
Principles of Matched Filtering with Auditory Examples from Selected Vertebrates

5

Peter M. Narins and Grace A. Clark

Contents

5.1	Part I Principles of Matched Filtering	112
5.1.1	The Correlation Detector	114
5.1.2	Section Organization	114
5.2	An Example Detection Problem	114
5.2.1	Define the Signal-to-Noise Ratio (SNR)	115
5.3	The Matched Filter Detector	116
5.3.1	Convolution and Filtering	117
5.3.2	Correlation vs. Convolution	117
5.3.3	The Matched Filter	118
5.3.4	Example: Applying the Matched Filter to Our Example Signals	118
5.4	Bayesian Binary Hypothesis Testing and Performance Measurement	119
5.4.1	The Confusion Matrix (or Contingency Table)	120
5.4.2	Bayesian Hypothesis Testing for Multidimensional (Vector) Data	122
5.4.3	Training and Testing Phases of the Detection Process	123
5.4.4	The Receiver Operating Characteristic (ROC) Curve	124
5.4.5	Statistical Confidence Interval About the Probability of Correct Classification	125
5.5	Processing for the Example Problem	126
5.5.1	Experiment Design for the Example Problem	126
5.5.2	Processing Results for the Example Problem	127
5.5.3	Conclusions	130

P.M. Narins (✉)

Department of Integrative Biology and Physiology, University of California, Los Angeles, 621
Charles E. Young Drive South, Los Angeles, CA 90095, USA

Department of Ecology and Evolutionary Biology, University of California, Los Angeles, 621
Charles E. Young Drive South, Los Angeles, CA 90095, USA

e-mail: pnarins@ucla.edu

G.A. Clark, Ph.D., IEEE Fellow

Grace Clark Signal Sciences, 532 Alden Lane, Livermore, CA 94550, USA

e-mail: GCSS_Grace@comcast.net

5.6	Part II Auditory Matched Filtering: Biological Examples from Selected Vertebrates .	131
5.6.1	Weakly Electric Fish	132
5.6.2	Anuran Amphibians	132
5.6.3	Birds	137
5.6.4	Conclusions	138
	References	139

Abstract

This chapter consists of two parts. In the first part, a compact tutorial exposition of the principles of matched filtering for the biological scientist is introduced. The concept of matched filtering for detecting desired signals buried in noisy measurement signals is presented. It is shown that the matched filter is another name for the correlation detector or replica-correlation detector, which exploits prior knowledge in the form of an exemplar (or replica) of the desired signal. An example detection problem is used to demonstrate the matched filtering approach. The detection methodology comes from hypothesis testing algorithms in Bayesian detection theory. This Bayesian approach provides very powerful methods for evaluating detection performance in the form of the Receiver Operating Characteristic (ROC) curve and the statistical confidence interval about the probability of correct classification. It is shown that the matched filter can be an effective detection tool when exemplars of the desired signal are available a priori. In the second part, several key examples of matched filters in the auditory systems of several selected vertebrates are provided.

5.1 Part I Principles of Matched Filtering

This section presents a compact tutorial exposition of the principles of matched filtering for the biological scientist. The matched filter is a concept from communications and radar theory (Van Trees 1968; Whalen 1971; Kay 1998; Papoulis 1965) that has been applied widely to various other applications in science and engineering (Carpranica and Moffat 1983; Wehner 1987; Clark et al. 1999, 2000, 2009; Waltz and Llinas 1990; Clark 1999; Jazwinski 1970; Candy 2006). It is a statistical signal processing algorithm designed for detecting the existence of a desired signal that is buried in a noisy measurement signal. In general and for our purposes, a “signal” can be interpreted to be a scalar or multidimensional construct; e.g., a time series, an image, a three-dimensional volume, a video sequence, etc. The fundamental mathematical approaches are common to all of these modalities. For tutorial purposes, we focus here on the fundamental matched filtering approach for a time series. The literature in matched filtering is vast, and a full understanding of the concept requires a great deal of study. This section endeavors to encapsulate the most important principles of matched filtering so as to aid the biologist in processing experimental data. The concepts are introduced with the idea that the reader can consult the referenced literature for in-depth treatments.

Imagine that we are conducting a scientific experiment involving a physical process that generates a noisy discrete-time temporal signal (time series). Our goal is to make a judgment or decision about whether or not the noisy measured signal contains a particular desired signal component of interest to us. We can say that we wish to *detect* the desired signal component. More specifically, imagine that we have a real noisy measured discrete-time signal

$$x(n) = a(n - n_0) + v(n) \quad (\text{Noisy Measurement Signal}) \quad (5.1)$$

where $a(n - n_0)$ is a time delayed version of the desired signal $a(n)$ that we wish to detect, $v(n)$ is noise (undesired component of the signal), n denotes the discrete time index, $n = 0, 1, 2, \dots, N - 1$, n_0 is the time delay, and the sampling interval T is normalized to one (i.e., $T = 1$ s, so $nT = n$). We assume that the arrival time n_0 of the desired signal $a(n - n_0)$ is unknown to the user.

Perhaps the simplest way to *detect* the signal $a(n)$ in the noisy measurement $x(n)$ is to choose various values of a threshold γ and compare the amplitude values of $x(n)$ to that threshold at each value of time n . This is often called a “threshold detector” that uses the raw signal $x(n)$ as the *decision statistic* (the quantity we compare to the threshold). At each time instant n , if $x(n)$ is less than the threshold γ , then we declare that $a(n)$ is not present in the measurement at time n . If $x(n)$ is greater than or equal to the threshold γ , then we declare that $a(n)$ is present in the measurement at time n . Of course, our declarations will vary as we vary the threshold value. This section discusses methods for dealing with the various declarations and measuring the performance of the detection algorithm.

Note that the terms “detection” and “classification” are often used interchangeably, for good reason. Detection theory is often regarded as a subset of classification theory. Classification generally describes methods for multidimensional hypothesis testing, but detection theory was originally developed for scalar (one dimensional) hypothesis testing. The distinction is not nearly as important as understanding their common underlying concepts.

Imagine now a scenario in which we have prior information about the desired signal $a(n)$. In the signal processing world, we always welcome prior knowledge because we are often able to incorporate it into our processing algorithms to give us an advantage. In many applications (especially in communication and radar systems), we have available exemplars of the signal $a(n)$ we wish to detect. In radar, for example, we have a pulse generator and antenna system that create waves in the form of a transient pulse which are propagated through a channel, reflect from a physical target, propagate back through the channel, and are measured by the antenna and radar system. In this scenario, we have prior knowledge of the transmitted transient pulse used to interrogate the target, because we generated it ourselves. We can, of course, generate additional exemplars of the transmitted pulse. The issue now is whether or not we can use this prior knowledge to help us detect such a transient pulse in a measured waveform. The idea behind the *correlation detector* is that we can.

5.1.1 The Correlation Detector

The term “correlation detector” refers to a signal detection algorithm that cross-correlates the measured data with a replica or exemplar of the desired signal. It is also sometimes called a “correlator” or a “replica correlator.” The basic idea behind the correlation detector is to use the mathematical operation called *cross-correlation* to scan the measured signal $x(n)$ with an exemplar of $a(n)$. The cross-correlation result is $R_{ax}(k)$, a time waveform that is a function of the time delay k between $a(n)$ and $x(n)$ during the correlation process. The premise of the correlation detector is that the cross-correlation waveform will be large when the measurement $x(n)$ contains nonzero $a(n)$ and small when it does not. The correlation detector uses the cross-correlation waveform $R_{ax}(k)$ (or a function of $R_{ax}(k)$) as the *decision statistic* in a threshold detection algorithm. The hope is that using $R_{ax}(k)$ as the decision statistic will give better detection performance than that which would be obtained by using the raw signal $x(n)$ as the decision statistic. This hoped-for result is generally realized in practice. Prior knowledge is very helpful.

5.1.2 Section Organization

In the remainder of this section, we introduce the matched filter by examining an example detection problem and showing the steps in the detection process. We show that the *matched filter* is really another name for the correlation detector and why. We then step through the process of detecting a signal buried in a noisy measurement and develop the appropriate measures for evaluating detector performance. We show that the matched filter can be an effective detection tool when exemplars of the desired signal are available a priori.

5.2 An Example Detection Problem

The data for an example event signal detection problem are depicted in Fig. 5.1. The *top signal* is a transient “event waveform” $a(n)$ representing a physical event that we wish to detect. In this particular case, this event is a dissolver acid time series from a chemical processing plant. The *middle signal* is a delayed version of the transient waveform with delay equal to n_0 . The *bottom signal* denotes the noisy measurement signal $x(n) = a(n - n_0) + v(n)$. The noise $v(n)$ is statistically white, zero mean and Gaussian distributed with noise variance $\sigma_v^2 = .599$. We denote this as $v(n) \sim N(0, \sigma_v^2)$. Clearly, the noise power is fairly large, so the desired event signal $a(n)$ is significantly obscured in the measurement. This leads us to define the concept of *signal-to-noise ratio*.

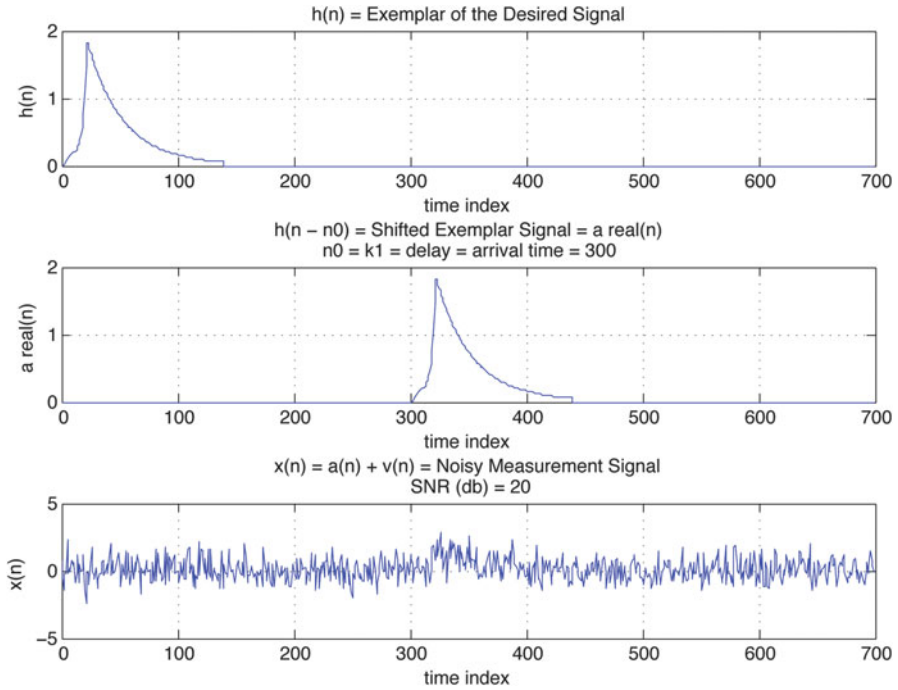


Fig. 5.1 Simulated measurement signal $x(n)$ created for our matched filtering example: (Top) Exemplar of the transient event waveform $a(n)$ we wish to detect in the noise measurement. (Middle) Time delayed version $a(n - n_0)$ of the transient waveform. (Bottom) Measured event plus noise signal $x(n) = a(n - n_0) + v(n)$. The signal-to-noise ratio is $\text{SNR} = 20$ dB. The magnitude units are arbitrary and the temporal sampling period $T = 1$

5.2.1 Define the Signal-to-Noise Ratio (SNR)

Consider a noisy measured signal as described in Eq. 5.1. In a nice theoretical simulation experiment, we can easily know the desired signal $a(n)$ and the noise $v(n)$ separately, because we create them ourselves. However, in many real-world experiments, we can measure only the sum in Eq. 5.1. In some experiments, the measured signal consists of pre-event noise (before a signal event occurs) followed in time by the sum of the event signal $a(n)$ and the noise $v(n)$. This occurs in, for example, seismic event signals. In this case, we can cut out a section of pre-event noise and compute its noise variance. In some rare applications, we have available a noiseless signal $a(n)$ before the physical system corrupts it (e.g., in communications and radar).

Let us assume for now that we have prior knowledge that allows us to separate the desired signal from the noise. In general, we define the signal-to-noise ratio (SNR) as follows (Kay 1998; Candy 2006):

$$\text{SNR} \triangleq \frac{\text{Signal energy}}{\text{Noise variance}} \quad (\text{Signal-to-noise ratio}) \quad (5.2)$$

$$\triangleq \frac{E_a}{\sigma_v^2} \quad (5.3)$$

where the energy in signal $a(n)$ is given by

$$E_a \triangleq \sum_{n=n_0}^{n_1} a^2(n) \quad (5.4)$$

and we calculate the energy in signal $a(n)$ over the time interval between appropriate time indices n_0 and n_1 . We denote the noise variance by σ_v^2 . We can express the SNR in the commonly used units of decibels (dB) by applying the following definition:

$$\text{SNR(dB)} \triangleq 10 \log_{10} \left[\frac{E_a}{\sigma_v^2} \right] \quad (5.5)$$

$$= 10 \log_{10}[R], \quad \text{where } R \triangleq E_a / \sigma_v^2 \quad (5.6)$$

Note that in a simulation experiment, once we have computed E_a for our particular signal and we know our desired SNR(dB), we can solve for the noise variance required to achieve that SNR(dB). If we define Q as follows, then we have:

$$Q \triangleq \text{SNR(dB)} / 10 \quad (5.7)$$

$$R = 10^Q \quad (5.8)$$

$$\sigma_v^2 = E_a / R \quad (5.9)$$

Consider a numerical example: Let $E_a = 4.2632$ and the desired SNR(dB) = 40. Then, we see that $Q = 4$, $R = 10^4$, and $\sigma_v^2 = 4.2532e - 4$. Note that in applications such as the seismic event signal described earlier, we can define an approximate SNR, call it *SNRE*, that consists of the ratio of the energy in the measured signal $x(n)$ and the noise variance (Clark and Rodgers 1981). This is one way in which we can cope with the lack of prior knowledge.

5.3 The Matched Filter Detector

The term *matched filter* is another name for the correlation detector. The fundamental principle of matched filtering is to exploit the prior knowledge we have about the desired signal of interest $a(n)$ to build a correlation detector. One might ask the question, "Then why do we call the correlation detector a matched filter?" The answer lies in the meaning of the mathematical operation of correlation, as we show next.

5.3.1 Convolution and Filtering

Given two discrete-time signals $a(n)$ of length M samples and $x(n)$ of length N samples, we can define the convolution $y(n)$ of the two signals as follows:

$$\begin{aligned} y(n) &= a(n) * x(n) = \sum_{k=-\infty}^{\infty} a(k)x(n-k) \\ &= x(n) * a(n) = \sum_{k=-\infty}^{\infty} x(k)a(n-k) \quad (\text{Convolution}) \quad (5.10) \end{aligned}$$

We see that convolution is commutative. The convolution operation can be interpreted as “flipping” (reversing) one of the two signals in time, then sliding it in time across the other signal and multiplying each of the values of the two signals together at each time sample and summing the products (Oppenheim and Schaffer 1975). The time-reversal operation is described mathematically by $x(-n)$ and $a(-n)$. The key concept is that a linear filtering operation in the time domain can be written as a convolution summation (Oppenheim and Schaffer 1975). Thus, if we filter a signal $x(n)$ with a linear filter impulse response $a(n)$, then that filtering operation is written as a convolution of the form in Eq. 5.11. Note that the convolution of a signal of length N with a signal of length M has length $N_c = N + M - 1$.

5.3.2 Correlation vs. Convolution

The correlation of two real discrete-time signals $a(n)$ of length M samples and $x(n)$ of length N samples is written as follows:

$$\begin{aligned} R_{ax}(n) &= a(n) * x(-n) = \sum_{k=-\infty}^{\infty} a(k)x(n+k) \\ &= x(n) * a(-n) = \sum_{k=-\infty}^{\infty} x(k)a(n+k) = R_{xa}(n) \quad (\text{Correlation}) \quad (5.11) \end{aligned}$$

We see from this equation that correlation is commutative, and the correlation operation can be written in terms of the convolution operation. If we do the convolution operation *without* reversing one of the two signals in time, then we get the correlation operation. We recall that the convolution operation reverses one of the signals in time before sliding it across the other signal. If we reverse one of the signals *before* doing the convolution, then the convolution reverses it again, so the result is an operation with a signal that has been reversed twice. This is equivalent to the *correlation* operation.

5.3.3 The Matched Filter

We are ready now to see why we call the correlation detector a *matched filter*. If we interpret the exemplar signal $a(n)$ to be the impulse response of a linear filter, then if we convolve $a(n)$ with $x(n)$, we are *filtering* the measurement $x(n)$ with a *filter matched to the signal of interest $a(n)$* . We have chosen a filter impulse response that is matched to the desired signal of interest $a(n)$. Because correlation is equivalent to convolution (with one of the signals flipped in time), we can think of the correlation operation as a filtering operation. Because the chosen filter impulse response is a signal matched to the desired signal, we call this filtering operation *matched filtering*. Of course, the term “matched filter” does not explicitly mention the fact that the filter output is used as the test statistic in a threshold detector, so the reader must infer that information without help from the name.

Figure 5.2 depicts the general block diagram for a matched filter detector, showing the filter, the test statistic, and the threshold test. In order to maximize detector performance and enable the proper measurement of performance, it can be shown (Kay 1998) that the threshold test must be conducted at the time sample corresponding to the largest value (peak) of the test statistic $r[y(n)]$. Letting n^* denote the time index corresponding to the peak of the test statistic, we conduct the threshold test at $r[y(n^*)]$.

5.3.4 Example: Applying the Matched Filter to Our Example Signals

Figure 5.3 depicts an example of a correlation detector scheme. We see that the scheme consists of a cross-correlation operation, a test statistic and a thresholding operation. Note the plots depicting the various signals at each step of the scheme. Note the very low noise level in the plot of the absolute value of the cross-correlation. This demonstrates a key property of the matched filter, that the matched filter maximizes the SNR at the output of the filter/correlator. We discuss the processing scheme in detail in the following sections.

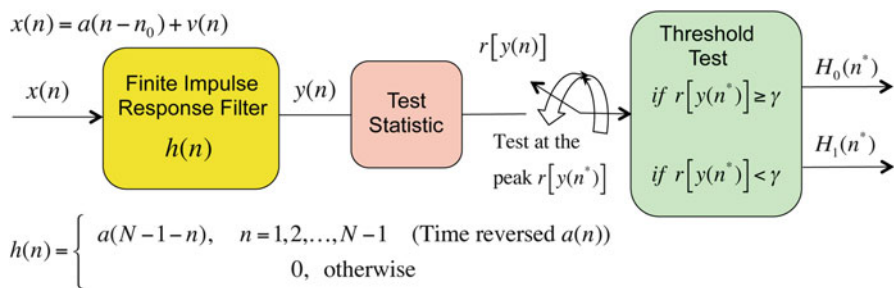


Fig. 5.2 General matched filter block diagram. Letting the time sample at which the test statistic is maximum be denoted by n^* , note that the threshold test is conducted at the peak $r[y(n^*)]$ of the test statistic r . Ideally, $n^* = n_0$. The decision threshold is denoted by γ

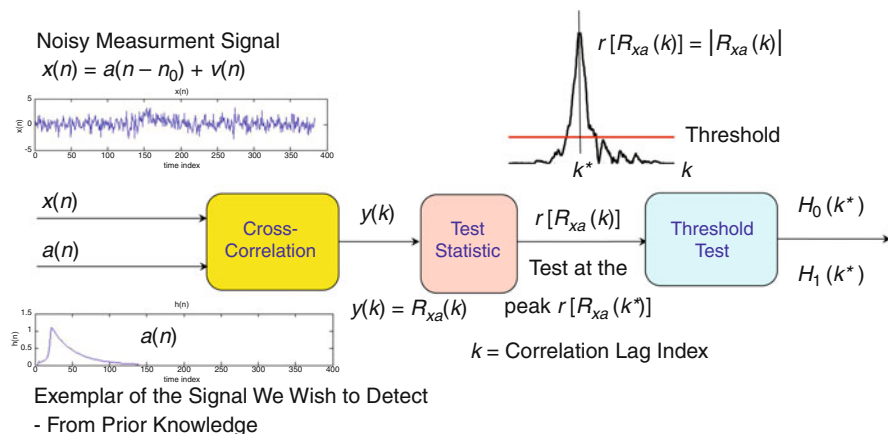


Fig. 5.3 Block diagram for the example problem. Letting the correlation lag index at which the test statistic is maximum be denoted by k^* , note that the threshold test is conducted at the peak of the test statistic, $r[R_{xa}(k^*)] = |R_{xa}(k^*)|$

5.4 Bayesian Binary Hypothesis Testing and Performance Measurement

Bayesian detection theory provides a rigorous foundation for evaluating detector performance (Whalen 1971; Van Trees 1968). Assume that we have a noisy one-dimensional (scalar) measured discrete-time signal (time series) as in Eq. 5.1. In general, the desired signal $a(n)$ can be deterministic or stochastic, and the noise $v(n)$ is modeled as stochastic and uncorrelated (statistically white) or correlated (statistically colored). In this tutorial, we model the desired signal $a(n)$ as deterministic (Papoulis 1965). We treat the noise as either uncorrelated or correlated. Most textbooks focus on the special case when the noise is Gaussian distributed (Whalen 1971; Kay 1998; Van Trees 1968), and for that case, the reader is directed to the references. For generality, this tutorial makes no assumptions about the form of the noise distribution.

We can use *binary hypothesis testing* to make decisions or *declarations* about whether or not the desired signal $a(n)$ is present in the measurement. The hypothesis that $a(n)$ is not present in measurement $x(n)$ is denoted H_0 , and the hypothesis that $a(n)$ is present in $x(n)$ is denoted H_1 .

$$\begin{aligned} H_0 : x(n) &= v(n) && \text{(Null Hypothesis: Desired Signal Not Present)} \\ H_1 : x(n) &= a(n - n_0) + v(n) && \text{(Alternative Hypothesis: Desired Signal Present)} \end{aligned} \quad (5.12)$$

Notice that this problem definition assumes that we do not know in advance the arrival time n_0 of the signal of interest $a(n - n_0)$. We must estimate the arrival time as part of the detection/classification process. As depicted in Figs. 5.2 and 5.3, the

general matched filter includes computing a test statistic $r[y(n)]$. This test statistic can take many forms; however, probably the most commonly used is the absolute value of the correlation result $y(n)$, as shown in the figure. Recall from before that this threshold test is conducted at the time sample at which the test statistic is maximum (Kay 1998). For our case, the test statistic is the absolute value of the cross-correlation. Because the correlation lag index is different from the signal time index n , we denote the lag index by k , and we denote the lag index at which the test statistic is maximum by k^* . This index k^* corresponds to the arrival time n_0 of the desired signal $a(n-n_0)$. Expressed mathematically, we define k^* as the correlation lag index that satisfies:

$$\max_{k \in [0, N-M]} |R_{xa}(k)| \quad (5.13)$$

Recall that N is the number of time samples in $x(n)$ and M is the length of $a(n)$. The set of k values over which to search is $[0, N - M]$ because that is the range over which the cross-correlation is affected by the signal $a(n)$ (Kay 1998). The decision is made when the test statistic evaluated at k^* is compared with the threshold γ :

$$r[R_{xa}(k^*)] \underset{H_0}{\overset{H_1}{\gtrless}} \gamma \quad (\text{Decision Rule}) \quad (5.14)$$

5.4.1 The Confusion Matrix (or Contingency Table)

Each time the hypothesis test is conducted, one of four events can occur: (1) H_0 is true and we declare H_0 , (2) H_0 is true and we declare H_1 , (3) H_1 is true and we declare H_1 , and (4) H_1 is true and we declare H_0 . The first and third alternatives correspond to correct choices. The second and fourth alternatives correspond to errors. The confusion matrix (or contingency table) depicted in Fig. 5.4 summarizes these four events, their associated probabilities, and the method for computing the four probabilities. The Bayes test assumes that there exist prior probabilities (priors) for the hypotheses and costs associated with the four courses of action. The priors $P(H_0)$ and $P(H_1)$ represent information available about the source prior to conducting the experiments. The costs for the four possible courses of action are given by C_{00} , C_{10} , C_{11} , and C_{01} , where C_{ij} is the cost of deciding H_i given that H_j is true. Once the costs have been assigned, the decision rule is based on minimizing the expected cost, which is known as the Bayes risk \mathfrak{R} (Van Trees 1968; Whalen 1971; Kay 1998):

$$\mathfrak{R} = \sum_{i=0}^1 \sum_{j=0}^1 C_{ij} P(H_i | H_j) P(H_j) \quad (\text{Bayes Risk}) \quad (5.15)$$

We assume throughout this discussion that the cost of an incorrect decision is higher than the cost of a correct decision. In other words, $C_{10} > C_{00}$ and $C_{01} > C_{11}$.

"Confusion Matrix" or Contingency Table for Binary Hypothesis Testing		
Truth Declaration	True Hypothesis H_0 (Null)	True Hypothesis H_1
Declared Hypothesis H_0 (Null)	$P(H_0 H_0) = P_{Spec} = \text{Specificity}$ $= \frac{\# H_0 \text{ Samples Declared } H_0}{\# H_0 \text{ Samples}}$	$P(H_0 H_1) = P_{Miss} = P(\text{Miss})$ $= \frac{\# H_1 \text{ Samples Declared } H_0}{\# H_1 \text{ Samples}}$
Declared Hypothesis H_1	$P(H_1 H_0) = P_{FA} = P(\text{False Alarm})$ $= \frac{\# H_0 \text{ Samples Declared } H_1}{\# H_0 \text{ Samples}}$	$P(H_1 H_1) = P_D = P(\text{Detection})$ $= \frac{\# H_1 \text{ Samples Declared } H_1}{\# H_1 \text{ Samples}}$
	$P(H_0 H_0) + P(H_1 H_0) = 1$	$P(H_0 H_1) + P(H_1 H_1) = 1$

$$P_{cc} = P(\text{Correct Classification}) = P(H_0 | H_0)P(H_0) + P(H_1 | H_1)P(H_1)$$

$$P_e = P(\text{Error}) = 1 - P_{cc} = P(H_0 | H_1)P(H_1) + P(H_1 | H_0)P(H_0)$$

Assume: Correct classification is given zero cost $\Rightarrow C_{00} = C_{11} = 0$

Incorrect classification is given full cost $\Rightarrow C_{01} = C_{10} = 1$

Fig. 5.4 Confusion matrix (contingency table): the two hypotheses are denoted H_0 , the null hypothesis, and H_1 , the alternative hypothesis. Note that for the special case in which the prior probabilities are equal ($P(H_0) = P(H_1) = \frac{1}{2}$), the probability of correct classification becomes $P_{CC} = \frac{1}{2}[P_D + (1 - P_{FA})]$. Note that we construct one confusion matrix for each value of the decision threshold γ

Note that in the confusion matrix, the probabilities in the two columns of the matrix each sum to one. An important special case of the Bayes criterion is that in which a correct classification is assigned zero cost and an incorrect classification is assigned full cost. In this case, we assign $C_{00} = C_{11} = 0$ and $C_{01} = C_{10} = 1$. Inserting these values in the Bayes Risk of Eq. 5.15 and using the fact that $P(\text{error}) + P(\text{correct classification}) = 1$, we obtain the probability of correct classification. Note that P_{cc} is the weighted sum along the main diagonal of the confusion matrix, weighted by the priors:

$$P(\text{Correct Classification}) = P_{CC} = P(H_1, H_1) + P(H_0, H_0) \tag{5.16}$$

$$= P(H_1 | H_1)P(H_1) + P(H_0 | H_0)P(H_0) \tag{5.17}$$

Often in practice, there exists insufficient information about an experiment to allow the user to assign the values of the prior probabilities P_0 and P_1 . In this case, it is common to assume that the priors are equal (no information), so $P(H_0) = P(H_1) = \frac{1}{2}$. Under this condition, the probability of correct classification becomes

$$P_{CC} = \frac{1}{2}[P(H_1|H_1) + P(H_0|H_0)] \quad (5.18)$$

This can now be written in terms of the probability of detection and probability of false alarm as follows:

$$P_{CC} = \frac{1}{2}[P_D + (1 - P_{FA})] \quad (\text{Probability of Correct Classification}) \quad (5.19)$$

5.4.2 Bayesian Hypothesis Testing for Multidimensional (Vector) Data

Let us now generalize our discussion. Our example problem specifies a scalar measurement signal $x(n)$. However, in general, we can have a vector of J observations denoted as follows.

$$\underline{X} = [x_1, x_2, \dots, x_J]^T \quad (\text{Observation Vector}) \quad (5.20)$$

The observations $x_j, j = 1, 2, \dots, J$ are called features of the physical process being observed, and T denotes the vector transpose. In our example problem, the vector has one element $x(n)$, the measurement signal. We assume throughout this discussion that the cost of an incorrect decision is higher than the cost of a correct decision. In other words, $C_{10} > C_{00}$ and $C_{01} > C_{11}$. Under this assumption, the detector that minimizes the Bayes risk is given by the following (Van Trees 1968):

$$\frac{f(\underline{X}|H_1)}{f(\underline{X}|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \frac{P(H_0)(C_{10} - C_{00})}{P(H_1)(C_{01} - C_{11})} \quad (\text{Bayes Decision Criterion}) \quad (5.21)$$

where $f(\underline{X}|H_1)$ denotes the conditional probability density function (pdf) of the observation vector \underline{X} given that hypothesis H_1 is true, and $f(\underline{X}|H_0)$ denotes the conditional pdf of \underline{X} given that hypothesis H_0 is true (Kay 1998). The ratio of the conditional densities is called the likelihood ratio and is denoted by $\Lambda(\underline{X})$:

$$\Lambda(\underline{X}) = \frac{f(\underline{X}|H_1)}{f(\underline{X}|H_0)} \quad (\text{Likelihood Ratio}) \quad (5.22)$$

Because this is a ratio of two functions of a random variable, the likelihood ratio is a random variable. A very important result is that regardless of the dimensionality of the observations \underline{X} , the likelihood ratio $\Lambda(\underline{X})$ is a one-dimensional variable. This idea is of fundamental importance in hypothesis testing. Regardless of the dimension of the observation space, the decision space is one dimensional. The quantity on the right-hand side of the relation (5.21) is the threshold of the test and is denoted by γ :

$$\gamma \triangleq \frac{P(H_0)(C_{10} - C_{00})}{P(H_1)(C_{01} - C_{11})} \quad (\text{Decision Threshold}) \quad (5.23)$$

Thus, the Bayes criterion leads to a likelihood ratio test:

$$\Lambda(\underline{X}) \underset{H_0}{\overset{H_1}{\gtrless}} \gamma \quad (\text{Likelihood Ratio Test}) \quad (5.24)$$

We see that the test threshold allows for weighting according to the priors and the costs. This allows the user flexibility in choosing a threshold that is best for the problem at hand. Note that if we have available the conditional pdfs or estimates of them, we can construct the confusion matrices (one for each value of the threshold) by integrating under the pdfs as depicted in Fig. 5.6. In most practical problems, we do not have the pdfs, so we construct the confusion matrices using the threshold method shown in Fig. 5.2.

5.4.3 Training and Testing Phases of the Detection Process

In order to measure detection performance, we must be able to conduct controlled experiments in which we know the correct experimental outcomes (“ground truth”) a priori (Duda and Hart 1973; Duda et al. 2001). This is a very important point that is often overlooked. Our use of the threshold detector occurs in phases. (1) First, in the *Training Phase*, we conduct controlled experiments in which the ground truth is known in order to construct a Receiver Operating Characteristic (ROC) curve from which we determine the appropriate operating threshold for the detector. We use a set of known “training data” during this phase. (2) Second, in the *Testing Phase*, we use the selected operating threshold on a set of “testing data,” from which we produce operational detection results. The remainder of this section concentrates on the training phase. Note that the training and testing process involves a considerable “leap of faith” in which we assume that the training data are representative of the test data. Therefore, extreme care must be taken to ensure that this assumption is valid. Otherwise, interpretations of the processing results may be meaningless.

The user should carefully design the experiments so the statistical sample size (number of data samples we can use for performance evaluation) is large enough to enable the computation of meaningful confidence intervals (see the following two subsections). Figure 5.5 summarizes some rules of thumb for picking the training and testing data sets. Using ground truth, we create a labeled data set in which each data sample is labeled H_0 or H_1 . *Rule of Thumb 1*: We generally want to have a large sample size (approximately 100 or 1,000 or more). However, the real world often does not allow us this luxury. *Rule of Thumb 2*: If the data are Gaussian distributed, we need a minimum of about 30 or more samples to obtain reasonable confidence

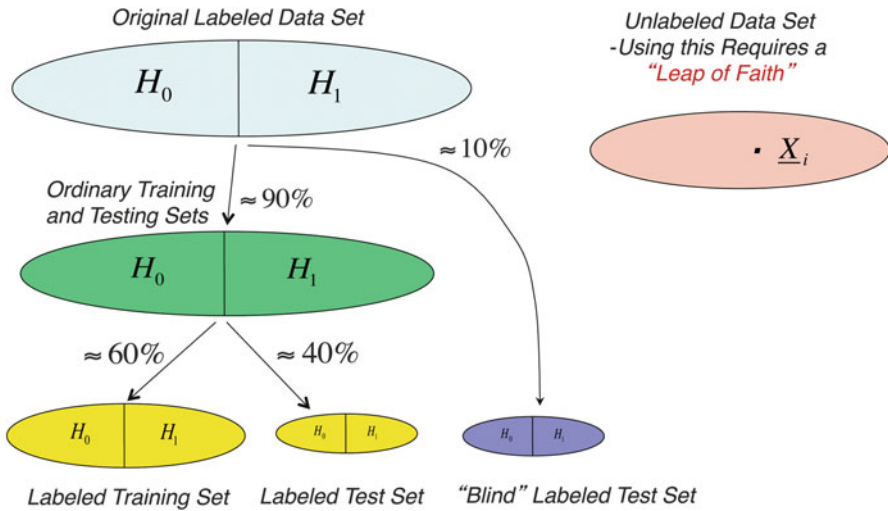


Fig. 5.5 (Left) Training data sets for controlled experiments are depicted on the left side. These include the standard labeled training and testing sets, as well as a possible blind test set that can be used if enough data are available. When using this, we can ask some unbiased users to conduct the experiment. (Right) The unlabeled test set is used after training and testing in the controlled experiments. To minimize the “leap of faith,” these data should be representative of the data in the training and testing sets

interval estimates (Hogg and Craig 1978). If our sample size is small, we turn to “hold one out” and bootstrap algorithms for estimating confidence intervals (Zoubir and Iskander 2004). Even with these methods, we cannot take too seriously any results with very small sample sizes. *Rule of Thumb 3*: Generally, we should set aside about 60 % of our labeled samples for the training phase and the other 40 % for the testing phase (Hand 1981; Devijver and Kittler 1987). *Hard Rule*: Never test on the training set. That is cheating and leads to improper inferences. Despite warnings, many people continue to do this. Please do not become one of them.

5.4.4 The Receiver Operating Characteristic (ROC) Curve

A Receiver Operating Characteristic (ROC) curve is constructed to quantify the tradeoff between the probability of detection P_D and the probability of false alarm P_{FA} versus the detection threshold γ , as depicted in Fig. 5.6 (Whalen 1971; Duda et al. 2001; Duda and Hart 1973; Van Trees 1968; Kay 1998). Note that this requires the construction of one confusion matrix for each value of the threshold. We vary the value of the decision threshold γ over the full range of values of the decision statistic $r[y(n)]$. For each measurement signal in the ensemble, at the time sample n^* , we compare the decision statistic to the decision threshold as in Fig. 5.2. If $r[y(n^*)] < \gamma$, we declare that H_0 is true. If $r[y(n^*)] \geq \gamma$, we declare that H_1 is true.

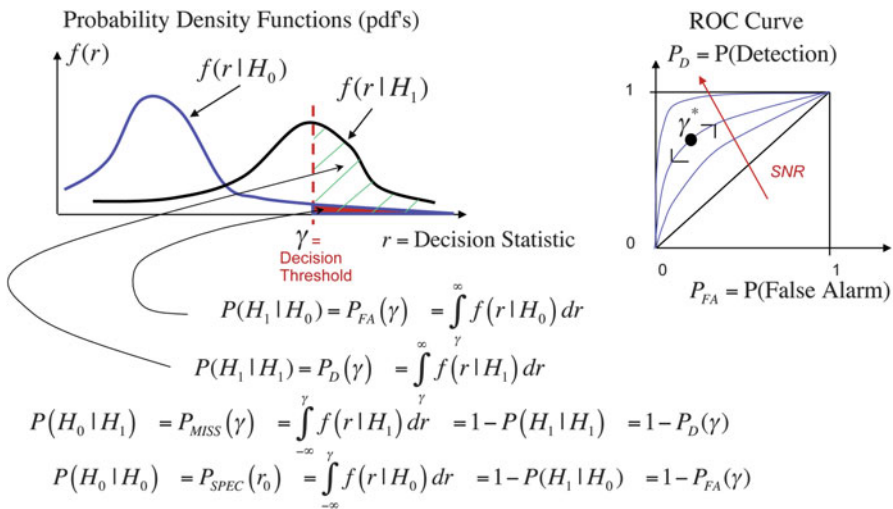


Fig. 5.6 The Receiver Operating Characteristic (ROC) curve can be constructed by integrating the conditional probability density functions depicted on the left. This, of course, assumes that the pdfs or estimates of them are available. If not, then we generally use the confusion matrix method

Once we have the ROC curve, we choose the “operating threshold” γ^* to be the one that maximizes $P_D(\gamma)$ and minimizes $P_{FA}(\gamma)$. This threshold value is found at the “knee” of the ROC curve at the appropriate SNR for the experiment. *Note that this knee occurs for the threshold value at which the probability of correct classification P_{CC} is maximum.* We denote this by $P_{cc}(\gamma^*)$.

5.4.5 Statistical Confidence Interval About the Probability of Correct Classification

The author believes that “the ROC curve is not finished until we have computed the confidence interval on the probability of correct classification.” Unfortunately, this last step is almost always overlooked by most practitioners. P_{cc} is not a deterministic quantity. It can be viewed as a random variable with an associated distribution.

The classifier/detector performs a random experiment, the outcome of which can be classified in one of two mutually exclusive and exhaustive ways: success or failure. Success means that the classification is correct. Failure means that the classification is incorrect. Let N equal the number of independent trials. Let $p = P_{cc}$ = the probability of correct classification. Assume that the true value p is the same on each repetition. Let $q = 1 - p$ = probability of error. For the classification problem in which we conduct an experiment, we can calculate the estimated quantities in the confusion matrix. The maximum likelihood estimate of p is given by \hat{p} and is the estimated \hat{P}_{CC} computed in the confusion matrix. We can

write the confidence interval about the true value of p as follows, where α , a probability, is the significance of the test.

$$P\{L < p < U\} = 1 - \alpha \quad (\text{Confidence Interval about } p) \quad (5.25)$$

where L and U are the lower and upper bounds, respectively, of the confidence interval. The most common interpretation is to read the confidence interval relation above as follows: “With confidence $1 - \alpha$, the true p lies between L and U .” However, this interpretation is not generally supported by statistical rigor. The preferred interpretation is: “Prior to the repeated independent performances of the random experiment, the probability is $1 - \alpha$ that the random interval (L , U) includes the unknown fixed point (parameter) p (Hogg and Craig 1978).” For our example problem, we arbitrarily choose $\alpha = 0.05$ so we have a “95 percent confidence interval.”

There exists a significantly large literature on how to compute the lower and upper bounds L and U . In most practical cases, the author prefers to use bounds that do not assume a particular distribution and which can be used for both small and large sample sizes N . To this end, a reasonable set of bounds is the following (Hogg and Craig 1978):

$$L = N\hat{p} + 2 - 2\sqrt{\frac{N\hat{p}(1-\hat{p})+1}{N+4}} \quad (\text{Lower Bound on } p) \quad (5.26)$$

$$U = N\hat{p} + 2 + 2\sqrt{\frac{N\hat{p}(1-\hat{p})+1}{N+4}} \quad (\text{Upper Bound on } p) \quad (5.27)$$

We can evaluate L and U and plot them versus the true p and the estimated \hat{p} , for various values of the sample size N , as in Figs. 5.8 and 5.10. These plots are very instructive in showing how confidence intervals tighten as the sample size increases. Bootstrap techniques for estimating confidence intervals from the data for small sample sizes are discussed in Zoubir and Iskander (2004).

5.5 Processing for the Example Problem

5.5.1 Experiment Design for the Example Problem

The *training data* were created so as to allow the computation of the confusion matrices necessary for computing the ROC curve and the confidence interval on P_{cc} . Please refer to Fig. 5.5. For the null hypothesis H_0 , we simulated an ensemble of $P = 300$ *labeled* realizations (each realization is labeled H_0) of the simulated measurement waveform $x(n)$ specified in Eq. 5.15 in which the desired signal $a(n) = 0$. Each realization of $x(n)$ is different because each has a different realization of the stochastic process $v(n)$. For hypothesis H_1 , we created an ensemble of $P = 300$ *labeled* realizations of the simulated measurement waveform $x(n)$ in which the desired signal $a(n)$ is nonzero. Again, each realization of $a(n)$ is different

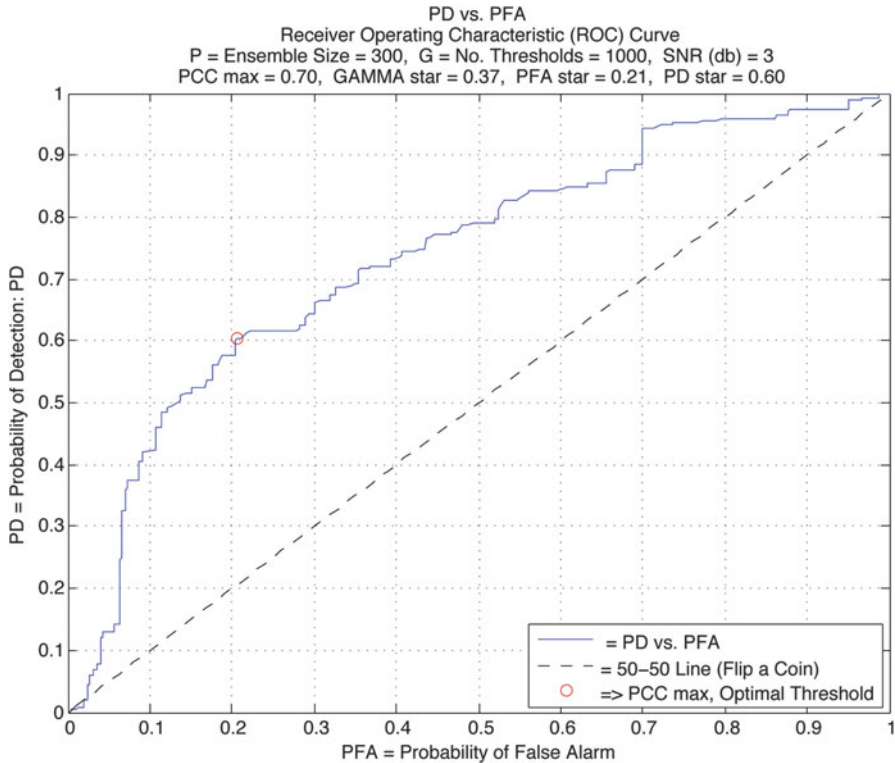


Fig. 5.7 Example problem for SNR = 3 dB: the ROC curve is plotted given the sample size $P = 300$ samples per class (300 for H_0 and 300 for H_1). The abscissa is probability of false alarm P_{FA} , and the ordinate is probability of detection P_D computed using confusion matrices. The knee of the curve (marked by a circle) occurs at $P_{CC}(\gamma^*) = .698$, $P_{FA}(\gamma^*) = .21$ and $P_D(\gamma^*) = .60$. The corresponding operating detection threshold is $\gamma^* = 0.37$

because each has a different realization of the stochastic process $v(n)$. These ensembles are sufficient to compute the performance indices. For *controlled testing* purposes, we can similarly simulate two more labeled ensembles of realizations of the stochastic process $x(n)$. For *blind controlled testing*, we can similarly simulate two additional *labeled* ensembles. For *uncontrolled testing*, we can similarly simulate two more *unlabeled* ensembles.

5.5.2 Processing Results for the Example Problem

For our example, we have no prior knowledge from which to derive prior probabilities $P(H_0)$ and $P(H_1)$ so we assume that the two hypotheses are equally likely, and $P(H_1) = P(H_0) = \frac{1}{2}$. Under this condition, the expression for P_{cc} is simplified, as shown in Eq. 5.19.

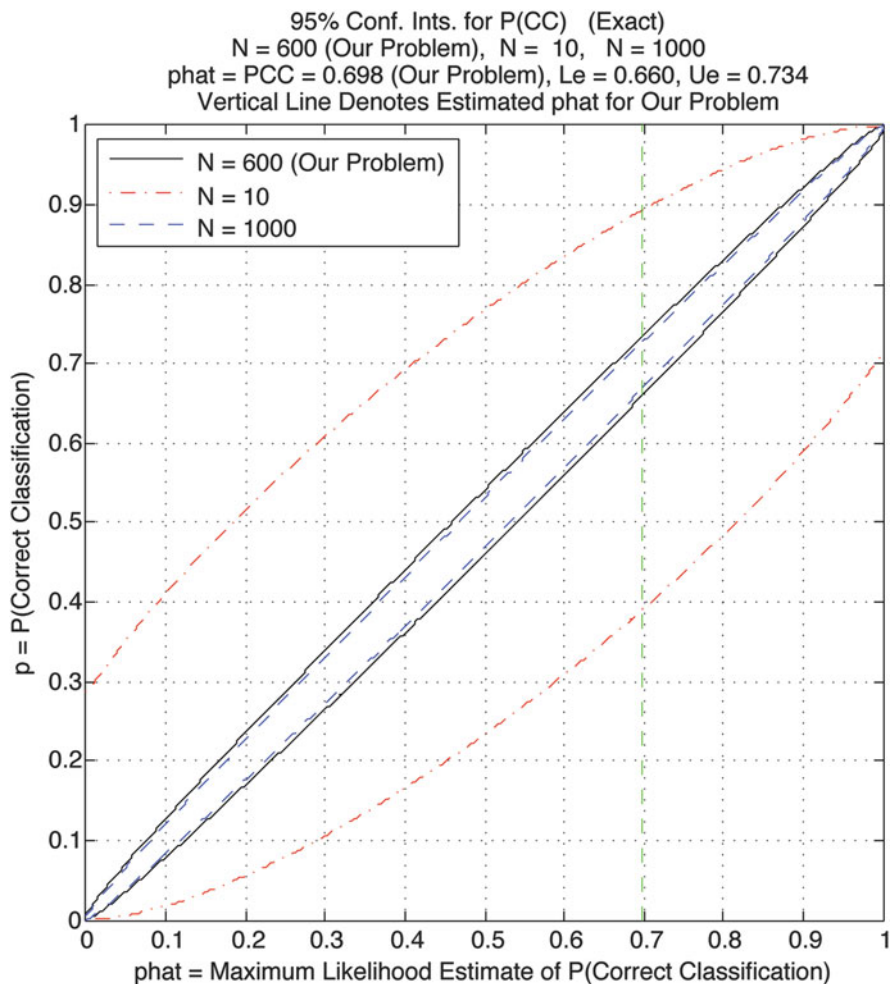


Fig. 5.8 Example problem for SNR=3 dB: the abscissa is \hat{p} = the maximum likelihood estimate of the probability of correct classification. The ordinate is p = the true value of the probability of correct classification. The 95 % ($1 - \alpha = .95$) confidence interval bounds (L, U) = (.66, .734) for the probability of correct classification are plotted, given the sample size $N = 600$ and $\hat{p}^* = P_{CC}(\gamma^*) = .698$ (the green vertical line). Note that the confidence interval tightens as the sample size N increases

Figure 5.7 depicts a ROC curve constructed for our threshold detector example in which the SNR is 3 dB. We see that with this low SNR, the curve is far away from the desired upper left-hand corner of the diagram, and the knee of the curve (marked by a circle) occurs at $P_{CC}(\gamma^*) = .667$, $P_{FA}(\gamma^*) = .36$ and $P_D(\gamma^*) = .76$. Figure 5.8 presents the confidence interval about the P_{CC} for this example in which the SNR is 3 dB. For our problem, the number of samples (signals) in the training data for both

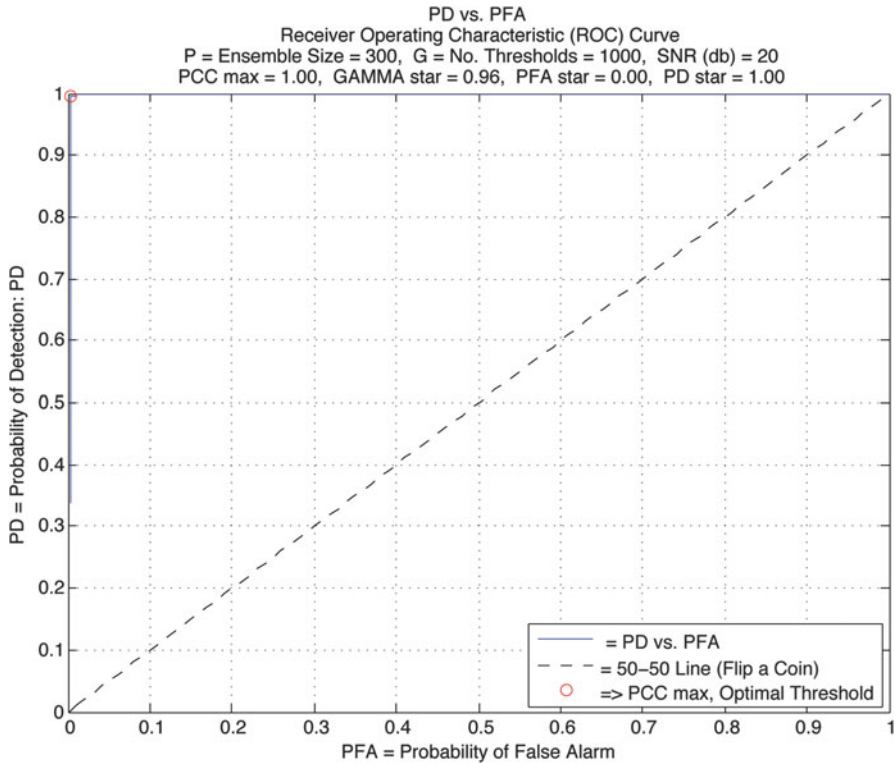


Fig. 5.9 Example problem for SNR = 20 dB: the ROC curve is plotted given the sample size $P = 300$ samples per class (300 for H_0 and 300 for H_1). The abscissa is probability of false alarm P_{FA} , and the ordinate is probability of detection P_D computed using confusion matrices. The knee of the curve (marked by a circle) occurs at $P_{CC}(\gamma^*) = .997$, $P_{FA}(\gamma^*) = .01$ and $P_D(\gamma^*) = 0.99$. The corresponding operating detection threshold is $\gamma^* = 0.96$

H_0 and H_1 is $N = 600$ and the estimate of $P_{CC}(\gamma^*) = \hat{p} = .667$. The green vertical line depicts \hat{p} , and the curves it crosses depict the lower and upper bounds on p . We see that the confidence interval is given by $P(.627 < p < .704) = .95$. The corresponding operating detection threshold is $\gamma^* = 0.36$. For tutorial purposes, the figure also shows what the bounds would be if $N = 10$ and $N = 1000$. We see that for small N , the bounds are very wide and for large N , the bounds are much narrower, as expected.

In Fig. 5.9, the SNR is 20 dB and the ROC curve lies in the desired upper left-hand corner. Here, the knee of the curve occurs at $P_{CC}(\gamma^*) = .997$, $P_{FA}(\gamma^*) = .01$ and $P_D(\gamma^*) = .99$. The corresponding operating detection threshold is $\gamma^* = 0.91$. Figure 5.10 shows the confidence interval for $N = 600$ and $\hat{p} = .993$. The estimated 95 % confidence interval is given by $P(.988 < p < .997) = .95$. We see the clear performance improvement that occurs with increased SNR.

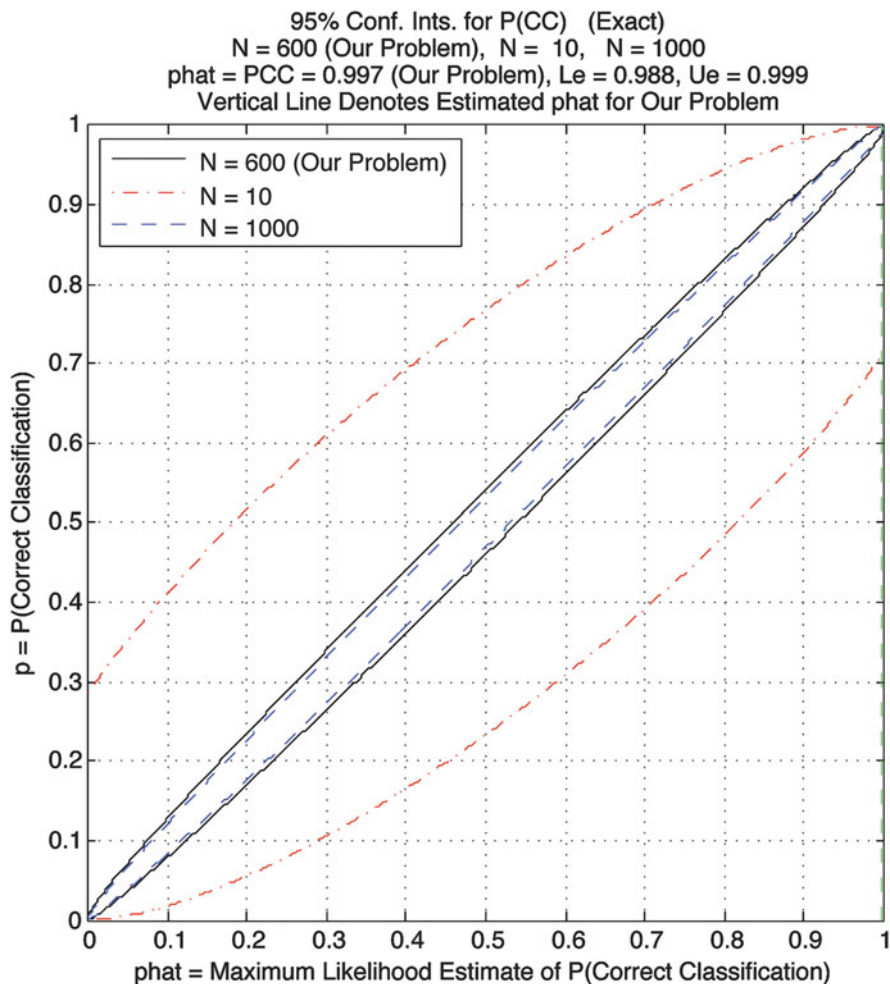


Fig. 5.10 Example problem for SNR=20 dB: The abscissa is \hat{p} = the maximum likelihood estimate of the probability of correct classification. The ordinate is p = the true value of the probability of correct classification. The 95 % ($1 - \alpha = 0.95$) confidence interval bounds $(L, U) = (.988, .999)$ for the probability of correct classification are plotted, given the sample size $N = 600$ and $\hat{p}^* = P_{CC}(\gamma^*) = .997$ (the *green vertical line*). Note that the confidence interval tightens as the sample size N increases

5.5.3 Conclusions

In this chapter, we introduce the concept of matched filtering for detecting desired signals buried in noisy measurement signals. We show that the *matched filter* is another name for the *correlation detector*, which exploits prior knowledge in the form of an exemplar of the desired signal. We use an example detection problem to

demonstrate the matched filtering approach. We see that the detection methodology comes from hypothesis testing algorithms in Bayesian detection theory. This Bayes approach gives us very powerful methods to choose the detection threshold and evaluate detection performance in the form of the Receiver Operating Characteristic (ROC) curve and the statistical confidence interval about the probability of correct classification. We show that the matched filter can be an effective detection tool when exemplars of the desired signal are available a priori.

5.6 Part II Auditory Matched Filtering: Biological Examples from Selected Vertebrates

According to the auditory matched filter hypothesis (Capranica and Moffat 1983), auditory information processing in sub-mammalian vertebrates (e.g., fishes, amphibians, reptiles, and to some extent birds) relies on extensive peripheral prefiltering. In other words, the auditory sensory filter (ear) is often tuned to signals of biological importance to the species so that less post-processing is required by the reduced central nervous systems in these species. In contrast, mammals can afford to “take in” all sensory input and rely on their superior brain power to sort out the meaning behind the message. The optimum receiver strategy, according to one formulation of the matched filter hypothesis, is to “design” a bias into the frequency response of the auditory system (Capranica and Moffat 1983; Wehner 1987). Rather than a high-fidelity flat frequency response (no bias), the receiver’s auditory system should have a frequency response which closely matches the envelope of the energy spectrum of the emitter’s call. This ensures that the receiver maximizes the signal-to-noise ratio in the frequency domain for that particular call. Figure 5.11 illustrates the decision criterion of such a matched filter detector. The received

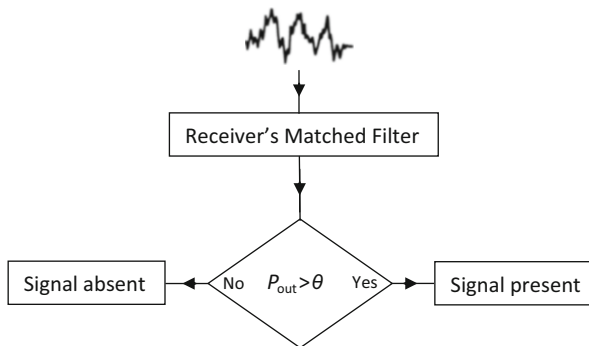


Fig. 5.11 Schematic diagram for the decision criterion of a matched filter detector. Signal propagation through the environment results in a noisy signal at the input to the receiver’s matched filter. If the power output (P_{out}) of the filter exceeds a certain threshold (θ), the receiver decides that the emitter’s signal is present. If P_{out} remains below this θ , the receiver decides that no signal is present. By adjusting the internal threshold, the reliability of the receiver’s decision can be modulated (Modified from Capranica and Rose (1983))

signal has been contaminated by unwanted noise. After passing through the filter, the receiver must make a decision if the signal is present or not. One decision criterion is to simply ask if the power (energy per unit time) output of the matched filter at any time exceeds some internal threshold (θ); if it has, the signal is present; if not, the signal is not present. But this statistical decision-making process depends critically on the setting of θ . If θ is set too low, the false alarm rate will be high; if θ is set too high, the receiver runs the risk of missing some of the signals (Capranica and Moffat 1983). In the following section, we have chosen several examples from the literature of sub-mammalian vertebrates which appear to exhibit a conspicuous match between the properties of their acoustic signals and the tuning of their auditory systems and thus provide evidence biological matched filters (Wehner 1987).

5.6.1 Weakly Electric Fish

5.6.1.1 Correspondence of Electric Organ Discharge and Electroreceptor Tuning

The wavelike electric organ discharge (EOD) in some gymnotoid species of South American weakly electric fish is one of the most regular of all known biological phenomena (Heiligenberg 1991). The internal fluctuation in the EOD rate is ca. 0.01 %. In these fish, the electric field is generated by the electric organ located in the tail, and the field is sensed by electroreceptors in the head region. In a landmark study of gymnotoids, the EOD frequencies of three species were recorded and compared to the best frequencies (BFs) of a population of tuning curves obtained electrophysiologically from individual fibers innervating electroreceptors (Hopkins 1976). Figure 5.12 is a plot of the electroreceptor BFs versus the EOD frequencies from the same fish, for three different species. Clearly, there is a high degree of matching between the electroreceptor tuning and the EOD rate.

5.6.1.2 Effect of Androgens on the Matched Filter

Meyer and Zakon (1982) confirmed the strong matching between the EOD discharge rate and electroreceptor tuning in another species of South American weakly electric fish—*Sternopygus*. Moreover, they demonstrated that systematic treatment of these fish with androgens—in this case 5 α -dihydrotestosterone (DHT)—lowered their EOD rate. Concomitantly, DHT caused decreases in electroreceptor best frequencies over a 2-week period, maintaining the close match between discharge frequency and receptor tuning. Thus, electroreceptor tuning is dynamic and it parallels natural shifts in the EOD frequency.

5.6.2 Anuran Amphibians

To help ensure that an appropriate behavioral response is evoked during acoustic communication, the anuran auditory system is often tuned to salient spectral and/or

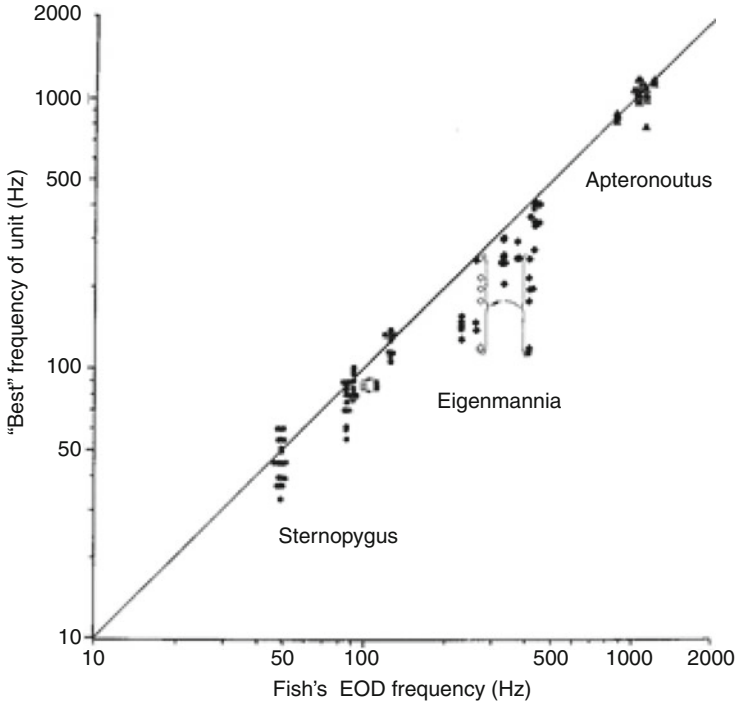


Fig. 5.12 Plot of the electroreceptor BFs versus the EOD frequencies from the same fish, for three different species of South American weakly electric fish (*Sternopygus macrurus*, *Eigenmannia virescens*, and *Aptereronotus albifrons*). Clearly, there is a high degree of matching between the electroreceptor tuning and the EOD rate. The *diagonal line* shows perfect correspondence between EOD frequency and BF (From Hopkins (1976))

temporal features of the conspecific call (Frishkopf et al. 1968; Capranica and Moffat 1975; Capranica and Rose 1983). The elegant coevolution of this relatively straightforward acoustic system has made anurans an extremely valuable neuroethological model for the study of acoustic communication.

5.6.2.1 The Puerto Rican Coqui (*Eleutherodactylus coqui*)

The Puerto Rican Coqui frog, *Eleutherodactylus coqui* (*Anura: Leptodactylidae*), is abundant in Puerto Rico where it can be found at altitudes from sea level to over 1,000 m above sea level (a.s.l.). Males of this arboreal amphibian are territorial, spaced several meters from each other, and call from tree branches or vegetation from sunset to shortly after midnight throughout 11 months of the year. They produce a characteristic two-note call (“Co-Qui”) in which each note has a different significance for each sex: males use the “Co”-note for territorial defense, while females are attracted to the “Qui”-note (Narins and Capranica 1976, 1978). In this species, the advertisement calls and snout-vent length (SVL) both exhibit an altitudinal gradient such that at 30 m.a.s.l., small males produce short, rapidly

repeating, high-pitched calls, whereas at 1,000 m.a.s.l., males are larger and the calls are longer, lower pitched, and repeated more slowly (Narins and Smith 1986). For example, the “Co”-note frequency produced by males at 30 m.a.s.l. is about a third of an octave higher in frequency than that produced by males at 1,000 m.a.s.l. More recently, it was found that the spectral contents of the males’ “Co”-note calls and the frequency to which the inner ear is most sensitive are tightly correlated and change concomitantly along this same altitudinal gradient (Meenderink et al. 2010). In that study, advertisement calls of males of *E. coqui* were recorded in situ along an altitudinal gradient ranging from 30 to 1,005 m.a.s.l. Following the recordings, males were captured and transported to a nearby lab in Puerto Rico where, the following day, distortion product otoacoustic emissions (DPOAEs) were measured from the anesthetized males. This was done by sweeping two primary tones (f_1 and f_2) from low frequencies to high frequencies and plotting the amplitude of the resulting third-order distortion product emission ($2f_1-f_2$) versus the lower primary frequency, f_1 , resulting in a DPOAE audiogram. From this audiogram, the frequency that results in maximum DPOAE amplitude (F_{maxDP}) can be identified and interpreted as the frequency to which the ear is most sensitive, or the frequency to which the ear is “tuned” (Meenderink et al. 2010). This frequency was then plotted against the frequency of the animal’s “Co”-note in its advertisement call. The resulting strong correlation (Fig. 5.13) is good evidence for a close match between the call note frequency (signal) and the peripheral auditory tuning (receiver characteristic) along the entire altitudinal gradient inhabited by these vocal animals. It was suggested that the animal’s body size, conditioned by the calling site temperature, determines the frequencies of the emitted calls and the best sensitivity of the inner ear (Meenderink et al. 2010).

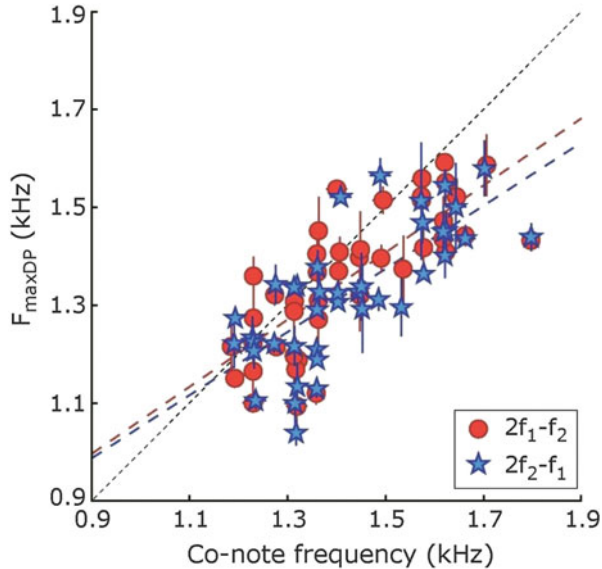
5.6.2.2 Ultrasonic Frogs

The Concave-Eared Torrent Frog (*Odorrana tormota*)

Odorrana tormota (previously *Amolops tormotus*) is known only from two provinces in central China (Zhou and Adler 1993). This species has unusually high-pitched calls containing substantial energy in the ultrasonic frequency range (above 20 kHz), and its hearing extends from less than or equal to 1 kHz to approximately 35 kHz (Narins et al. 2004; Feng et al. 2006), dramatically exceeding previously reported upper limits of anuran frequency sensitivity¹ (e.g., 8 kHz, Loftus-Hills and Johnstone 1970; 5 kHz, Fay 1988). Playback experiments in the animal’s natural habitat demonstrated that the ultrasonic elements are behaviorally relevant, and thus this extraordinary upward extension into the ultrasonic range of both the harmonic content of the advertisement calls and the frog’s hearing sensitivity is likely to have coevolved in response to the intense, predominantly

¹ Another sympatric species, *Odorrana graminea* (formerly *O. livida*), also has an extended high-frequency range (to 22–24 kHz) despite lacking the sunken tympana of *O. tormota* (Feng et al. 2006; Liu et al. 2014).

Fig. 5.13 Scatter plot showing the relationship between the frequency at maximum DPOAE amplitude ($F_{\max DP}$) and the dominant Co-note frequency in the call of the Puerto Rican Coqui frog, *Eleutherodactylus coqui*. Data points indicate median values \pm interquartile ranges. Gray circles, $2f_1-f_2$; black stars, $2f_2-f_1$. The diagonal (dashed) line represents equality between Co-note frequency and $F_{\max DP}$ (From Meenderink et al. (2010))



low-frequency ambient noise from local streams (Narins et al. 2004; Feng et al. 2006; Feng and Narins 2008). Because amphibians are a distinct evolutionary lineage from microchiropterans and cetaceans [which have evolved ultrasonic hearing to minimize spectral overlap in the frequency bands used for sound communication (Sales and Pye 1974) and to increase hunting efficacy in darkness (Bradbury and Vehrencamp 2011)], ultrasonic perception in amphibians represents a new example of independent evolution. Moreover, this example illustrates how a matched filter, when subject to selection pressure in the form of ambient noise, can respond appropriately to maintain the signal-to-noise ratio necessary for communication of biologically significant signals.

The Hole-in-the-Head Frog (*Huia cavitympanum*)

In addition to *Odorrana tormota*, only one other anuran species, *Huia cavitympanum*, is currently known to have recessed tympanic membranes (Inger 1966). *Odorrana tormota* and *H. cavitympanum* are both southeast Asian species in the family Ranidae, yet they do not overlap in geographical distribution and are unrelated at the generic level (Cai et al. 2007; Stuart 2008). The habitats in which the frogs are found, however, are remarkably similar; males of both species call in close proximity to rushing streams that produce substantial broadband background noise. Given the similarity of the species' acoustic environment and peripheral auditory morphology, Arch et al. (2008) predicted that they may have converged on the use of ultrasound for intraspecific communication. Recordings of the calls of *H. cavitympanum* in their natural habitat in Borneo obtained with ultrasonic detection and recording equipment demonstrated that males of this species are able to produce calls that are comprised entirely of ultrasound (Arch et al. 2008). To test

the hypothesis that these frogs use purely ultrasonic vocalizations for intraspecific communication, a series of acoustic playback experiments with male frogs in their natural calling sites was performed (Arch et al. 2009). These workers found that the frogs responded with increased calling to broadcasts of conspecific calls containing only ultrasound. The field study was complemented by electrophysiological recordings from the auditory midbrain and by laser Doppler vibrometer measurements of the tympanic membrane's response to acoustic stimulation. These measurements revealed that the frog's auditory system is broadly tuned over high frequencies, with peak sensitivity occurring within the ultrasonic frequency range (>20 kHz). Thus, *H. cavitympanum* is the first non-mammalian vertebrate reported to communicate with purely ultrasonic acoustic signals.

Structural Basis of the Matched Filter

In a detailed study of the inner ears of the three species of frog known to detect ultrasounds (*Odorrana tormota*, *Odorrana graminea*, and *Huia cavitympanum*), Arch et al. (2012) attempted to identify morphological correlates of high-frequency sound detection. These workers found that the three ultrasound-detecting species have converged on a series of small-scale functional modifications of the basilar papilla (BP), the high-frequency hearing organ in the frog inner ear. These modifications include (1) reduced BP chamber volume, (2) reduced tectorial membrane mass, (3) reduced hair bundle length, and (4) reduced hair cell soma length. While none of these factors on its own could completely account for the ultrasonic sensitivity of the inner ears of these species, the combination of these factors appears to extend their hearing bandwidth and thus facilitate high frequency/ultrasound detection. Similar morphological modifications are seen in the inner ears of *O. chloronota*—a poorly known species from the mountains of Laos. In fact, the striking similarity of the BP features of *O. chloronota* to those of the three amphibian species known to detect ultrasound suggests that this species is a potential candidate for high-frequency hearing sensitivity. These data form the foundation for future functional work probing the physiological bases of ultrasound detection by a non-mammalian ear (Fig. 5.14).

5.6.2.3 *Eupsophus roseus*: A Leptodactylid Frog from the South American Temperate Forest

In a recent test of the matched filter hypothesis, Moreno-Gomez et al. (2013) sought to test the concordance between the acoustic sensitivity of female frogs of *E. roseus* and (a) the spectral characteristics of the advertisement calls of conspecific males and (b) the spectral characteristics of the ambient noise in which these frogs breed. Audiograms measured from the torus semicircularis in the midbrain of anesthetized females exhibited two sensitivity peaks: one in the low-frequency range (LFR <700 Hz) and the second in the high-frequency range (HFR >700 Hz). Advertisement calls of conspecific males were characterized by three dominant harmonics of which the second and third fell within the bandwidth of the lowest thresholds in the female's HFR. In fact, the mean cross-correlation coefficient between the audiograms and the conspecific vocalization spectra was 0.4 (95 % CI: 0.3–0.5).

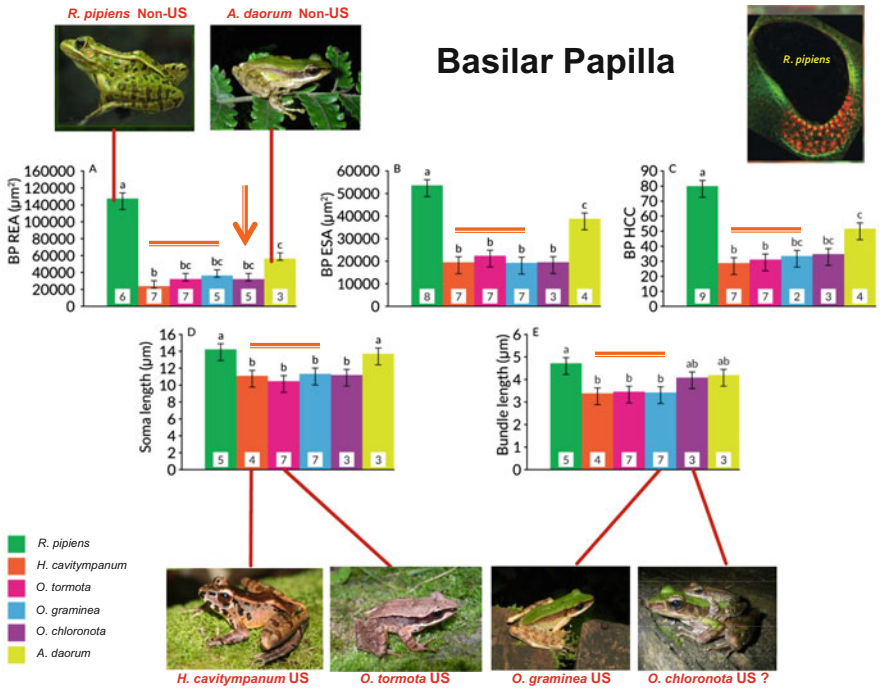


Fig. 5.14 Comparison of morphometric data from the basilar papillae of six amphibians. (a) Recess entrance area (REA); (b) Epithelium surface area (ESA); (c) Hair cell count (HCC); (d) Hair cell soma length; (e) Hair cell bundle length. Numbers indicate sample sizes. Letters denote statistically significant differences in pairwise comparisons using Tukey’s post hoc analysis with $\alpha = 0.05$. If a pair of species shares a common letter, they are not significantly different in that trait. Horizontal bars indicate the three amphibian species known to detect ultrasound (US). Vertical arrow indicates putative ultrasound detector—*O. chloronota*. Inset (upper right) shows the BP from *Rana pipiens*, a North American species known not to detect ultrasound (Non-US) (Modified from Arch et al. (2012))

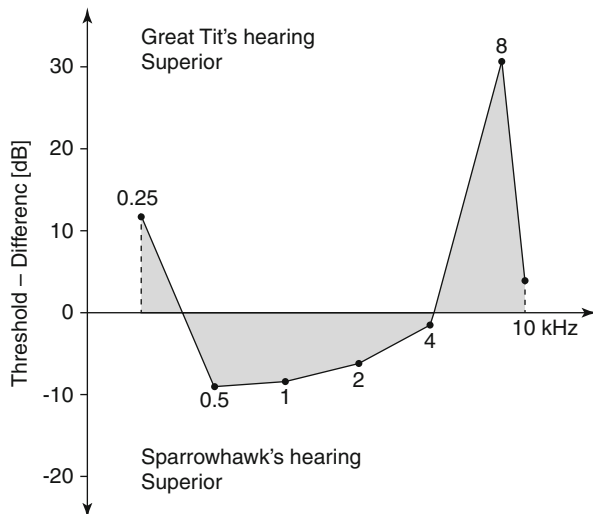
This coefficient was significantly higher than that between the audiograms and the background noise spectra over the 4 months for which data were available (Moreno-Gomez et al. 2013). Both this measured concordance and the mismatch between the auditory sensitivity of *E. roseus* females with both the local abiotic and biotic background noise are interpreted as adaptations to increase the signal-to-noise ratio in this communication system.

5.6.3 Birds

5.6.3.1 Unmatched Filters Between Predators and Prey

By exploiting call frequencies heard well by conspecifics but poorly by a prey species, animals may use a species-specific “private channel” to their advantage.

Fig. 5.15 The difference in the absolute thresholds between the great tit and the European sparrowhawk (thresholds in both species calculated using a $d' = 1.5$) (From Klump et al. (1986))



Using the method of constant stimuli in an operant positive reinforcement conditioning procedure, behavioral audiograms of the great tit (*Parus major*) and its principal avian predator, the European sparrowhawk (*Accipiter nisus*), were determined (Klump et al. 1986). The hawk was 6.5 dB more sensitive than the tit at 2 kHz—the best frequency of both species. Although the high-frequency cutoff was very similar in both species, at 8 kHz, the great tit was about 30 dB more sensitive than the sparrowhawk. Figure 5.15 shows the differences in the unmasked thresholds between the sparrowhawk and the great tit at various frequencies. When confronted by a European sparrowhawk, the great tit uses three different vocalizations: (a) the mobbing call (dominant frequency: 4.5 kHz), (b) the scolding call (dominant frequency: 4 kHz), and (c) the “seeet” call (dominant frequency: 8 kHz). The latter call is mainly used by the great tit when it detects a sparrowhawk flying at some distance (Klump et al. 1986) and is the aerial predator call described by Marler (1955). At the dominant frequency of the “seeet” call, the hearing of the great tit is 30 dB more sensitive than that of the European sparrowhawk. This example illustrates that prey species may warn other potential prey of an impending predator by exploiting the mismatch between predator and prey auditory sensitivities.

5.6.4 Conclusions

In this chapter, the concept of matched filtering for detecting desired signals buried in noisy measurement signals is presented. It is shown that the matched filter is another name for the correlation detector or replica-correlation detector, which exploits prior knowledge in the form of an exemplar (or replica) of the desired signal. An example detection problem is used to demonstrate the matched filtering

approach. The matched filter can be an effective detection tool when exemplars of the desired signal are available a priori. In the second part, several key examples of matched filters in the auditory systems of several selected sub-mammalian vertebrates are provided. These auditory systems implement matched filters by sculpting the receiver characteristics to the spectral and temporal features of the species-specific signals of biological importance. With the examples provided, it is hoped that the reader will more fully appreciate the adaptive value of the matched filter concept for reducing the effective noise and thus maximizing the signal-to-noise ratio.

References

- Arch VS, Grafe TU, Narins PM (2008) Ultrasonic signaling by a Bornean frog. *Biol Lett* 4:19–22
- Arch VS, Grafe TU, Gridi-Papp M, Narins PM (2009) Pure ultrasonic communication in an endemic Bornean frog. *PLoS One* 4:e5413. doi:10.1371/journal.pone.0005413
- Arch VS, Simmons DD, Quiñones PM, Feng AS, Jiang J, Stuart B, Shen J-X, Blair C, Narins PM (2012) Inner ear morphological correlates of ultrasonic hearing in frogs. *Hear Res* 283:70–79
- Bradbury JW, Vehrencamp SL (2011) Principles of animal communication, 2nd edn. Sinauer Associates, Sunderland
- Cai H-X, Che J, Pang J-F, Zhou E-M, Zhang Y-P (2007) Paraphyly of Chinese *Amolops* (Anura, Ranidae) and phylogenetic position of the rare Chinese frog, *Amolops tormotus*. *Zootaxa* 1531:49–55
- Candy JV (2006) Model-based signal processing. IEEE Press/Wiley-Interscience, Hoboken, New Jersey, USA
- Capranica RR, Moffat AJM (1975) Selectivity of the peripheral auditory system of spadefoot toads (*Scaphiopus couchi*) for sounds of biological significance. *J Comp Physiol* 100:231–249
- Capranica RR, Moffat AJM (1983) Neurobehavioral correlates of sound communication in amphibians. In: Ewert JP, Capranica RR, Ingle DJ (eds) *Advances in vertebrate neuroethology*. Plenum Press, New York, pp 701–730
- Capranica RR, Rose G (1983) Frequency and temporal processing in the auditory system of anurans. In: Huber F, Markl H (eds) *Neuroethology and behavioral physiology*. Springer, Berlin, pp 136–152
- Clark GA (1999) The revelations of acoustic waves. Science and technology review. Lawrence Livermore National Laboratory, UCRL-52000-99-5, U.S. Government Printing Office 1999/783-046-80013
- Clark GA, Rodgers PW (1981) Adaptive prediction applied to seismic event detection. *Proc IEEE* 69(9):1166–1168
- Clark GA, Axelrod MC, Scott DD (1999) Classification of heart valve sounds from experiments in an anechoic water tank. Lawrence Livermore National Laboratory report UCRL-JC-134634
- Clark GA, Sengupta SK, Aimonetti WD, Roeske JG, Donetti JG (2000) Multispectral image feature selection for land mine detection. *IEEE Trans Geosci Remote Sens* 38(1):304–311
- Clark GA, Robbins CL, Wade KA, Sousa PR (2009) Cable damage detection system and algorithms using time domain reflectometry. Lawrence Livermore National Laboratory report LLNL-TR-576314
- Devijver PA, Kittler J (1987) Pattern recognition theory and applications. In: Jain AK (ed) *Advances in statistical pattern recognition*. Springer, Berlin
- Duda RO, Hart PE (1973) *Pattern classification and scene analysis*. Wiley, New York
- Duda RO, Hart PE, Stork DG (2001) *Pattern classification*, 2nd edn. Wiley, New York
- Fay RR (1988) *Hearing in vertebrates: a psychophysics databook*. Hill-Fay Associates, Winnetka

- Feng AS, Narins PM (2008) Ultrasonic communication in concave-eared torrent frogs (*Amolops tormotus*). *J Comp Physiol* 194:159–167
- Feng AS, Narins PM, Xu C-H, Lin W-Y, Yu Z-L, Qiu Q, Xu Z-M, Shen J-X (2006) Ultrasonic communication in frogs. *Nature* 440:333–336
- Frishkopf LS, Capranica RR, Goldstein MH (1968) Neural coding in the bullfrog's auditory system – a teleological approach. *Proc IEEE* 56:969–980
- Hand DJ (1981) Discrimination and classification. Wiley, New York, p 218
- Heiligenberg W (1991) Neural nets in electric fish. MIT Press, Cambridge
- Hogg RV, Craig AT (1978) Introduction to mathematical statistics. Macmillan, New York
- Hopkins CD (1976) Stimulus filtering and electroreception: tuberous electroreceptors in three species of gymnotoid fish. *J Comp Physiol* 111:171–207
- Inger RF (1966) The systematics and zoogeography of the Amphibia of Borneo, vol 52, Fieldiana: Zoology. Field Museum of Natural History, Chicago
- Jazwinski A (1970) Stochastic processes and filtering theory. Academic, New York
- Kay SM (1998) Fundamentals of statistical signal processing, vol II, Detection theory. Prentice Hall, Upper Saddle River
- Klump GM, Kretzschmar E, Curio E (1986) The hearing of an avian predator and its avian prey. *Behav Ecol Sociobiol* 18:317–323
- Liu W-R, Shen J-X, Zhang Y-J, Xu Z-M, Qi Z, Xue M-Q (2014) Auditory sexual difference in the large odorous frog *Odorrana graminea*. *J Comp Physiol* 200:311–316
- Loftus-Hills JJ, Johnstone BM (1970) Auditory function, communication, and the brain-evoked response in anuran amphibians. *J Acoust Soc Am* 47:1131–1138
- Marler P (1955) Characteristics of some animal calls. *Nature* 176:6–8
- Meenderink SWF, Kits M, Narins PM (2010) Frequency matching of vocalizations to inner-ear sensitivity along an altitudinal gradient in the coqui frog. *Biol Lett* 6:278–281
- Meyer JH, Zakon HH (1982) Androgens alter the tuning of electroreceptors. *Science* 217:635–637
- Moreno-Gómez FN, Sueur J, Soto-Gamboa M, Penna M (2013) Female frog auditory sensitivity, male calls, and background noise: potential influences on the evolution of a peculiar matched filter. *Biol J Linn Soc* 110:814–827
- Narins PM, Capranica RR (1976) Sexual differences in the auditory system of the tree frog, *Eleutherodactylus coqui*. *Science* 192:378–380
- Narins PM, Capranica RR (1978) Communicative significance of the two – note call of the treefrog *Eleutherodactylus coqui*. *J Comp Physiol* 127:1–9
- Narins PM, Smith SL (1986) Clinal variation in anuran advertisement calls: basis for acoustic isolation? *Behav Ecol Sociobiol* 19:135–141
- Narins PM, Feng AS, Schnitzler H-U, Denzinger A, Suthers RA, Lin W, Xu C-H (2004) Old World frog and bird vocalizations contain prominent ultrasonic harmonics. *J Acoust Soc Am* 115:910–913
- Oppenheim AV, Schaffer RW (1975) Digital signal processing. Prentice-Hall, Englewood Cliff
- Papoulis A (1965) Probability, random variables, and stochastic processes. McGraw-Hill, New York
- Sales G, Pye D (1974) Ultrasonic communication by animals. Chapman and Hall, London
- Stuart BL (2008) The phylogenetic problem of *Huia* (Amphibia: Ranidae). *Mol Phylogenet Evol* 46:49–60
- Van Trees HL (1968) Detection, estimation and modulation theory, part I. Detection, estimation and linear modulation theory. Wiley, New York
- Waltz E, Llinas J (1990) Multisensor data fusion. Artech House, Norwood
- Wehner R (1987) 'Matched filters' – neural models of the external world. *J Comp Physiol* 161:511–531
- Whalen A (1971) Detection of signals in noise. Academic, New York
- Zhou EM, Adler K (1993) Herpetology of China. Society for the Study of Amphibians and Reptiles, Oxford
- Zoubir AM, Iskander DR (2004) Bootstrap techniques for signal processing. Cambridge University Press, Cambridge