiting

cases. For the case of detector noise limited decay, we have the fractional shift in the fitted ring-down rate induced by the mode beating is given by

$$\begin{aligned} \frac{\delta k'}{k} &= \left( \frac{16k^2\omega^2}{(4k^2 + \omega^2)^2} \right) c_{\text{mb}} + \left( \frac{4k\omega(\omega^2 - 4k^2)}{(4k^2 + \omega^2)^2} \right) s_{\text{mb}} \\ &\xrightarrow{\omega \gg k} \left( \frac{16k^2}{\omega^2} \right) c_{\text{mb}} + \left( \frac{4k}{\omega} \right) s_{\text{mb}}. \end{aligned} \quad (73)$$

While in the shot noise limited case we have

$$\begin{aligned} \frac{\delta k'}{k} &= \left( \frac{2k^2\omega^2}{(k^2 + \omega^2)^2} \right) c_{\text{mb}} + \left( \frac{k\omega(\omega^2 - k^2)}{(k^2 + \omega^2)^2} \right) s_{\text{mb}} \\ &\xrightarrow{\omega \gg k} \left( \frac{2k^2}{\omega^2} \right) c_{\text{mb}} + \left( \frac{k}{\omega} \right) s_{\text{mb}}. \end{aligned} \quad (74)$$

If we had several different frequency beating terms in the decay, then under the linear correction terms used here, the total shift in  $k'$  would just have additive contributions. From these expressions, it is clear that for the same fractional amplitude of mode beating in the decay, the shot noise weighted fit is considerably less perturbed. These results were derived for the case of uncorrelated fits to the data (i.e. diagonal weight matrix the correction factor  $G$ ). For the typical case, where  $\omega \gg k$ , including the data correlation, which can be approximated by adding corrections proportional to the second derivative with respect to time of  $\bar{F}(t)$ , will tend to enhance the effect of the mode beating! This is an example where the ideal procedure for the case of random statistical error is not the optimal procedure to use in the case of nonstatistical error.

It is clear from Equations (73) and (74) that the effect of the mode beating will be second order if we start the data collection at a time corresponding to an antinode of the mode beating. This suggests one strategy to minimize the effect of mode beating (assuming we have a single dominant beating frequency) by picking the starting point of the fit to minimize the  $\chi^2$  of the fit. The phase of all the excited modes of the cavity is fixed at  $t=0$  for the case of impulsive excitation of the cavity, and the beat frequencies fixed by the cavity spectrum. Thus, this “best” time should be constant as the excitation spectrum is scanned, and the starting point can be fixed after a series of preliminary measurements. The “optimal” method of data analysis is to include the amplitude of the known mode beating terms in the least squares fit of the data. Including the extra parameter will raise the predicted dispersion of  $k$  due to parameter correlation, but simple calculations indicate that this will be by a minor amount, 1% if  $\omega = 10(2\pi k)$ .

## 12. CONCLUSIONS

This work has presented a statistical analysis of data processing in CRDS. In particular, it is shown that to approach the “ideal” limits in the standard deviation or dispersion of the ring-down decay rate, one must filter the input signals to ensure that the noise is sufficiently well sampled, that is, the noise density aliased to near-zero frequency should be negligible. This filtering, however, introduces correlation of the data. As long as the filter time constant is much less than the ring-down time, the data correlation is found to have only a modest effect on the ring-down time extracted by a least squares fit. However, a least squares fit that ignores the data correlation will predict parameter dispersions substantially below the correct statistical value. The effect of “linearization” of the fit is considered, and it is found to have minimal effect of the least squares fit, provided the signal to noise of the decay is initially sufficiently large. Other methods used to determine cavity decay rates from CRDS data and their noise and bias have been analyzed. As could have been anticipated, all give higher dispersion in the extracted decay rate, which translates into decreased sensitivity in CRDS experiments.

A natural extension of this chapter would be a similar analysis of related cavity-enhanced spectroscopic methods, particularly the cavity-enhanced absorption method [25,26] which is gaining popularity. These alternative methods are not immune to source amplitude and frequency (or phase) noise. The last is most important as one has interference between light stored in the cavity and the instantaneous excitation field. Noise analysis will depend on the details of the temporal coherence of the laser used to excite the cavity.

## ACKNOWLEDGMENTS

The authors’ work in CRDS has been supported by the National Science Foundation, the New Jersey Commission on Science and Technology, and the University of Virginia.

## REFERENCES

- [1] A. O’Keefe and D. A. G. Deacon, *Review of Scientific Instrument* **59**(12), 2544 (1988).
- [2] K. Busch and M. Busch, eds., *Cavity Ringdown Spectroscopy – An Ultratrace-Absorption Measurement Technique*, vol. 720 of *ACS Symposium Series*, (American Chemical Society, Washington, D.C., 1999).
- [3] G. Berden and G. Engel, *Cavity Ring-down Spectroscopy: Techniques and Applications* (Blackwell Publishing, Oxford, 2008).
- [4] J. J. Scherer, J. B. Paul, A. O’Keefe, and R. J. Saykally, *Chemical Reviews* **97**(1), 25 (1997).
- [5] G. Berden, R. Peeters, and G. Meijer, *International Reviews in Physical Chemistry* **19**(4), 565 (2000).
- [6] C. Vallance, *New Journal of Chemistry* **29**(7), 867 (2005).

- [7] D. Romanini and K. K. Lehmann, *Journal of Chemical Physics* **99**(9), 6287 (1993).
- [8] D. L. Albritton, A. L. Schmeltekopf, and R. N. Zare, chap. 1 in Rao, ed., *Molecular Spectroscopy: Modern Research*, vol. II, (Academic Press, New York, 1974).
- [9] T. von Lerber and M. W. Sigrist, *Chemical Physics Letters* **353**(1–2), 131 (2002).
- [10] M. Mazurenka, R. Wada, A. J. L. Shillings, T. J. A. Butler, J. M. Beames, and A. J. Orr-Ewing, *Applied Physics B* **81**(1), 135 (2005).
- [11] D. Halmer, G. Basum, P. Hering, and M. Murtz, *Review of Scientific Instruments* **75**(6), 2187 (2004).
- [12] J. M. Herbelin, J. A. McKay, M. A. Kwok, R. H. Ueunten, D. S. Urevig, D. J. Spencer, and D. J. Benard, *Applied Optics* **19**(1), 144 (1980).
- [13] R. Engeln, G. vonHelden, G. Berden, and G. Meijer, *Chemical Physics Letters* **262**(1–2), 105 (1996).
- [14] T. G. Spence, C. C. Harb, B. A. Paldus, R. N. Zare, B. Willke, and R. L. Byer, *Review of Scientific Instruments* **71**(2), 347 (2000).
- [15] P. R. Bevington and D. K. Robinson, *Data Reduction and Error Analysis for the Physical Sciences* 2nd Edition. (McGraw-Hill Inc., New York, 1992).
- [16] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. Vetterling, *Numerical Recipes* (Cambridge University Press, Cambridge, 1986).
- [17] H. Huang and K. K. Lehmann, *Optics Express* **15**(14), 8745 (2007).
- [18] J. L. Miller and A. J. Orr-Ewing, *Journal of Chemical Physics* **126**(17), 174303 (2007).
- [19] A. A. Istratov and O. F. Vyvenko, *Review of Scientific Instruments* **70**(2), 1233 (1999).
- [20] I. B. C. Matheson, *Analytical Instrumentation* **16**(3), 345 (1987).
- [21] K. K. Lehmann and D. Romanini, *Journal of Chemical Physics* **105**(23), 10263 (1996).
- [22] G. Meijer, M. G. H. Boogaarts, R. T. Jongma, D. H. Parker, and A. M. Wodtke, *Chemical Physics Letters* **217**(1–2), 112 (1994).
- [23] J. Morville, D. Romanini, M. Chenevier, and A. Kachanov, *Applied Optics* **41**(33), 6980 (2002).
- [24] J. B. Dudek, P. B. Tarsa, A. Velasquez, M. Wladyslawski, P. Rabinowitz, and K. K. Lehmann, *Analytical Chemistry* **75**(17), 4599 (2003).
- [25] R. Engeln, G. Berden, R. Peeters, and G. Meijer, *Review of Scientific Instruments* **69**(11), 3763 (1998).
- [26] A. O’Keefe, J. J. Scherer, and J. B. Paul, *Chemical Physics Letters* **307**(5–6), 343 (1999).
- [27] J. E. Thompson and K. Myers, *Measurement Science and Technology* **18**(1), 147 (2007).
- [28] D. Walls and G. Milburn, *Quantum Optics* (Springer, 1995).
- [29] L. Lewin, *Polylogarithms and Associated Functions* (North Holland, New York, 1981).
- [30] L. Vepstas, *An Efficient Algorithm for Accelerating the Convergence of Oscillatory Series, Useful for Computing the Polylogarithm and Hurwitz Zeta Functions* (2007), <http://arxiv.org/abs/math/0702243>.
- [31] A. E. Siegman, *Lasers* (University Science Books, Mill Valley, California, 1986).