23. Fundamentals of Data Processing

In this chapter the fundamentals of statistical pa-
rameter estimation are reviewed for applications
typical in experimental fluid mechanics. The chap-
ter begins with a review of the probability density
function and its moments and continues with
common estimators for the mean and variance of
stationary random processes. A brief introduction
to signal noise is given as a prelude to a rigor-
ous discussion of the Cramér–Rao Lower Bound
(CRLB). The CRLB represents the lower bound of
variance of unbiased estimators of a parameter.
This concept is deepened using illustrations from
the laser Doppler, phase Doppler and PIV mea-
surement techniques. The chapter closes with a
short discussion about the propagation of errors in
a measurement chain.

Part D | 23

23.1 Statistical Principles

The first concept to be introduced is the probability
distribution function $P(x)$, which is the probability as-
signed to a set of points k, such that the random variable
$x(k)$ satisfies $x(k) \leq x$, where x is some fixed value.

$$P(x) = \text{Prob}[x(k) \leq x] . \tag{23.1}$$

This set of points $x(k) \leq x$ is a subset of all points
satisfying $x(k) \leq \infty$. Thus

$$P(-\infty) = 0 , \quad P(\infty) = 1 . \tag{23.2}$$

The *probability density* function (PDF) $p(x)$ is defined
by the relation

$$p(x) = \lim_{\Delta x \to 0} \left(\frac{\text{Prob}\,[x < x(k) \leq x + \Delta x]}{\Delta x} \right) . \tag{23.3}$$

Thus,

$$p(x) \geq 0 , \tag{23.4}$$

$$\int_{-\infty}^{\infty} p(x)\,\mathrm{d}x = 1 , \tag{23.5}$$

$$P(x) = \int_{-\infty}^{x} p(\xi)\,\mathrm{d}\xi ; \quad \frac{\mathrm{d}P(x)}{\mathrm{d}x} = p(x) . \tag{23.6}$$

The next concept to be discussed is that of expected
values. The expected value for any real, single-valued,
continuous function $g(x)$ of the random variable $x(k)$ is
given by

$$E\{g\,[x(k)]\} = \int_{-\infty}^{\infty} g(x)p(x)\,\mathrm{d}x \tag{23.7}$$

in particular, for $g(x) = x$, the mean value of $x(k)$ is
obtained by

$$E\,[x(k)] = \mu_x = \int_{-\infty}^{\infty} x p(x)\,\mathrm{d}x \tag{23.8}$$

and for $g(x) = x^2$, the mean square value of $x(k)$ is
obtained by

$$E\big[x^2(k)\big] = \psi_x^2 = \int_{-\infty}^{\infty} x^2 p(x)\,\mathrm{d}x . \tag{23.9}$$

The quantities defined in (23.8) and (23.9) are also known as the first and second *moments* of the random variable $x(k)$. Note that often $\bar{x}$ is used instead of μ_x for the mean value of $x(k)$. Furthermore, often the variance of $x(k)$, σ_x^2, is used rather than the mean square value,

$$\sigma_x^2 = \psi_x^2 - \mu_x^2 = \int_{-\infty}^{\infty} (x - \mu_x)^2 \, p(x) \, dx . \qquad (23.10)$$

The standard deviation σ_x of $x(k)$ is the square root of the variance. Equation (23.10) is one example of the more general r-th-order *central moment*

$$\mu_r = \int_{-\infty}^{\infty} (x - \mu_x)^r \, p(x) \, dx \qquad (23.11)$$

which quantifies deviations of $x(k)$ about its mean value.

Similar expressions can be written for the bivariate case, in which two random variables $x(k)$ and $y(k)$ are considered. The joint probability function is defined by

$$P(x, y) = \text{Prob}\,[x(k) \le x \text{ and } y(k) \le y] \qquad (23.12)$$

and the associated joint probability density function by

$$p(x, y)$$
$$= \lim_{\substack{\Delta x \to 0 \\ \Delta y \to 0}} \left(\frac{\text{Prob}\left[\begin{smallmatrix} x < x(k) \le x + \Delta x \\ \text{and } y < y(k) \le y + \Delta y \end{smallmatrix}\right]}{\Delta x \, \Delta y} \right) \qquad (23.13)$$

yielding also

$$p(x, y) \ge 0 , \qquad (23.14)$$

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x, y) \, dx \, dy = 1 , \qquad (23.15)$$

$$P(x, y) = \int_{-\infty}^{y} \int_{-\infty}^{x} p(\xi, \eta) \, d\xi \, d\eta , $$

$$\frac{\partial^2 P(x, y)}{\partial x \partial y} = p(x, y) . \qquad (23.16)$$

The two random variables are said to be statistically independent if

$$p(x, y) = p(x)p(y) . \qquad (23.17)$$

The expected value of any real, single-valued, continuous function $g(x, y)$ of two random variables $x(k)$ and $y(k)$ is given by

$$E\,[g(x, y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y)p(x, y) \, dx \, dy \qquad (23.18)$$

One special example is when $g(x, y) = [x(k) - \mu_x] \times [y(k) - \mu_y]$, where μ_x and μ_y are the respective mean values. The expected value is known as the covariance

$$\begin{aligned} C_{xy} &= E\left\{[x(k) - \mu_x][y(k) - \mu_y]\right\} \\ &= E\,[x(k)y(k)] - E\,[x(k)]\,E\,[y(k)] \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [x(k) - \mu_x][y(k) - \mu_y]p(x, y) \, dx \, dy . \end{aligned}$$
$$(23.19)$$

The correlation coefficient is then defined by

$$\rho_{xy} = \frac{C_{xy}}{\sigma_x \sigma_y} \qquad (23.20)$$

which lies between -1 and $+1$.

Data processing deals with the *estimation* of relevant process statistics from the primary measurement quantities. The term *estimation*, rather than determination or computation, is used, since in almost all cases, the physical process has a stochastic part, meaning that the result of an estimation is a random variable (even an exact replication of the experiment would yield a slightly different answer). The procedure or computational algorithm used to obtain the estimation is known as the *estimator*. ("stimatore" in italiano)

Estimators are evaluated on the basis of three properties. First, the expected value of the estimation should be equal to the parameter being estimated

$$E\left(\hat{\phi}\right) = \phi . \qquad (23.21)$$

If this is true, the estimator is *unbiased*. Note that an estimator is often signified by the hat symbol. Second, the mean square error of the estimator should be smaller than for any other possible estimator.

$$E[(\hat{\phi}_1 - \phi)^2] \le E[(\hat{\phi}_i - \phi)^2] . \qquad (23.22)$$

In this case the estimator $\hat{\phi}_1$ is said to be *efficient*. Note that the smallest possible estimation variance for any unbiased estimator is given by the Cramér–Rao lower bound (CRLB). Finally, the estimate should converge to the parameter being estimated for a large sample number or for a long observation time

$$\lim_{N \to \infty} \text{Prob}(|\hat{\phi} - \phi| \ge \varepsilon) = 0 . \qquad (23.23)$$

For an arbitrarily small $\varepsilon > 0$, the estimator is said to be *consistent*. A sufficient condition to meet this requirement is

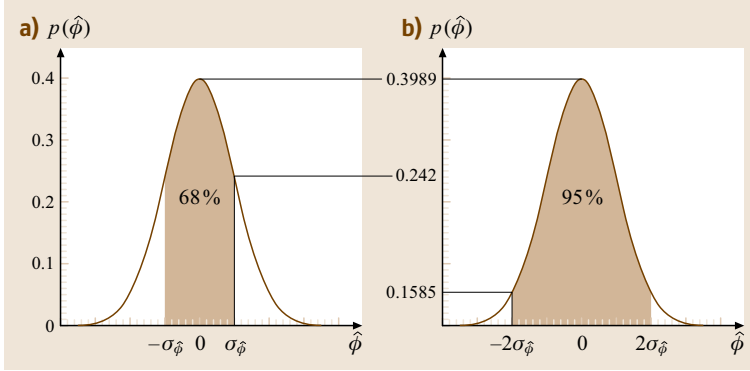$$\lim_{N \to \infty} [(\hat{\phi} - \phi)^2] = 0 . \qquad (23.24)$$

Fig. 23.1a,b Gaussian (normal) distribution illustrating confidence limits. **(a)** For $\pm\sigma$ (68%), **(b)** For $\pm 2\sigma$ (95%)

The mean square error used above can be expanded to yield

$$E[(\hat{\phi} - \phi)^2] = E\{[\hat{\phi} - E(\hat{\phi}) + E(\hat{\phi}) - \phi]^2\}$$
$$= E\{[\hat{\phi} - E(\hat{\phi})]^2\}$$
$$+ E\{[E(\hat{\phi}) - \phi]^2\} . \qquad (23.25)$$

Hence, the mean square error is the sum of two parts: the first part is a variance term that describes the random part of the error

$$\mathrm{var}(\hat{\phi}) = E\{[\hat{\phi} - E(\hat{\phi})]^2\} = E(\hat{\phi}^2) - E^2(\hat{\phi}) , \qquad (23.26)$$

which can be made arbitrarily small by increasing the sample size. The second part is the square of a bias term describing the systematic portion of the error

$$b^2(\hat{\phi}) = E\{[E(\hat{\phi}) - \phi]^2\} . \qquad (23.27)$$

This part is not influenced directly by the sample size and can arise from many sources, often found outside of the data processing. Often special calibration procedures are required to quantify such errors; however, these will not be considered further here. In fact, the bias error will be assumed to be negligible in the following discussion.

Under these conditions and for a *small* normalized random error

$$\varepsilon = \frac{\sigma(\hat{\phi})}{\phi} = \frac{\sqrt{\mathrm{var}(\hat{\phi})}}{\phi} , \qquad (23.28)$$

the probability density function for the estimates, $p(\hat{\phi})$, can often be approximated by a Gaussian distribution with the mean value $E(\hat{\phi}) = \phi$ and a standard deviation $\sigma_{\hat{\phi}} = \varepsilon\phi$

$$p\left(\hat{\phi}\right) = \frac{1}{\varepsilon\phi\sqrt{2\pi}} \exp\left(\frac{-\left(\hat{\phi} - \phi\right)^2}{2\left(\varepsilon\phi\right)^2}\right) . \qquad (23.29)$$

Probability statements about the bounds in which future estimates $\hat{\phi}$ will lie can thus be made as follows

$$\mathrm{Prob}[\phi(1 - \varepsilon) \le \hat{\phi} < \phi(1 + \varepsilon)] \approx 0.68 ,$$
$$\mathrm{Prob}[\phi(1 - 2\varepsilon) \le \hat{\phi} < \phi(1 + 2\varepsilon)] \approx 0.95 , \qquad (23.30)$$

since for a Gaussian distribution $\pm\sigma$ or $\pm 2\sigma$ about the mean contains, respectively, 68% or 95% of the probability mass, as sketched in Fig. 23.1.

This leads directly to the concept of *confidence intervals*, i.e., the interval in which the true value will lie with a given probability (valid for small ε).

$$\hat{\phi}\,(1 - \varepsilon) \le \phi \le \hat{\phi}\,(1 + \varepsilon) \quad \text{with 68% confidence,}$$
$$\hat{\phi}\,(1 - 2\varepsilon) \le \phi \le \hat{\phi}\,(1 + 2\varepsilon) \quad \text{with 95% confidence.} \qquad (23.31)$$

The value of ε can be estimated directly from the sampled data, as discussed in the next section.

23.2 Stationary Random Processes

Given some random phenomena, any single time history of this function is called a sample function. The collection of all possible sample functions, possibly an infinite number, is known as a *random process* or *stochastic process*.

The mean value (first moment) of the ensemble of sample functions at time t_1 is then the arithmetic mean over the instantaneous values of the sample functions at time t_1, as illustrated in Fig. 23.2. A correlation or joint moment of the process at two different times can be

computed by taking the ensemble average of the product of instantaneous values at two times t_1 and $t_1 + \tau$. These values can be written as

$$\mu_x(t_1) = \lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} x_k(t_1), \qquad (23.32)$$

$$R_{xx}(t_1, \tau) = \lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} x_k(t_1)x_k(t_1 + \tau), \quad (23.33)$$

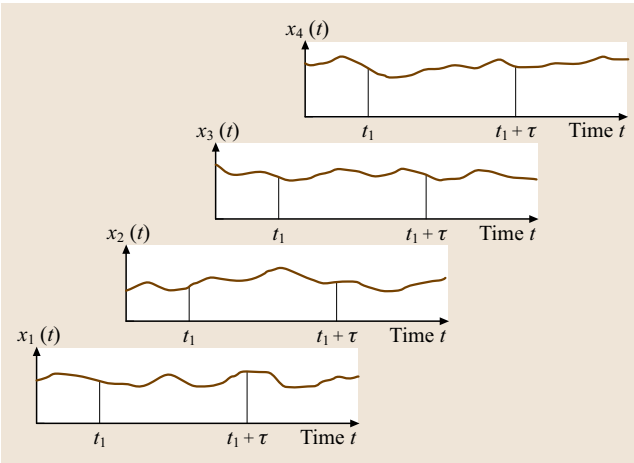where R_{xx} is known as the autocorrelation function.

Fig. 23.2 Ensemble of sample functions defining a random process

A random process is known as *weakly stationary* when the value defined by (23.32) is independent of t_1 and the autocorrelation is only a function τ, the variance is limited and the mean value is constant. The process is known as *strongly stationary* when also the entire probability density function is independent of t_1. Otherwise the process is *instationary*.

Generally, however, statistics of a stationary random process are not computed over an ensemble of sample functions but over a time average. For example,

$$\mu_x(k) = \lim_{T\to\infty} \frac{1}{T} \int_0^T x_k(t)\,dt = \mu_x, \qquad (23.34)$$

$$R_{xx}(\tau, k) = \lim_{T\to\infty} \frac{1}{T} \int_0^T x_k(t)x_k(t + \tau)\,dt = R_{xx}(\tau). \qquad (23.35)$$

If these values do not differ from those in (23.32) and (23.33), then the process is said to be *ergodic*, in which case the index k is dropped. All stationary processes encountered in fluid mechanics can be considered ergodic.

Note that the covariance function is simply the autocorrelation function with the mean removed and the cross-covariance function is the cross-correlation function with the product of the means removed

$$C_{xx}(\tau) = R_{xx}(\tau) - \mu_x^2,$$
$$C_{xy}(\tau) = R_{xy}(\tau) - \mu_x\mu_y. \qquad (23.36)$$

23.3 Estimator Expectation and Variance

In many cases the expectation and variance of an estimator can be derived analytically and several examples are given below. For more complicated quantities, this is not always possible and other strategies can be followed. The jackknife algorithm will be introduced as one such approach.

23.3.1 Estimators for the Mean

The first estimator to be examined is the mean value. The most common sample mean estimator is given by

$$\hat{\mu}_x = \frac{1}{N} \sum_{i=1}^{N} x_i, \qquad (23.37)$$

where x_i are individual samples of the process x. Instead of $\hat{\mu}_x$ the alternative expression $\bar{x}$ is also commonly used.

However, the difference between the estimated value and the true mean value can be shown better using $\hat{\mu}_x$ and μ_x respectively. The estimator of the mean value given in (23.37) is non-biased, since $E(\hat{\mu}_x) = \mu_x$ [23.1]. The mean square error, or variance, of this estimator is then given by

$$\text{var}(\hat{\mu}_x) = \sigma_{\hat{\mu}_x}^2 = E\left[(\hat{\mu}_x - \mu_x)^2\right]. \qquad (23.38)$$

Substituting (23.37) into (23.38) leads to

$$\sigma_{\hat{\mu}_x}^2 = E\left[\left(\frac{1}{N}\sum_{i=1}^{N} x_i - \mu_x\right)^2\right]$$

$$= \frac{1}{N^2} E\left[\left(\sum_{i=1}^{N}(x_i - \mu_x)\right)^2\right]. \qquad (23.39)$$

ERR.

If the condition $E(x_i x_j) = 0$ is satisfied, i. e., consecutive samples are uncorrelated or statistically independent, (23.39) can be further reduced to

$$\sigma_{\hat{\mu}_x}^2 = \frac{1}{N^2} E\left[\sum_{i=1}^{N} (x_i - \mu_x)^2\right]$$

$$= \frac{\sigma_x^2}{N}, \tag{23.40}$$

which states that the variance of the mean estimator decreases with increasing number of samples.

This analysis has been performed for an estimator based on discrete samples x_i; however, a similar analysis could be made for a mean estimator based on the continuous signal $x(t)$

$$\hat{\mu}_x = \frac{1}{T} \int_0^T x(t)\,dt, \tag{23.41}$$

which differs from the true mean μ_x, since the integral is performed only over a finite time T. The variance of this estimator becomes

$$\text{var}(\hat{\mu}_x) = \sigma_{\hat{\mu}_x}^2 = E\left[(\hat{\mu}_x - \mu_x)^2\right] = E\left(\hat{\mu}_x^2\right) - \mu_x^2. \tag{23.42}$$

In terms of the autocovariance function, this can be written as [23.1]

$$\sigma_{\hat{\mu}_x}^2 = \frac{1}{T} \int_{-T}^{T} \left(1 - \frac{|\tau|}{T}\right) C_{xx}(\tau)\,d\tau \tag{23.43}$$

for a stationary random process. For small τ only C_{xx} remains in the integral and for large τ, C_{xx} goes to zero, thus the integral can be expressed as

$$\sigma_{\hat{\mu}_x}^2 = \frac{2\sigma_x^2 T_x}{T} \tag{23.44}$$

with the integral time scale

$$T_x = \frac{1}{\sigma_x^2} \int_0^{\infty} C_{xx}(\tau)\,d\tau. \tag{23.45}$$

As pointed out by *George* [23.2], if the results given by (23.40) and (23.44) are equated, the condition for statistically independent samples can be obtained, namely

$$N = \frac{T}{2T_x}. \tag{23.46}$$

This is graphically represented in Fig. 23.3 and leads to two very insightful interpretations.

- Samples are statistically independent if they are separated by a period of the least $2T_x$ in time.
- Segments of the continuous signal $2T_x$ in length contribute to the mean estimate as one, statistically independent sample.

The manifestation of this relation is that sampling a signal with time intervals less than $2T_x$ will not accelerate the convergence of the mean estimator. At this point, the difference between *data* and *information* should become very clear. New information (with respect to the mean estimate), comes only every $2T_x$ time periods.

Equation (23.44) makes a statement about the necessary observation or measurement time to achieve a given statistical uncertainty (variance of the mean estimator). However, to use this equation the integral time scale, as defined using the autocovariance function, must be known beforehand. Moreover, the integral time scale may change by orders of magnitude, e.g., in flow fields between different points of a single velocity profile. Often, however, a simple estimate of T_x suffices. This will be illustrated with the following example of how (23.44) can be used in practice.

The example chosen is a velocity measurement in the recirculation zone of a backward-facing step water flow. In a preliminary measurement the local variance of the velocity fluctuations is estimated to be $0.2\,\text{m}^2/\text{s}^2$ at point A (Fig. 23.4). The requirement is that the mean velocity at point A be determined to within $\pm0.04\,\text{m/s}$ with 95% confidence.

The integral time scale of the velocity fluctuations can be estimated from appropriate velocity and length scales, in this case $U_0 = 2\,\text{m/s}$ and x_R, which is approximately $8H$ or $0.4\,\text{m}$. Thus, $T_u = x_R/U_0 = 0.2\,\text{s}$. Note that the subscript u for the integral time scale is used, since the process being measured is the velocity u. Assuming a normal distribution for the scatter of the estimates, the probability of being within $\pm\sigma_{\hat{\mu}}$ of the true mean value would be about 68%. This would increase to the required 95% for $\pm2\sigma_{\hat{\mu}_u}$.

[margin annotation: ERR. $\sigma_{\hat{\mu}_u}$]

$$2\sigma_{\hat{\mu}_u} = 0.04\,\text{m/s}, \quad \sigma_{\hat{\mu}_u}^2 = 0.0004\,\text{m}^2/\text{s}^2. \tag{23.47}$$

Equation (23.44) can now be solved for the required measurement time to fulfill this condition

$$T = \frac{2\,\sigma_u^2 T_u}{\sigma_{\hat{\mu}_u}^2} = 200\,\text{s}. \tag{23.48}$$

[margin annotation: ← ERR.]

Note that this calculation has been performed independent of the choice of measurement technique. In fact, no measurement technique can shorten the necessary observation time given in (23.48), since this

Part D | 23.3

$x(t)$
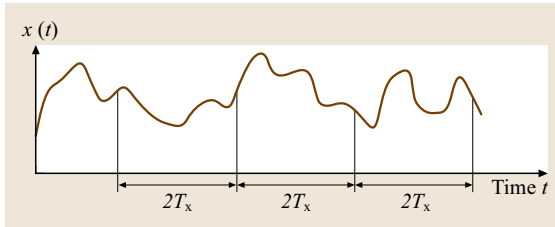
Time t

$2T_x$ $2T_x$ $2T_x$

Fig. 23.3 Graphical interpretation of statistical independence of consecutive samples of a continuous process

describes the fundamental statistical behavior of a random process.

In practice, it is unusual to make such calculations prior to every measurement. It is more convenient to display the current measured mean velocity online, accumulated over all samples up to that time, and then to allow the user to terminate the measurement when the fluctuations of the mean are below an acceptable level. Indeed, from the necessary measurement duration, and from the fluctuation level of the mean, a rough estimate of the integral time scale can often be made. This technique of user intervention does not lend itself to automation, so that still a third approach is often used, in which a fixed number of samples is used for each point, whereby the number is chosen to be very large to ensure sufficient convergence for all measurement points. In many flows there are regions where data rates decrease dramatically, e.g., near walls. In such cases there is often no choice but to accept a higher degree of statistical uncertainty, since otherwise the data collection time becomes excessive.

Alternatively, (23.40) could have been used if the velocity data were available in discrete form at regular time intervals. Assuming the sample rate was not faster than every $2T_u$, the number of samples required to insure

$u_0 = 2\,\mathrm{m/s}$ A $\sigma_u^2 \approx 0.2\,\mathrm{m^2/s^2}$

$H = 5\,\mathrm{cm}$

x

$x_R \approx 0.4\,\mathrm{m}$

Fig. 23.4 Sketch of example backward facing step flow. x_R is the mean reattachment length

the requested accuracy would be

$$N = \frac{\sigma_u^2}{\sigma_{\hat{\mu}_u}^2} = \frac{0.2\,\mathrm{m^2/s^2}}{0.0004\,\mathrm{m^2/s^2}} = 500 \,. \qquad (23.49)$$

This discussion puts into perspective expressions such as 'high' or 'low' data rates or 'many' or 'few' samples. The data rate, or the number of samples, must always be considered with respect to the integral time scale of the process at the particular measurement point. This explains the preferred use of data *density* rather than data *rate*. It should also be apparent that, for the same Reynolds number, measurements performed in air flows will typically be much shorter in duration than in water flows, given the same target accuracy. The reason for this lies in the fact that, for the same Reynolds number, the integral time scale of an air flow is generally shorter.

Further guidelines for reporting measurement uncertainties can be found in *Kline* and *McClintock* [23.3], *Kline* [23.4] or *Moffat* [23.5, 6].

23.3.2 Estimators for Higher-Order Statistics

In the study of turbulence, statistics of not only the mean velocity but also of higher-order moments are required. General formula for the estimator variance for higher-order statistics have been given by *Stuart* and *Ord* [23.7] and *Kendall* and *Stuart* [23.8]. *Benedict* and *Gould* [23.9] have summarized their results in the following manner.

An unbiased estimator of the r-th central moment μ_r (23.11) is given by

$$\hat{\mu}_r = \frac{1}{N} \sum_{i=1}^{N} \left(x_i - \hat{\mu}_x \right)^r \qquad (23.50)$$

in which the true mean has been replaced by the sample mean (23.37). Strictly this estimator is unbiased only for $r = 1$, however this also applies for higher moments when N is large. The sampling variance of $\hat{\mu}_r$ is given by

$$\mathrm{var}\left(\hat{\mu}_r \right) = \sigma_{\hat{\mu}_r}^2 = \frac{1}{N} \left(\mu_{2r} - \mu_r^2 + r^2 \mu_{r-1} \mu_2 \right.$$
$$\left. - 2r\mu_{r+1}\mu_{r-1} \right), \qquad (23.51)$$

where terms of order N^{-2} and higher have been neglected. 95% confidence intervals are then $\hat{\mu}_r \pm 2\sigma_{\hat{\mu}_r}$. Note that (23.51) uses $\mu_1 = \mu_0 = \mu_{-1}$ and $\mu_2 = \sigma_x^2$. Furthermore, it uses the exact central moments μ_r,
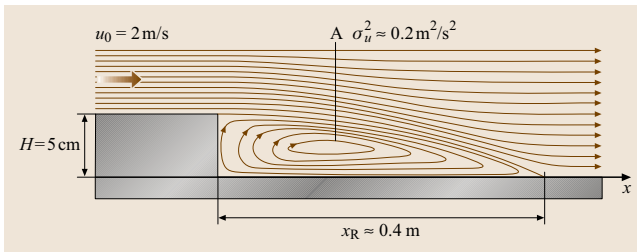
which are actually unknown. However if N is suitably large, typically $N = 100$, these can be replaced by the central moment sampling statistics, $\hat{\mu}_r$, for practical computations.

Similarly, the mixed central moment

$$\mu_{r,s} = \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} (x - \mu_x)^r \, (y - \mu_y)^s \, p(x)p(y)\,\mathrm{d}x\,\mathrm{d}y \tag{23.52}$$

can be estimated using

$$\hat{\mu}_{r,s} = \frac{1}{N} \sum_{i=1}^{N} (x_i - \hat{\mu}_x)^r (x_i - \hat{\mu}_y)^s \tag{23.53}$$

which exhibits the variance

$$\begin{aligned}
\mathrm{var}(\hat{\mu}_{r,s}) &= \sigma_{\hat{\mu}_r}^2 \\
&= \frac{1}{N}\big(\mu_{2r,2s} - \mu_{r,s}^2 + r^2\mu_{2,0}\mu_{r-1,s} \\
&\quad + s^2\mu_{0,2}\mu_{r,s-1}^2 + 2rs\mu_{1,1}\mu_{r-1,s}\mu_{r,s-1} \\
&\quad - 2r\mu_{r+1,s}\mu_{r-1,s} - 2s\mu_{r,s+1}\mu_{r,s-1}\big) .
\end{aligned} \tag{23.54}$$

Note that $\mu_{10} = \mu_{01} = 0$, $\mu_{r,-1} = \mu_{-1,s} = 0$, $\mu_{20} = \sigma_x^2$ and $\mu_{02} = \sigma_y^2$. Equation (23.54) can be simplified for normally distributed processes, since then all odd moments are zero and the second, fourth, sixth and eighth moments are 1, 3, 15 and 105 times σ_x^2, respectively.

The variances of the most common statistics in turbulence research are summarized in Table 23.1, for both an arbitrary and a normal distribution of the process. Note that the formulas given in Table 23.1 are multiplied by N. The u and v velocity components have been used for illustration. As an example, the variance of the

mean estimator is given as $\sigma_u^2 N^{-1}$, which agrees with (23.40).

The expressions in Table 23.1 all assume statistical independence between samples, as specified by (23.46). If the sample rate is too high to insure statistical independence, the total number of samples N must be adjusted so that the total observation time yields the desired confidence bounds, according to (23.44). Furthermore, it should be noted that turbulence quantities are seldom normally distributed, so that the simplifications given in Table 23.1 can lead to significant errors if normality is not previously established.

For more-complex estimators, there exist several resampling algorithms with which the uncertainty of the measured quantity can be estimated. In particular the jackknife algorithm will be discussed, as first introduced by *Tukey* [23.10]. Notes on its practical implementation are given by *Efron* and *Tibshirani* [23.11] and an evaluation of its potential with laser Doppler data is given by *Benedict* and *Gould* [23.9]. This algorithm also assumes statistical independence in the data set $\boldsymbol{x} = (x_1, x_2, \ldots, x_N)$ when computing some statistical estimator. The jackknife samples

$$\boldsymbol{x}_{\mathrm{jack},i} = (x_1, x_2, \ldots, x_{i-1}, x_{i+1}, \ldots, x_N) \tag{23.55}$$

are obtained by leaving out in turn one of the data samples. The jackknife samples are then used to compute N estimates $\hat{\phi}_{\mathrm{jack},i}$ with $i = 0, 1, \ldots, (N-1)$. The jackknife variance for $\hat{\phi}$ is then given by

$$\mathrm{var}(\hat{\phi})_{\mathrm{jack},i} = \frac{N-1}{N} \sum_{i=1}^{N} (\hat{\phi}_{\mathrm{jack},i} - \mu_{\hat{\phi}_{\mathrm{jack}}})^2 , \tag{23.56}$$

Table 23.1 Estimator variances multiplied by N [23.9]

Statistic	Variance for any distribution	Normal assumption
μ_u	σ_u^2	σ_u^2
σ_u	$\frac{\mu_4 - \sigma_u^4}{4\sigma_u^2}$	$\frac{\sigma_u^2}{2}$
$R_{uv} = \mu_{1,1}$	$\mu_{2,2} - \mu_{1,1}^2$	$(1 + \rho_{uv}^2)\sigma_u^2\sigma_v^2$
$\rho_{uv} = \frac{\mu_{1,1}}{\sigma_u \sigma_v}$	$\rho_{uv}^2 \left[\frac{\mu_{2,2}}{\mu_{1,1}^2} + \frac{1}{4}\left(\frac{\mu_{4,0}}{\sigma_u^4} + \frac{\mu_{0,4}}{\sigma_v^4} + \frac{2\mu_{2,2}}{\sigma_u^2\sigma_v^2}\right) - \left(\frac{\mu_{3,1}}{\mu_{1,1}\sigma_u^2} + \frac{\mu_{1,3}}{\mu_{1,1}\sigma_v^2}\right) \right]$	$(1 - \rho_{uv}^2)^2$
σ_u^2	$\mu_4 - \sigma_u^4$	$2\sigma_u^4$
μ_3	$\mu_6 - \mu_3^2 - 6\mu_4\sigma_u^2 + 9\sigma_u^6$	$6\sigma_u^6$
$\mu_{2,1}$	$\mu_{4,2} - \mu_{2,1}^2 + \sigma_u^4\sigma_v^2 + 8\sigma_u^2\mu_{1,1}^2 - 2\sigma_u^2\mu_{2,2} - 4\mu_{1,1}\mu_{3,1}$	$2(1 + 2\rho_{uv}^2)\sigma_u^4\sigma_v^2$
μ_4	$\mu_8 - \mu_4^2 - 8\mu_5\mu_3 + 16\mu_3^2\sigma_u^2$	$96\sigma_u^8$

where

$$\mu_{\hat{\phi}_{\mathrm{jack}}} = \frac{1}{N} \sum_{i=1}^{N} \hat{\phi}_{\mathrm{jack},i} \; . \tag{23.57}$$

The 95% confidence interval for the estimator is given by $\hat{\phi} \pm 2\mathrm{var}((\hat{\phi})_{\mathrm{jack}})^{1/2}$.

The jackknife algorithm requires N^2 calculations per variance estimate. This computational load can be greatly reduced if the programming is modified specifically for each statistic to be studied. For example, if the mean square of the velocity fluctuations σ_u^2, is being studied, the jackknife sample can be written

$$\hat{\phi}_{\mathrm{jack},i} = \hat{\sigma}_{u,\mathrm{jack},i}^2 = \frac{1}{N-1} \sum_{\substack{j=1 \\ j \neq i}}^{N} (u_j - \hat{\mu}_{u,\mathrm{jack},i})^2 \; . \tag{23.58}$$

This equation can be rewritten as

$$\hat{\sigma}_{u,\mathrm{jack},i}^2 = \frac{1}{N-1} \left[\sum_{\substack{j=1 \\ j \neq i}}^{N} u_j^2 - 2\hat{\mu}_{u,\mathrm{jack},i} \sum_{\substack{j=1 \\ j \neq i}}^{N} u_j \right.$$
$$\left. + (N-1)(\hat{\mu}_{u,\mathrm{jack},i})^2 \right] \; . \tag{23.59}$$

Each term in the brackets is summed only once over all $j = 1, \ldots N$ and then decremented by u_j^2 and u_j respectively for each jackknife replication.

It can be shown theoretically that the jackknife is biased high on its estimation of uncertainty and thus, it will never underestimate the uncertainty of a statistic.

23.4 Signal Noise

Noise is essentially any amplitude deviation of an individual realization of a signal from its ideal model. Sources of noise are manifold, e.g., shot noise or thermal noise in any electronics, additional generation-recombination and modulation noise in semiconductors, the photon noise for optical components or quantization noise.

The power of signal fluctuations σ_s^2 put into relation with the power of noise fluctuations σ_n^2 is known as the signal-to-noise ratio (SNR) and is generally expressed in decibels:

$$\mathrm{SNR}/\mathrm{dB} = +10 \log \left(\frac{\sigma_s^2}{\sigma_n^2} \right) \; . \tag{23.60}$$

The estimation of the SNR from a given signal segment is often required in the detection/validation step of signal processing to indicate whether a result can be expected to be reliable or not. Unfortunately, the estimation of the SNR directly from the signal is complicated since the signal fluctuations and the noise are superimposed. However, noise contributions in the system are usually considered to be spectrally white. This refers to the fact that the total noise power is distributed evenly over all frequencies up to the upper bandwidth of the system. Still, the spectral distributions of the useful signal and the noise are superimposed. However, if the bandwidth of the expected signal model is limited, the noise power can be derived from the frequency spectrum.

As an example, in Fig. 23.5, a laser Doppler signal, a noise signal and the summation of the two in time, spectral and correlation domain is illustrated. It becomes obvious from Fig. 23.5 that the power spectral density (PSD) or the autocorrelation function (ACF) offer excellent means to monitor SNR and to determine whether a particle signal is present or not.

An idealized graphical interpretation of SNR is given in Fig. 23.6, which shows schematically the PSD of a Doppler signal logarithmically scaled. The SNR is given by the ratio of the areas A to B. A more detailed estimation procedure is given by *Tropea* [23.12]. The noise appears as a base line floor of width Δf, the bandwidth of the system, and of amplitude σ_n^2 / f_s. Any filtering of the signal, for instance using a low-pass filter, will directly decrease area B and thus increase the SNR, since more of the noise is removed. The use of a band-pass filter to increase SNR increases the reliability of the signal detection, since the SNR acceptance threshold can be chosen higher. In contrast, the variance of the frequency estimation remains constant because the peak in the spectrum still has the same width. Indeed, such adjustable input filters are usually an integral part of any Doppler signal processor. On the other hand there is a danger in filtering with too narrow a bandwidth, since in general the signal frequency is not known a priori. This can lead to truncation of the velocity distribution and to a bias of the estimated moments.
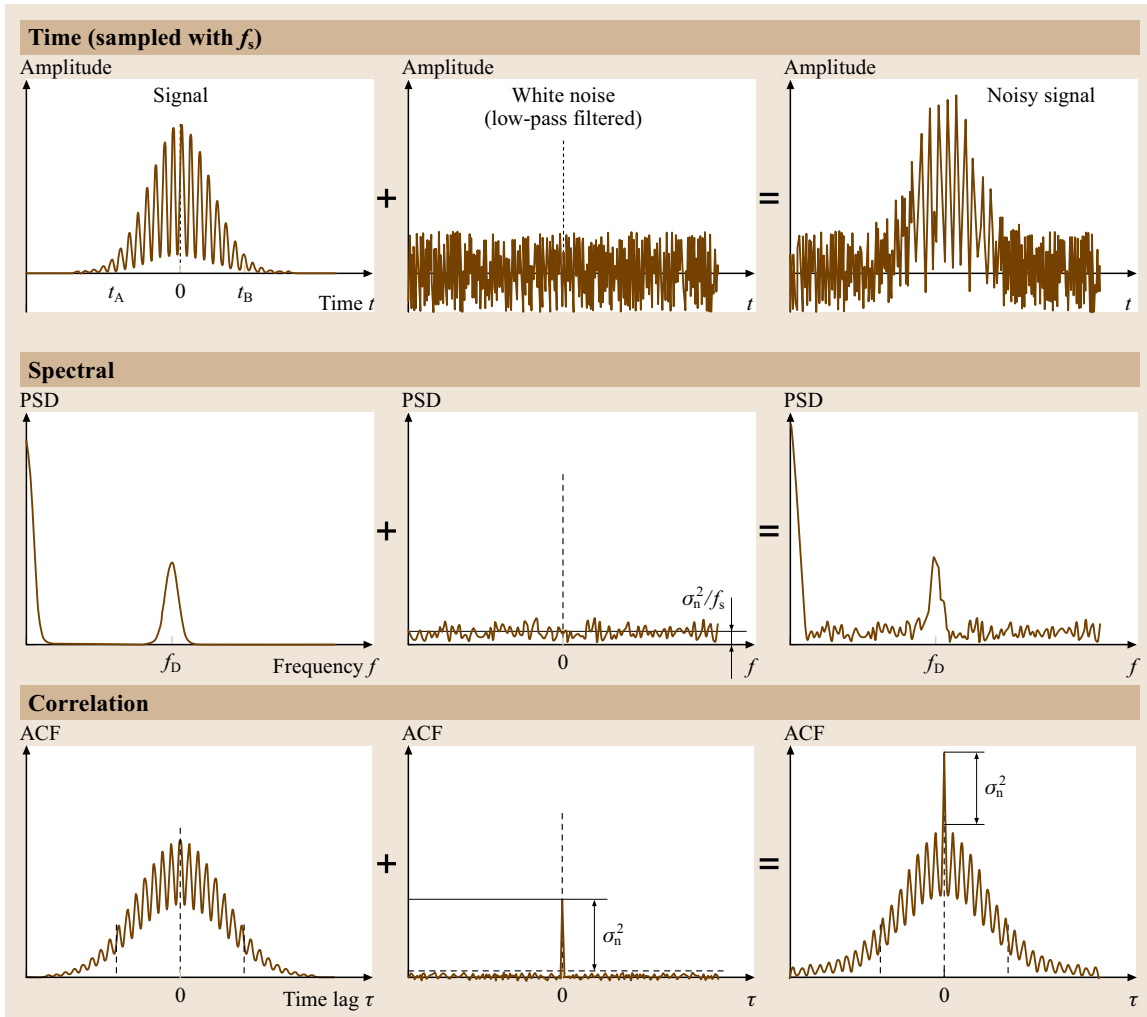
Fig. 23.5 Representation of a laser Doppler signal, a noise signal and a combination of the two in time, spectral and correlation domain

The SNR can also be estimated from the auto-correlation function. Noise, being fully stochastic and having zero correlation duration, appears only in the first autocorrelation coefficient, i. e., $R(\tau = 0)$. The statistical scatter (error) of the autocorrelation coefficients increases with SNR for every $\Delta\tau$, given a finite number of samples. Thus, the SNR can be estimated by comparing the amplitude of the autocorrelation function at $\tau = 0$ to the maximum peak amplitude of the remaining periodicity, exemplary shown in Fig. 22.6c, Chap. 22 for a high-pass filtered burst signal with added noise. If the frequency of the periodicity f has already been determined, the amplitude of the signal A_R
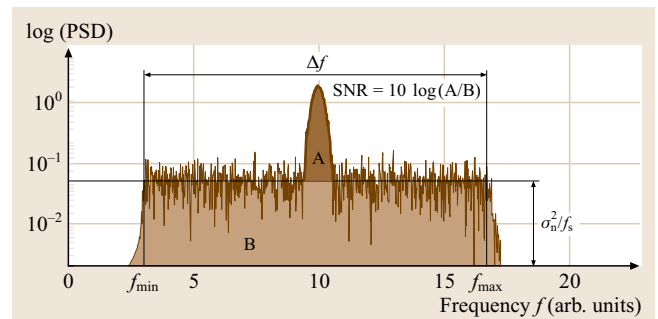

Fig. 23.6 Graphical representation of SNR using the power spectral density (PSD)

(Fig. 22.6, Chap. 22) can be estimated by fitting a co-sine wave to the measured correlation function at points removed from $\tau = 0$. This can be computed using the expression

$$A_R \approx \frac{R(n\Delta\tau)}{\cos(2\pi \, fn\Delta\tau)} \tag{23.61}$$

from which the variance of the noise portion of the signal can be computed

$$\sigma_n^2 = R(0) - A_{R^2} . \tag{23.62}$$

The index n should ideally be chosen at the first maximum or minimum removed from $\tau = 0$.

The SNR is then given as

$$\text{SNR/dB} = 10\log\left(\frac{A_{R^2}}{\sigma_n^2}\right) . \tag{23.63}$$

The presence of noise in the signal can have different effects on the estimation of signal parameters or statistics. First of all, noise can directly affect the signal processing and cause systematic errors of the parameter extraction from individual signals. Furthermore, noise can disturb the signal detection, which often leads to a biased selection of signal realizations. In both cases, the derived signal statistics are biased. But even without systematic errors, the possible accuracy of the parameter extraction is limited, known as the Cramér–Rao lower bound (CRLB), which is discussed in more detail in the following section.

23.5 Cramér–Rao Lower Bound (CRLB)

The goal of analyzing an acquired signal is to derive several signal parameters according to a given model representing the physical basis of the signal generating process, e.g., a signal frequency, phase or amplitude. In practical cases, the recorded signals are equidistantly sampled and time limited, so that the amount of available information is also finite. Furthermore, the signal is influenced by noise and this introduces uncertainty into any parameter determined from the signal. The calculation of signal parameters is therefore called *estimation*, since it contains a random component.

The true values of the parameters to be estimated are seldom known and different estimation algorithms (*estimators*) will also yield different results. Therefore, it is of interest to quantify the accuracy of each estimator statistically. To begin with, the *expectation* of the estimator should be equal to the true value, i.e., non-biased. Second, the estimator should be *efficient*, meaning that it uses all available information to estimate the required parameter as accurately as possible. In Sect. 23.1, features of estimators are discussed in more detail. The efficiency of an estimator is quantified by its *variance*. While the bias should be zero, the finite amount of information yields a lower bound of achievable accuracy and thus, a finite variance. For unbiased estimators this lower bound of variance is given by the CRLB [23.13, 14].

No unbiased estimator can obtain estimates with a variance smaller than the CRLB, thus this quantity can be used to evaluate the performance of a spe-cific algorithm. On the other hand, the CRLB gives no information about how an algorithm should process a measured signal to reach this lower bound. However, based on estimation theory, and closely related to the CRLB, the maximum-likelihood (ML) estimator can be derived. If any unbiased estimator reaches the CRLB, then the ML estimator will also reach it, at least asymptotically [23.13, 15].

For a signal

$$x(t = t_i) = x_i = m_i + n_i , \quad i = 0, 1, \ldots, (N-1) \tag{23.64}$$

with

$$\boldsymbol{x} = \begin{pmatrix} 0 \\ x_1 \\ \vdots \\ x_{N-1} \end{pmatrix}, \quad \boldsymbol{m} = \begin{pmatrix} m_0 \\ m_1 \\ \vdots \\ m_{N-1} \end{pmatrix},$$

$$\boldsymbol{n} = \begin{pmatrix} n_0 \\ n_1 \\ \vdots \\ n_{N-1} \end{pmatrix}, \quad \boldsymbol{x} = \boldsymbol{m} + \boldsymbol{n} \tag{23.65}$$

consisting of the model signal $\boldsymbol{m}$ of known type, the noise $\boldsymbol{n}$ and the unknown (scalar) parameter a, the CRLB is

given by

$$\sigma_a^2 \geq \frac{1}{E\left[\left(\frac{\partial \ln p(x,a)}{\partial a}\right)^2\right]} = -\frac{1}{E\left[\frac{\partial^2 \ln p(x,a)}{\partial a^2}\right]},$$

(23.66)

where $p(x, a)$ is the joint probability density function, (Sect. 23.1), of the measured signal x for a given parameter a. Since a is normally a vector, (23.66) is the inverse of a matrix, the Fisher information matrix $\mathbf{J}$, whose typical element is given by

$$J_{ij} = E(H_{a_i} H_{a_j}) = -E(H_{a_i a_j})$$

(23.67)

with

$$H_{a_i} = \frac{\partial \ln p(x, a)}{\partial a_i} \ . \quad a = \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_{A-1} \end{pmatrix}$$

(23.68)

The bound of the i-th unknown element of the parameter vector a is given by the i-th diagonal element with index ii of the inverse Fisher information matrix

$$\sigma_{a_i}^2 \geq \left(\mathbf{J}^{-1}\right)_{ii} ,$$

(23.69)

where no summation is implied. For uncorrelated and signal independent noise with power σ_n^2, and with a Gaussian distribution, the joint probability density function $p(x, a)$ becomes

$$p(x, a) = \left(\frac{1}{2\pi\sigma_n^2}\right)^{\frac{N}{2}} \exp\left[-\frac{1}{2\sigma_n^2} \sum_{i=0}^{N-1} (x_i - m_i)^2\right]$$

(23.70)

and the elements of the Fisher information matrix become [23.16]

$$J_{ij} = \frac{1}{\sigma_n^2} \sum_{k=0}^{N-1} \left(\frac{\partial m_k}{\partial a_i} \frac{\partial m_k}{\partial a_j}\right).$$

(23.71)

To derive the lower bounds for a given Doppler burst this can be calculated and inverted, at least numerically.

23.5.1 Laser Doppler and Phase Doppler Signals

As an example, the lower bounds for the estimation of the frequency ω and the phase φ will be derived from a laser Doppler-like signal. Since the parameter vector a must contain all unknown parameters, including

those that are not estimated (hidden parameters), a constant amplitude of unity during the observation time is assumed for simplification. The time-dependent signal

$$x(t) = m(t) + n(t)$$

(23.72)

is composed of the model signal

$$m(t) = \cos(\omega t + \varphi)$$

(23.73)

and the time-dependent noise $n(t)$. The measured signal after sampling is therefore

$$x_i = x(t = t_i) = \cos(\omega t_i + \varphi) + n_i , \\ i = 0, 1, \dots, (N-1)$$

(23.74)

and each sample is a function of the two model parameters $x_i(\omega, \varphi)$ and the noise. The sampling times are given by $t_i = i/f_s$. The noise n is uncorrelated and Gaussian distributed. The parameter vector is

$$a = \begin{pmatrix} \omega \\ \varphi \end{pmatrix} .$$

(23.75)

The derivatives of the model-parameter-dependent samples $m_i = \cos(\omega t_i + \varphi)$ are

$$\frac{\partial m_i}{\partial \omega} = -t_i \sin(\omega t_i + \varphi) , \\ \frac{\partial m_i}{\partial \varphi} = -\sin(\omega t_i + \varphi) .$$

(23.76)

The Fisher information matrix becomes

$$\mathbf{J} = \frac{1}{\sigma_n^2} \\ \times \begin{pmatrix} \sum_{i=0}^{N-1} t_i^2 \sin^2(\omega t_i + \varphi) & \sum_{i=0}^{N-1} t_i \sin^2(\omega t_i + \varphi) \\ \sum_{i=0}^{N-1} t_i \sin^2(\omega t_i + \varphi) & \sum_{i=0}^{N-1} \sin^2(\omega t_i + \varphi) \end{pmatrix} .$$

(23.77)

The inverse of the Fisher information matrix is

$$\mathbf{J}^{-1} = \frac{\sigma_n^2}{\det(\mathbf{J})} \\ \times \begin{pmatrix} \sum_{i=0}^{N-1} \sin^2(\omega t_i + \varphi) & -\sum_{i=0}^{N-1} t_i \sin^2(\omega t_i + \varphi) \\ -\sum_{i=0}^{N-1} t_i \sin^2(\omega t_i + \varphi) & \sum_{i=0}^{N-1} t_i^2 \sin^2(\omega t_i + \varphi) \end{pmatrix} .$$

(23.78)

As an example, a numerical simulation was performed for the (true) parameters $\omega = 2$, $\varphi = 1.1$, $f_s = 10$ and

$N = 256$. For this and subsequent examples in this section, frequencies have been non-dimensionalized using 2π. Thus, for $\omega = 2$, $f_s = 10$, $\omega = 2$, $f_s = 10$ corresponds to 10π samples per cycle. The noise power was varied logarithmically in 25 steps from e^{-10} to e^{+10}, which corresponds to 25 equal steps of SNR, expressed in dB. The noise power has been normalized with the signal variance, thus $\sigma_n^2 = 1$ corresponds to SNR $= 0$ dB. For each noise level, 1000 independent realizations were generated. The individual signals were processed by a least mean square estimation routine, which for Gaussian distributed noise is equal to the maximum-likelihood estimation. The Fisher information matrix and its inverse were calculated to be

$$\mathbf{J} = \frac{1}{\sigma_n^2} \begin{pmatrix} 130 & 1653 \\ 1653 & 28356 \end{pmatrix} , \tag{23.79}$$

$$\mathbf{J}^{-1} = \sigma_n^2 \begin{pmatrix} 0.029787 & -0.001737 \\ -0.001737 & 0.000137 \end{pmatrix} . \tag{23.80}$$

In Fig. 23.7a a sample signal with SNR $= 15$ dB is illustrated. The results presented in Fig. 23.7b show that the maximum-likelihood estimator meets the calculated CRLB. Furthermore, a threshold noise power $\sigma_{n,max}^2$ can be seen for the frequency estimate. Above this limit, the noise dominates the spectrum and the algorithm estimates the frequency randomly from the entire frequency range. The frequency of the threshold depends not only on the signal characteristics, but also on the capability of the estimation procedure to find the correct peak in the spectrum. The phase range is limited by $\pm\pi$. Therefore, the variance of the phase estimation is also limited.

In the case of phase Doppler signals, the Doppler frequency and the phase difference between two signals

$$x_i = \cos(\omega t_i + \varphi_x) + n_{x,i} , \quad i = 0, 1, 2 \ldots , (N-1) , \tag{23.81}$$

$$y_i = \cos(\omega t_i + \varphi_y) + n_{y,i} , \quad i = 0, 1, 2 \ldots , (N-1) , \tag{23.82}$$

with independent noise components n_x and n_y are of interest. To derive the CRLB for the phase difference, it is convenient to rewrite these signals as

$$x_i = m_{x,i} + n_{x,i} = \cos(\omega t_i + \varphi) + n_{x,i} , \\ i = 0, 1, 2 \ldots , (N-1) , \tag{23.83}$$

$$y_i = m_{y,i} + n_{y,i} = \cos(\omega t_i + \varphi + \Delta\varphi) + n_{y,i} \\ i = 0, 1, 2 \ldots , (N-1) . \tag{23.84}$$

Since the signals are of the same length with independent noise components, the joint probability density function $p(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{a})$ now becomes

$$p(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{a}) = \left(\frac{1}{2\pi\sigma_n^2}\right)^N \exp\left(-\frac{1}{2\sigma_n^2} \sum_{i=0}^{N-1} \left[(x_i - m_{x,i})^2 \right.\right. \\ \left.\left. + (y_i - m_{y,i})^2\right]\right) \tag{23.85}$$

with

$$m_{x,i} = \cos(\omega t_i + \varphi) , \tag{23.86}$$

$$m_{y,i} = \cos(\omega t_i + \varphi + \Delta\varphi) , \tag{23.87}$$
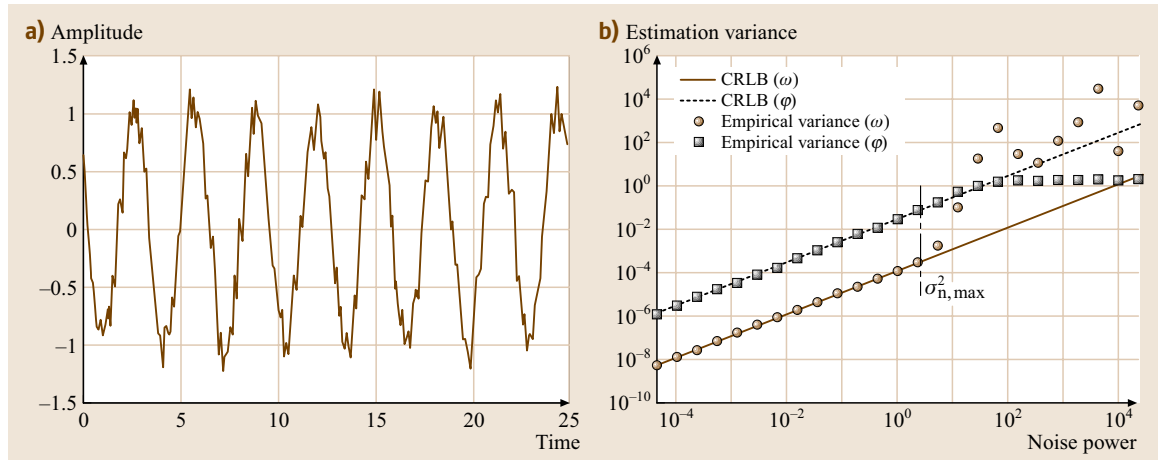

Fig. 23.7a,b Single-tone parameter estimation. (**a**) Sample input signal. (**b**) Comparison of the CRLB with the computed variance for frequency and phase estimates

and the elements of the Fisher information matrix become

$$J_{ij} = \frac{1}{\sigma_n^2} \sum_{k=0}^{N-1} \left(\frac{\partial m_{x,k}}{\partial a_i} \frac{\partial m_{x,k}}{\partial a_j} + \frac{\partial m_{y,k}}{\partial a_i} \frac{\partial m_{y,k}}{\partial a_j} \right).$$

(23.88)

The vector of unknown parameters is

$$a = \begin{pmatrix} \omega \\ \varphi \\ \Delta\varphi \end{pmatrix}.$$

(23.89)

Note that φ is included in the parameter vector since it is unknown, even though it is not used. The derivatives of m_x and m_y are

$$\frac{\partial m_{x,i}}{\partial \omega} = -t_i \sin(\omega t_i + \varphi),$$

(23.90)

$$\frac{\partial m_{x,i}}{\partial \varphi} = -\sin(\omega t_i + \varphi),$$

(23.91)

$$\frac{\partial m_{x,i}}{\partial \Delta\varphi} = 0,$$

(23.92)

$$\frac{\partial m_{y,i}}{\partial \omega} = -t_i \sin(\omega t_i + \varphi + \Delta\varphi),$$

(23.93)

$$\frac{\partial m_{y,i}}{\partial \varphi} = -\sin(\omega t_i + \varphi + \Delta\varphi),$$

(23.94)

$$\frac{\partial m_{y,i}}{\partial \Delta\varphi} = -\sin(\omega t_i + \varphi + \Delta\varphi).$$

(23.95)

The Fisher information matrix becomes

$$\mathbf{J} = \frac{1}{\sigma_n^2} \begin{pmatrix} P_2 + Q_2 & P_1 + Q_1 & Q_1 \\ P_1 + Q_1 & P_0 + Q_0 & Q_0 \\ Q_1 & Q_0 & Q_0 \end{pmatrix}$$

(23.96)

with

$$P_k = \sum_{i=0}^{N-1} t_i^k \sin^2(\omega t_i + \varphi),$$

(23.97)

$$Q_k = \sum_{i=0}^{N-1} t_i^k \sin^2(\omega t_i + \varphi + \Delta\varphi).$$

(23.98)

The inverse of the Fisher information matrix is

$$\mathbf{J}^{-1} = \sigma_n^2 \begin{pmatrix} P_2 + Q_2 & P_1 + Q_1 & Q_1 \\ P_1 + Q_1 & P_0 + Q_0 & Q_0 \\ Q_1 & Q_0 & Q_0 \end{pmatrix}^{-1}.$$

(23.99)

As an example, a numerical simulation was performed for the (true) parameters $\omega = 2$, $\varphi = 1.1$, $\Delta\varphi = -0.8$, $f_s = 10$ and $N = 256$. The noise power varied logarithmically in 25 steps from e^{-10} to e^{+10}. For each noise level 1000 independent realizations were generated. The individual signals (Fig. 23.8a) were processed by a maximum-likelihood estimation routine. The Fisher information matrix and its inverse were
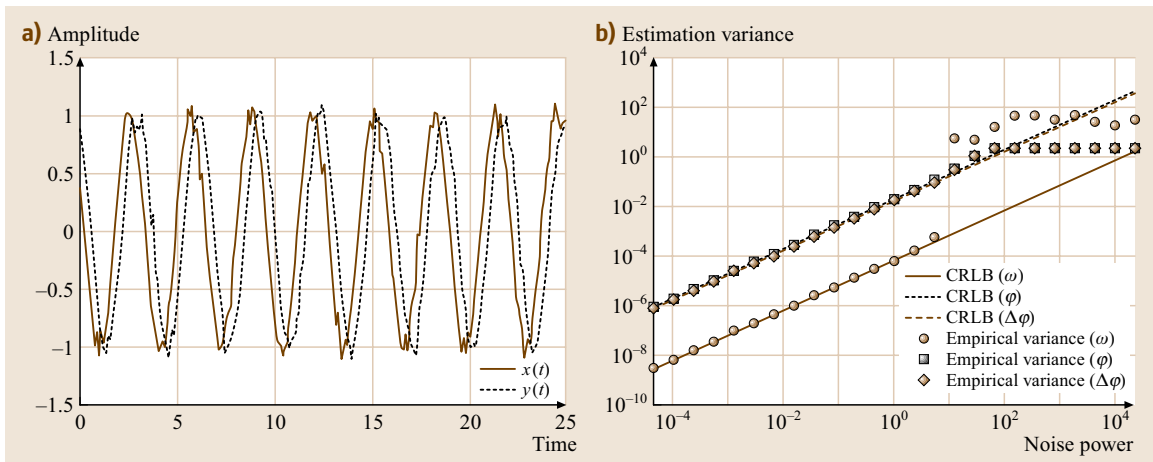

Fig. 23.8a,b Single-tone parameter estimation from two orthogonal signals. **(a)** Sample input signals. **(b)** Comparison of the CRLB with computed variance for frequency, phase and phase difference estimates

calculated to be

$$J = \frac{1}{\sigma_n^2} \begin{pmatrix} 55536 & 3261 & 1608 \\ 3261 & 258 & 128 \\ 1608 & 128 & 128 \end{pmatrix} , \quad (23.100)$$

$$J^{-1} = \sigma_n^2 \begin{pmatrix} 0.000070 & -0.000894 & 0.000008 \\ -0.000894 & 0.019063 & -0.007794 \\ 0.000008 & -0.007794 & 0.015536 \end{pmatrix} . \quad (23.101)$$

The results presented in Fig. 23.8b show that the maximum-likelihood estimator meets the calculated CRLB. Again, a threshold noise power can be seen for the frequency estimate and the phase range is limited by $\pm\pi$. Note that the CRLB and the empirically derived estimation variance of the phase φ and the phase difference $\Delta\varphi$ are different.

The expressions of the CRLB derived above are not convenient for practical use in setting up a signal processor. Explicit expressions of the CRLB are required. To derive these, a set of two orthogonal signals x and $\tilde{x}$ with independent noise components n and $\tilde{n}$ are considered.

$$x_i = m_i + n_i , \quad i = 0, 1, \dots, (N-1), \quad (23.102)$$
$$\tilde{x}_i = \tilde{m}_i + \tilde{n}_i , \quad i = 0, 1, \dots, (N-1), \quad (23.103)$$

with

$$m_i = A \cos(\omega t_i + \varphi) , \quad i = 0, 1, \dots, (N-1), \quad (23.104)$$
$$\tilde{m}_i = A \sin(\omega t_i + \varphi) , \quad i = 0, 1, \dots, (N-1), \quad (23.105)$$

where additionally the amplitude A is unknown. The joint probability density function becomes

$$p(x, \tilde{x}, a) = \left(\frac{1}{2\pi\sigma_n^2}\right)^N \exp\left\{-\frac{1}{2\sigma_n^2}\sum_{i=0}^{N-1}\left[(x_i - m_i)^2 + (\tilde{x}_i - \tilde{m}_i)^2\right]\right\} \quad (23.106)$$

and the elements of the Fisher information matrix are

$$J_{ij} = \frac{1}{\sigma_n^2}\sum_{k=0}^{N-1}\left(\frac{\partial m_k}{\partial a_i}\frac{\partial m_k}{\partial a_j} + \frac{\partial \tilde{m}_k}{\partial a_i}\frac{\partial \tilde{m}_k}{\partial a_j}\right) . \quad (23.107)$$

The vector of unknown parameters is

$$a = \begin{pmatrix} \omega \\ \varphi \\ A \end{pmatrix} . \quad (23.108)$$

Using the fact that $m_i^2 + \tilde{m}_i^2 = A^2$, the Fisher information matrix can be expressed explicitly as

$$J = \frac{1}{\sigma_n^2}\begin{pmatrix} A^2\sum_{i=0}^{N-1}t_i^2 & A^2\sum_{i=0}^{N-1}t_i & 0 \\ A^2\sum_{i=0}^{N-1}t_i & A^2\sum_{i=0}^{N-1}1 & 0 \\ 0 & 0 & N \end{pmatrix} = \frac{A^2}{6\sigma_n^2 f_s^2}$$

$$\times \begin{pmatrix} N(N-1)(2N-1) & 2N(N-1)f_s & 0 \\ 2N(N-1)f_s & 6Nf_s^2 & 0 \\ 0 & 0 & \frac{6Nf_s^2}{A^2} \end{pmatrix} . \quad (23.109)$$

The zero elements in this matrix indicate that the amplitude can be estimated completely independent of the frequency and the phase. Thus, the amplitude can be presumed to be known without changing the lower bounds of the frequency and phase estimator variance. The inverse of the Fisher information matrix becomes

$$J^{-1} = \frac{2\sigma_n^2}{A^2 N^2(N^2-1)}$$

$$\times \begin{pmatrix} 6Nf_s^2 & -3N(N-1)f_s & 0 \\ -3N(N-1)f_s & N(N-1)(2N-1) & 0 \\ 0 & 0 & \frac{A^2N(N^2-1)}{2} \end{pmatrix} \quad (23.110)$$

leading to the CRLB for the frequency

$$\sigma_\omega^2 \geq \frac{12\sigma_n^2 f_s^2}{A^2 N(N^2-1)} . \quad (23.111)$$

This is the CRLB for two signals with independent noise components. For only one signal, the information content is approximately one half, leading to

$$\sigma_\omega^2 \geq \frac{24\sigma_n^2 f_s^2}{A^2 N(N^2-1)} . \quad (23.112)$$

Using

$$SNR = \frac{A^2}{2\sigma_n^2} , \quad (23.113)$$

the CRLB for ω can be expressed as [23.17–19]

$$\sigma_\omega^2 \geq \frac{12 f_s^2}{N(N^2-1)SNR} , \quad (23.114)$$

or, if $\omega = 2\pi f$ is used, this variance reduces to

$$\sigma_f^2 \geq \frac{3 f_s^2}{\pi^2 N(N^2-1)SNR} . \quad (23.115)$$

This expression was derived assuming that the noise is spectrally white. Any filtering used to reduce the signal noise violates this assumption and (23.115) no longer strictly holds. Thus, while (band-pass) filtering may improve the SNR, the estimator variance may not be reduced.

For the case of phase Doppler signals, a second signal pair is required.

$$x_i = m_{x,i} + n_{x,i} \quad m_{x,i} = A\cos(\omega t_i + \varphi),$$
$$i = 0, 1, \ldots, (N-1), \tag{23.116}$$
$$\tilde{x}_i = \tilde{m}_{x,i} + \tilde{n}_{x,i} \quad \tilde{m}_{x,i} = A\sin(\omega t_i + \varphi),$$
$$i = 0, 1, \ldots, (N-1), \tag{23.117}$$
$$y_i = m_{y,i} + n_{y,i} \quad m_{y,i} = A\cos(\omega t_i + \varphi + \Delta\varphi),$$
$$i = 0, 1, \ldots, (N-1), \tag{23.118}$$
$$\tilde{y}_i = \tilde{m}_{y,i} + \tilde{n}_{y,i} \quad \tilde{m}_{y,i} = A\sin(\omega t_i + \varphi + \Delta\varphi),$$
$$i = 0, 1, \ldots, (N-1). \tag{23.119}$$

The joint probability density function becomes

$$p(x, \tilde{x}, y, \tilde{y}, a) = \left(\frac{1}{2\pi\sigma_n^2}\right)^{2N}$$
$$\exp\left(-\frac{1}{2\sigma_n^2}\sum_{i=0}^{N-1}\left[(x_i - m_{x,i})^2 + (\tilde{x}_i - \tilde{m}_{x,i})^2\right.\right.$$
$$\left.\left. + (y_i - m_{y,i})^2 + (\tilde{y}_i - \tilde{m}_{y,i})^2\right]\right) \tag{23.120}$$

and the elements of the Fisher information matrix are

$$J_{ij} = \frac{1}{\sigma_n^2}\sum_{k=0}^{N-1}\left(\frac{\partial m_{x,k}}{\partial a_i}\frac{\partial m_{x,k}}{\partial a_j} + \frac{\partial \tilde{m}_{x,k}}{\partial a_i}\frac{\partial \tilde{m}_{x,k}}{\partial a_j}\right.$$
$$\left. + \frac{\partial m_{y,k}}{\partial a_i}\frac{\partial m_{y,k}}{\partial a_j} + \frac{\partial \tilde{m}_{y,k}}{\partial a_i}\frac{\partial \tilde{m}_{y,k}}{\partial a_j}\right). \tag{23.121}$$

The vector of unknown parameters is

$$a = \begin{pmatrix} \omega \\ \varphi \\ \Delta\varphi \\ A \end{pmatrix}. \tag{23.122}$$

The Fisher information matrix becomes

$$J = \frac{1}{\sigma_n^2}\begin{pmatrix} 2A^2\sum_{i=0}^{N-1}t_i^2 & 2A^2\sum_{i=0}^{N-1}t_i & A^2\sum_{i=0}^{N-1}t_i & 0 \\ 2A^2\sum_{i=0}^{N-1}t_i & 2A^2\sum_{i=0}^{N-1}1 & A^2\sum_{i=0}^{N-1}1 & 0 \\ A^2\sum_{i=0}^{N-1}t_i & A^2\sum_{i=0}^{N-1}1 & A^2\sum_{i=0}^{N-1}1 & 0 \\ 0 & 0 & 0 & 2N \end{pmatrix}, \tag{23.123}$$

$$J = \frac{A^2}{6\sigma_n^2 f_s^2}$$
$$\times \begin{pmatrix} \begin{matrix} 2N(N-1) \\ \times(2N-1) \end{matrix} & \begin{matrix} 6N \\ \times(N-1)f_s \end{matrix} & \begin{matrix} 3N \\ \times(N-1)f_s \end{matrix} & 0 \\ 6N(N-1)f_s & 12Nf_s^2 & 6Nf_s^2 & 0 \\ 3N(N-1)f_s & 6Nf_s^2 & 6Nf_s^2 & 0 \\ 0 & 0 & 0 & \frac{12Nf_s^2}{A^2} \end{pmatrix}. \tag{23.124}$$

The inverse of the Fisher information matrix is

$$J^{-1} = \frac{\sigma_n^2}{2A^2 N}\begin{pmatrix} \frac{12f_s^2}{N^2-1} & -\frac{6f_s}{N+1} & 0 & 0 \\ -\frac{6f}{N+1} & \frac{5N-1}{N+1} & -2 & 0 \\ 0 & -2 & 4 & 0 \\ 0 & 0 & 0 & A^2 \end{pmatrix} \tag{23.125}$$

leading to the following CRLBs [23.20]

$$\sigma_\omega^2 \geq \frac{6\sigma_n^2 f_s^2}{A^2 N(N^2-1)}, \tag{23.126}$$
$$\sigma_\varphi^2 \geq \frac{\sigma_n^2(5N-1)}{2A^2 N(N+1)}, \tag{23.127}$$
$$\sigma_{\Delta\varphi}^2 \geq \frac{2\sigma_n^2}{A^2 N}. \tag{23.128}$$

Note that these are the lower bounds for the four signals with independent noise components. If only the two phase Doppler signals are given, then the information content is one half and the bounds become

$$\sigma_\omega^2 \geq \frac{12\sigma_n^2 f_s^2}{A^2 N(N^2-1)} = \frac{6f_s^2}{N(N^2-1)\mathrm{SNR}}, \tag{23.129}$$
$$\sigma_\varphi^2 \geq \frac{\sigma_n^2(5N-1)}{A^2 N(N+1)} = \frac{5N-1}{2N(N+1)\mathrm{SNR}}, \tag{23.130}$$
$$\sigma_{\Delta\varphi}^2 \geq \frac{4\sigma_n^2}{A^2 N} = \frac{2}{N\mathrm{SNR}}. \tag{23.131}$$

The CRLB is exactly half of the value for a Doppler signal, since the frequency information content in the
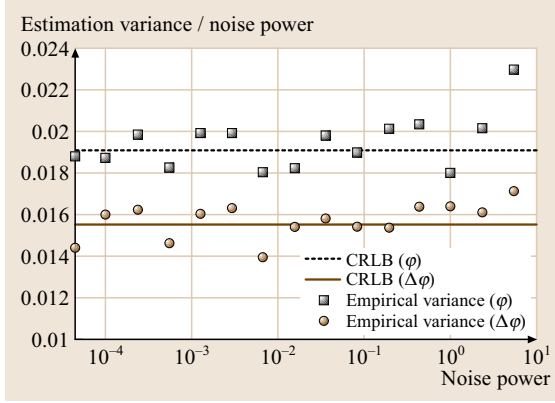
Fig. 23.9 Variance of the maximum-likelihood estimator, normalized by the noise power for phase and phase difference. Comparison to the respective CRLBs (Simulation parameters as used in Fig. 23.8)

two phase Doppler signals is twice as large. The CRLB for the phase difference between the two phase Doppler signals is lower than that for the absolute phase by a factor of 0.8, for large N. This can be seen in Fig. 23.9, which presents the estimator variance results from Fig. 23.8b, normalized by the noise power.

All of the above derivations were based on a signal model of constant amplitude. In fact, laser Doppler and phase Doppler signals exhibit an amplitude described by a Gaussian envelope and for this case, the derivation of CRLB is somewhat more tedious. Results for the frequency, amplitude, arrival time and residence time are presented by *Høst-Madsen* and *Gjelstrup* [23.21].

23.5.2 Particle Imaging

The task of deriving the two-dimensional position from images of small particles is addressed. For simplification, a sampled two-dimensional Gaussian function

$$z(x, y) = A \exp\left\{-\eta\left[(x-\xi)^2 + (y-\psi)^2\right]\right\}$$
(23.132)

with the particle position $[\xi, \psi]$, the maximum intensity amplitude A and the parameter η, defining the width of the particle image, is used for the intensity profile of the particle image. The originally continuous intensity profile is given at discrete positions (pixels) z_{ij}. For simplicity, an ideas sampling is assumed with $z_{ij} = z(x_i, y_j)$.

To derive the particle position from its sampled image, several methods can be applied such as the centroid method or a Gaussian fit. For the given condition

with a sampled Gaussian intensity profile, the centroid method is biased while the Gaussian fit is bias-free. However, all methods are affected by noise.

Assuming a Poisson-distributed fluctuation of the intensity values due to the photon noise as in *Wernet* and *Pline* [23.22], the elements of the Fisher information matrix become

$$J_{ij} = \sum_k \sum_l \frac{1}{z_{kl}} \frac{\partial z_{kl}}{\partial a_i} \frac{\partial z_{kl}}{\partial a_j} .$$
(23.133)

Note that the limits of the sums are not given here. A summation over all pixels that contribute to the particle image is assumed. Since the particle images are usually much smaller then the dimensions of the observed imaging area, the limits of the sums can be dropped off as long as the particle images do not overlap. The parameter vector is given by

$$\boldsymbol{a} = \begin{pmatrix} A \\ \eta \\ \xi \\ \psi \end{pmatrix} .$$
(23.134)

The element $J_{\xi\xi}$ of the Fisher information matrix then becomes

$$
\begin{aligned}
J_{\xi\xi} &= \sum_k \sum_l \frac{1}{z_{kl}} \left(\frac{\partial z_{kl}}{\partial \xi}\right)^2 \\
&= \sum_k \sum_l z_{kl}[-2\eta(x_k - \xi)]^2 \\
&= \sum_k \sum_l 4A\eta^2(x_k - \xi)^2 \exp\{-\eta[(x_k - \xi)^2 \\
&\quad + (y_l - \psi)^2]\} .
\end{aligned}
$$
(23.135)

A separation in two sums yields

$$
J_{\xi\xi} = 4A\eta^2 \left\{ \sum_k (x_k - \xi)^2 \exp\left[-\eta(x_k - \xi)^2\right] \right. \\
\left. \left| \sum_l \exp\left[-\eta(y_l - \psi)^2\right] \right\} .
$$
(23.136)

Approximating the sums by integrals yields

$$
J_{\xi\xi} = \frac{4A\eta^2}{\Delta x \Delta y} \left\{ \int (x - \xi)^2 \exp\left[-\eta(x - \xi)^2\right] dx \right. \\
\left. \left| \int \exp\left[-\eta(y - \psi)^2\right] dy \right\}
$$
(23.137)

with the sampling intervals Δx and Δy, which are one pixel ($\Delta x = \Delta y \equiv 1$). Assuming furthermore an unbounded particle image and therefore integrals within

the range $[-\infty, \infty]$, the first integral becomes $\sqrt{\pi/(4\eta^3)}$ and the second $\sqrt{\pi/\eta}$, yielding

$$J_{\xi\xi} = 2\pi A . \tag{23.138}$$

Assuming that the parameters of the parameter vector $\boldsymbol{a}$ can be estimated independently, the appropriate element of the inverse Fisher information matrix can be approximated by

$$\left(J^{-1}\right)_{\xi\xi} \approx \left(J_{\xi\xi}\right)^{-1} = \frac{1}{2\pi A} \tag{23.139}$$

and finally the CRLB for estimating the component ξ of the particle position in the x-direction becomes

$$\sigma_\xi^2 \geq \frac{1}{2\pi A} . \tag{23.140}$$

Similarly, for estimating the component ψ of the particle position in the y-direction is

$$\sigma_\psi^2 \geq \frac{1}{2\pi A} . \tag{23.141}$$

In Fig. 23.10 the estimation variances are shown for the centroid method, the Gaussian fit and an alternative maximum-likelihood estimator (MLE), obtained empirically based on a computer simulation. The deviations of the two estimated components $\hat{\xi}$ and $\hat{\psi}$ from the correct values ξ and ψ are combined to a common deviation

$$\sqrt{(\hat{\xi}-\xi)^2 + (\hat{\psi}-\psi)^2} \tag{23.142}$$

as well as the CRLB

$$\sigma_\xi^2 + \sigma_\psi^2 \geq \frac{1}{\pi A} . \tag{23.143}$$
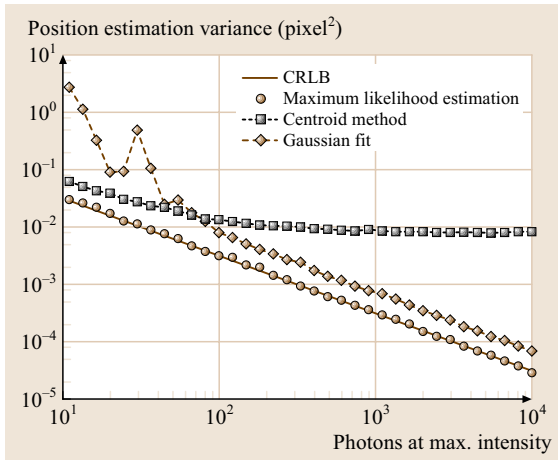

Fig. 23.10 CRLB for the position estimation of particle images for Poisson-distributed noise

As expected the MLE obviously meets the CRLB, while the Gaussian fit has a significantly larger estimation variance. This is a result of the different image information used for the two estimators. While the MLE uses the entire particle image for its position estimation, the position estimate of the Gaussian fit is based only on the nine central pixels of the particle image. However, extending the number of pixels to be used for the fit is not useful. On the one hand this would reduce the estimation variance. But on the other hand, systematic errors due to overlaps of particles images would increase.

Compared to those both estimators, the centroid method is dominated by systematic errors for large numbers of photons (low noise). However, for small number of photons (high noise) the centroid method is more robust then the Gaussian fit.

Due to the common use of powerful pulsed lasers the accuracy of the position estimation is nowadays not limited by the number of photons but by thermal noise and the discrete gray values of the images. Both can be approximated by Gaussian-distributed noise [23.23]. For a Gaussian-distributed noise with the variance σ_n^2 the elements of the Fisher information matrix become

$$J_{ij} = \frac{1}{\sigma_n^2} \sum_k \sum_l \frac{\partial z_{kl}}{\partial a_i} \frac{\partial z_{kl}}{\partial a_j} . \tag{23.144}$$

The element $J_{\xi\xi}$ of the Fisher information matrix then becomes

$$
\begin{aligned}
J_{\xi\xi} &= \frac{1}{\sigma_n^2} \sum_k \sum_l \left(\frac{\partial z_{kl}}{\partial \xi}\right)^2 \\
&= \frac{1}{\sigma_n^2} \sum_k \sum_l z_{kl}^2 [-2\eta(x_k-\xi)]^2 \\
&= \frac{1}{\sigma_n^2} \sum_k \sum_l 4A^2\eta^2(x_k-\xi)^2 \exp\{-2\eta[(x_k-\xi)^2 \\
&\quad + (y_l-\psi)^2]\} .
\end{aligned}
\tag{23.145}
$$

A separation in two sums yields

$$
\begin{aligned}
J_{\xi\xi} = \frac{4A^2\eta^2}{\sigma_n^2} &\left\{ \sum_k (x_k-\xi)^2 \exp\left[-2\eta(x_k-\xi)^2\right] \right. \\
&\left. \times \left|\sum_l \exp\left[-2\eta(y_l-\psi)^2\right]\right| \right\} .
\end{aligned}
\tag{23.146}
$$

Approximating the sums by integrals yields

$$
\begin{aligned}
J_{\xi\xi} = \frac{4A^2\eta^2}{\sigma_n^2 \Delta x \Delta y} &\left\{ \int (x-\xi)^2 \exp\left[-2\eta(x-\xi)^2\right] \mathrm{d}x \right. \\
&\left. \times \left|\int \exp\left[-2\eta(y-\psi)^2\right] \mathrm{d}y\right| \right\}
\end{aligned}
\tag{23.147}
$$

with the sampling intervals Δx and Δy, which again are 1 pixel ($\Delta x = \Delta y \equiv 1$). Assuming an unbounded particle image and therefore integrals within the range $[-\infty, \infty]$, the first integral becomes $\sqrt{\pi/(32\eta^3)}$ and the second $\sqrt{\pi/(2\eta)}$, yielding

$$J_{\xi\xi} = \frac{\pi A^2}{2\sigma_n^2} \qquad (23.148)$$

Assuming that the parameters of the parameter vector $\boldsymbol{a}$ can be estimated independently, the appropriate element of the inverse Fisher information matrix can be approximated by

$$\left(J^{-1}\right)_{\xi\xi} \approx \left(J_{\xi\xi}\right)^{-1} = \frac{2\sigma_n^2}{\pi A^2} \qquad (23.149)$$

and finally, the CRLB for estimating the component ξ of the particle position in the x-direction becomes

$$\sigma_\xi^2 \geq \frac{2\sigma_n^2}{\pi A^2} \, . \qquad (23.150)$$

Similarly, for estimating the component ψ of the particle position in the y-direction is

$$\sigma_\psi^2 \geq \frac{2\sigma_n^2}{\pi A^2} \, . \qquad (23.151)$$

Position estimation variance (pixel^2)

— CRLB
○ Maximum likelihood estimation
--■-- Centroid method
--◇-- Gaussian fit

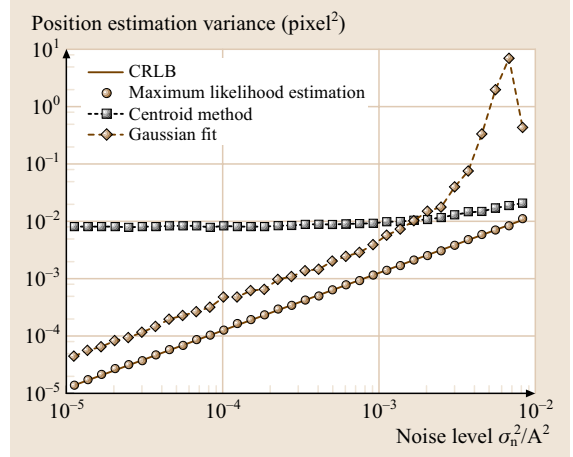Noise level σ_n^2/A^2

Fig. 23.11 CRLB for the position estimation of particle images for Gaussian-distributed noise

In Fig. 23.11 the results of the computer simulation are shown for Gaussian-distributed noise. The MLE again meets the CRLB, while the Gaussian fit has a larger estimation variance due to the reduced information used for the estimation procedure. The centroid method again is limited by systematic errors for low noise levels, but is more robust than the Gaussian fit for high noise levels.

23.6 Propagation of Errors

The concepts of stochastic and systematic errors for a given measurement quantity have already been introduced in (23.26) and (23.27), respectively. If a derived quantity y depends on several individual measurement quantities x_i, the question arises as to the measurement error in y

$$y = f(x_1, x_2, \ldots, x_n) \, . \qquad (23.152)$$

The propagation of errors from the quantities x_i to y is treated separately for systematic and stochastic errors. The resulting systematic error in y is found by using a first-order Taylor expansion

$$\delta y = \frac{\partial f}{\partial x_1}\delta x_1 + \frac{\partial f}{\partial x_2}\delta x_2 + \ldots + \frac{\partial f}{\partial x_n}\delta x_n \, , \qquad (23.153)$$

where δx_i are the systematic errors for each measurement quantity x_i and δy is the overall systematic error.

Note that all δx_i quantities are signed and, as such, systematic errors may be compensating in nature.

Stochastic errors are treated in the mean square, leading to the relation

$$\sigma_y = \sqrt{\left(\frac{\partial f}{\partial x_1}\sigma_{x_1}\right)^2 + \left(\frac{\partial f}{\partial x_2}\sigma_{x_2}\right)^2 + \ldots + \left(\frac{\partial f}{\partial x_n}\sigma_{x_n}\right)^2} \qquad (23.154)$$

where the individual estimator variances $\sigma_{x_i}^2$ have been evaluated using techniques described in the previous section. This formula assumes that all of the individual stochastic errors are normally distributed and that the standard deviations are all evaluated with the same confidence intervals. An extensive discussion of error propagation can be found in *Kline* and *McClintock* [23.3], *Kline* [23.4] and *Moffat* [23.5].

References

23.1 J.S. Bendat, A.G. Piersol: *Random Data: Analysis and Measurement Procedures* (Wiley, New York 1986)

23.2 W.K. George Jr.: Processing of random signals, Proc. Dyn. Flow Conf. (1978) 20–63

23.3 S.J. Kline, F.A. McClintock: Describing uncertainties in single-sample experiments, Mech Eng **75**, 3–8 (1953)

23.4 S.J. Kline: The purposes of uncertainty analysis, ASME J. Fluid. Eng. **107**, 153–160 (1985)

23.5 R.J. Moffat: Using uncertainty analysis in the planning of an experiment, ASME J. Fluid. Eng. **107**, 173–182 (1985)

23.6 R.J. Moffat: Describing uncertainties in experimental results, Exp. Thermal Fluid Sci. **1**, 3–17 (1988)

23.7 A. Stuart, J.K. Ord: *Kendall's Advanced Theory of Statistics* (Edward Arnold, London 1994)

23.8 M. Kendall, A. Stuart: *The Advanced Theory of Statistics*, Vol. Charles Griffen (Charles Griffen, London 1963)

23.9 L.H. Benedict, R.D. Gould: Towards better uncertainty estimates for turbulence statistics, Exp. Fluids **22**, 129–136 (1996)

23.10 J.W. Tukey: Bias and confidence in not-quite large samples, Ann. Math. Statist. **29**, 614 (1958)

23.11 B. Efron, R.J. Tibshirani: *An Introduction to the Bootstrap* (Chapman Hall, New York 1993)

23.12 C. Tropea: Performance testing of LDA/PDA signal processing systems, Int. Conf. on Laser Anemom – Adv. and Appl., Swansea, UK: paper II 4 (1989)

23.13 M. Kendall, A. Stuart: *The Advanced Theory of Statistics*, Vol. 2 (Charles Griffen, London 1963)

23.14 A. Papoulis: *Probability, Random Variables and Stochastic Processes* (McGraw-Hill, New York 1988)

23.15 S.M. Kay: *Fundamentals of Statistical Signal Processing: Estimation theory* (Prentice Hall, Englewood Cliffs 1993)

23.16 A.D. Whalen: *Detection of Signals in Noise* (Academic, New York 1971)

23.17 K. M. Ibrahim, D. Werthimer, W. D. Bachalo: Signal Processing Considerations for Laser Doppler and Phase Doppler Applications. Appl. of Laser Techn. Fluid. Mech. 1990 291–316

23.18 D.C. Rife, R.R. Boortstyn: Single tone parameter estimation from discrete time observations, IEEE Trans. on Inform. Theory **20**, 591–596 (1974)

23.19 T. Wriedt, K.A. Bauckhage, A. Schöne: Application of Fourier analysis to phase Doppler-signals generated by rough metal particles, IEEE Trans. Instrum. Meas. **38**, 984–990 (1989)

23.20 A. Høst-Madsen, K. Andersen: Lower bounds for estimation of frequency and phase of Doppler signals, Meas. Sci. **6**, 637–652 (1995)

23.21 A. Høst-Madsen, P. Gjelstrup: New processing methods for LDA signals, Proc. 8th Int. Symp. on Appl. of Laser Techn. to Fluid Mech., Lisbon, Portugal: paper 31.3 (1996)

23.22 M.P. Wernet, A. Pline: Particle displacement tracking technique and Cramer-Rao lower bound error in centroid estimates from CCD imagery, Exp. Fluids **15**, 295–307 (1993)

23.23 J. Westerweel: Theoretical analysis of the measurement precision in particle image velocimetry, Exp. Fluids **29**, S3–S12 (2000)