Chapter 21

Detection

ch,det

Contents

s,det,intro

21.1 Introduction (s,det,intro)

Classification is the task of assigning an object or event to one of a set of prespecified classes. This is performed by observing one or more **features** of the object or event and deciding which class it belongs to through the evaluation of those features [1]. Medical diagnosis tasks can very often be described as classification problems, *e.g.* classifying a patient as normal or diseased, classifying a condition as mild, moderate or acute, etc. Our focus here is the case where the features used for the purpose of such a classification are a set of tomographic data, or more appropriately an image reconstructed from such data.

We will refer to the entity making the decision as the **observer**. In a clinical setting that would usually be a human, although it may also be a computer program or a combination of the two. However, our ultimate goal in studying observers here is a fast, automated means for adjusting the parameters of image reconstruction methods prior to the reconstruction so as to improve performance in classification tasks. Therefore, we turn our attention to computer (numerical) observers, which are simpler to study analytically. Of course, we would like them to perform in a manner correlated to human observer performance, if our results are to be more broadly applicable.

This *chapter* analyzes the performance of numerical observers in signal detection tasks for the purpose of trying to develop regularization methods that optimize this performance. (Much of this was published in [2, 3].)

21.2 The detection task

Let $\boldsymbol{f}$ be the true object being imaged (or an approximation of the true object in $\mathbb{R}^{n_\mathrm{p}}$). To express our uncertainty about the object, we allow it to be a random field. In emission tomography this uncertainty stems from the variability in patient physiology and radiotracer uptake.

Assume that the true object $\boldsymbol{f}$ belongs to exactly one of the classes $C_i, i = 0, 1, \ldots, L-1$ and H_i is the hypothesis that $\boldsymbol{f}$ belongs to class C_i. We consider an observer that has to decide among the hypotheses $H_i, i = 0, 1, \ldots, L-1$ based on an observed feature vector $\boldsymbol{v}$. The decision rules we focus on are deterministic, that is the observer must make the same decision every time it is provided with the same $\boldsymbol{v}$. Furthermore, the decision has to be on exactly one of the H_i's. Classification then corresponds to a partition of the observation space into non-overlapping regions, each corresponding to one of the H_i's [1, 4]. The decision rule can be represented as a comparison of a set of **test statistics** t_i, each of which is a function of the feature vector $\boldsymbol{v}$, $t_i = g_i(\boldsymbol{v}), i = 0, 1, \ldots, L-1$. The $g_i(\cdot)$'s are called **discriminant functions**. The decision rule is then:

$$\text{Decide } H_i \text{ if } g_i(\boldsymbol{v}) > g_j(\boldsymbol{v}) \text{ for all } j \neq i.$$

In the binary-hypothesis case where $L = 2$, this reduces to comparing a single test statistic $t = g(\boldsymbol{v})$ to some **threshold** T, where T is independent of $\boldsymbol{v}$:

$$\text{Decide } H_1 \text{ if } g(\boldsymbol{v}) > T, \text{ otherwise decide } H_0.$$

In the following we focus on the binary-hypothesis problem, since our main interest is optimizing reconstruction methods with respect to performance in the task of detecting a signal of interest in the object. We therefore define the task as a decision between the **signal present** hypothesis H_1 and the **signal absent** hypothesis H_0. Let $\boldsymbol{f}_s$ be the **target signal**, which we assume to be highly localized in space (*e.g.*, a lesion). Thus the term "signal" does not refer here to the entire object $\boldsymbol{f}$. When the signal $\boldsymbol{f}_s$ is contained in $\boldsymbol{f}$, then we define $\boldsymbol{f}_b \triangleq \boldsymbol{f} - \boldsymbol{f}_s$ as the object background. Otherwise, $\boldsymbol{f}$ consists only of the background. The detection task at hand is thus to determine between the following pair of hypotheses:

$$\begin{aligned} H_0 &: \boldsymbol{f} = \boldsymbol{f}_b && \text{(signal absent)} \\ H_1 &: \boldsymbol{f} = \boldsymbol{f}_b + \boldsymbol{f}_s && \text{(signal present).} \end{aligned} \qquad (21.2.1)$$

e,hypo

In emission tomography, where the target signal is typically a region of higher radioactivity concentration than the normal (background) region of an organ, the additive model $\boldsymbol{f} = \boldsymbol{f}_b + \boldsymbol{f}_s$ is indeed a reasonable one [5].

Let $\boldsymbol{y} \in \mathbb{R}^{n_\mathrm{d}}$ be the observed tomographic data. Then we can write without loss of generality

e,raw

$$\boldsymbol{y} = \boldsymbol{\mathcal{A}} \boldsymbol{f} + \boldsymbol{\varepsilon}, \qquad (21.2.2)$$

where $\boldsymbol{\mathcal{A}}$ is a linear operator modeling the imaging system (considered known) and $\boldsymbol{\varepsilon} \in \mathbb{R}^{n_\mathrm{d}}$ is a random vector representing imaging noise, which may or may not be dependent on $\boldsymbol{f}$. If $\boldsymbol{\varepsilon}$ is independent of $\boldsymbol{f}$, then we have the additive noise problem. If it is not, then we could still simply define $\boldsymbol{\varepsilon}$ as $\boldsymbol{\varepsilon} \triangleq \boldsymbol{y} - \boldsymbol{\mathcal{A}} \boldsymbol{f}$. For a given instance of the object $\boldsymbol{f}$, *i.e.*, a given instance of patient physiology and radiotracer uptake, the measurement $\boldsymbol{y}$ is random due to variability inherent in the imaging system. In emission tomography the conditional probability distribution of $\boldsymbol{y}$ given $\boldsymbol{f}$ is Poisson as described in Chapter 8.

Depending on the model we adopt for the observer, the feature vector $\boldsymbol{v}$ on which the observer's decisions are based may be the tomographic data, an image of the object reconstructed from the data, or the output of a set of filtering operations on the image, as we will see later. Observers applied on the tomographic data $\boldsymbol{y}$ would be very likely to grossly overestimate human observer performance, since humans have difficulty distinguishing small features in a sinogram. Observers that are applied on reconstructed images are more realistic, since this situation better corresponds to common imaging practice. An image reconstruction method is a mapping of the measurements $\boldsymbol{y}$ into an estimated image $\hat{\boldsymbol{f}} \in \mathbb{R}^{n_p}$ and therefore an observer applied on the reconstructed image would use a test statistic of the form $t = t(\hat{\boldsymbol{f}}) = t(\hat{\boldsymbol{f}}(\boldsymbol{y}))$. In the following, we will denote discriminants that are applied on the reconstructed image with a hat.

21.3 Figures of merit

One can quantify the detection performance of an observer by tracing its **Receiver Operating Characteristic** (ROC) curve, a plot of the probability of a **true positive** (deciding that H_1 is true when H_1 is actually true) versus the probability of a **false positive** (deciding that H_1 is true when H_0 is actually true). The trade-off between these two probabilities can be plotted by varying the decision threshold T. A very common figure of merit for observers is the **Area Under the Curve** (AUC).

Another figure of merit is the **Signal-to-Noise Ratio** (SNR), defined as

$$\text{SNR} = \frac{\mathsf{E}[t|H_1] - \mathsf{E}[t|H_0]}{\sqrt{p_1 \operatorname{Var}\{t|H_1\} + p_0 \operatorname{Var}\{t|H_0\}}}, \tag{21.3.1}$$

e,obs,snr

where p_i is the **a priori** probability of the hypothesis H_i, and $\mathsf{E}[\cdot|H_i]$ and $\operatorname{Var}\{\cdot|H_i\}$ denote the expected value and variance respectively of some random variable under hypothesis H_i. In the case where the test statistic t is Gaussian-distributed under both hypotheses, the SNR has a special significance, since it can be used to directly calculate the AUC [5, p.819]:

$$\text{AUC} = \frac{1}{2}\left[1 + \operatorname{erf}\left(\frac{\text{SNR}}{2}\right)\right]. \tag{21.3.2}$$

e,obs,auc,snr

By inverting (21.3.2), one can also define the **detectability index** d_A as

$$d_A = 2\operatorname{erf}^{-1}(2(\text{AUC}) - 1).$$

Obviously, $\text{SNR} = d_A$ when t is Gaussian-distributed under both hypotheses.

21.4 The ideal observer

The **ideal observer** uses all available information on the feature vector to make a decision that minimizes an average Bayes cost (*e.g.*, average probability of error). It can be shown that the corresponding discriminant function is the well-known **likelihood ratio** [6]. When the feature vector considered is the data $\boldsymbol{y}$, this is

$$g_\circ(\boldsymbol{y}) = \frac{p(\boldsymbol{y}|H_1)}{p(\boldsymbol{y}|H_0)}, \tag{21.4.1}$$

e,obs,ideal

where $p(\cdot|H_i)$ denotes the probability distribution of a random vector under hypothesis H_i. The threshold is equal to

$$T = \frac{(c_{10} - c_{00})p_1}{(c_{01} - c_{11})p_0},$$

where c_{ij} is the Bayes cost associated with deciding on H_i when the truth is H_j. Conducting this test requires full knowledge of the statistics of both the object and the imaging system, as well as the a priori probabilities of the hypotheses, referred to in medical diagnosis as **disease prevalence**. In practice, this much information is never known, so the ideal observer serves as a golden rule against which the detection performance of other observers can be compared.

The ideal observer for feature vectors in the reconstructed image space uses the discriminant function

$$t_\circ(\boldsymbol{y}) = \frac{p(\hat{\boldsymbol{f}}(\boldsymbol{y})|H_1)}{p(\hat{\boldsymbol{f}}(\boldsymbol{y})|H_0)}. \tag{21.4.2}$$

e,obs,ideal,rec

However, the ideal observer is invariant to nonsingular transformations of the data. Thus the performance of the observer with discriminant (21.4.2) would only be inferior to the one with discriminant (21.4.1) if the image reconstruction method discarded some of the information in the data by performing a noninvertible transformation. Since a reasonable reconstruction method would not discard information, the ideal observer is not used to evaluate and rank reconstruction methods. Instead, suboptimal observer models such as the ones mentioned below are used for this purpose.

21.5 Linear observers

The ideal discriminants from a Bayesian point of view for several detection problems are of complexity higher than linear. However, researchers have found that human observers do not perform ideally and have proposed observer models that account for human suboptimality [7–10]. As a result, various popular linear observer models exist in the literature, either in their ideal form [11], or with the addition of frequency-selective channels and internal noise [5, §14.2.2].

In general, a **linear discriminant** can be defined as the scalar product of some **template** $\boldsymbol{w} \in \mathbb{R}^{n_{\mathrm{p}}}$ with the observation $\hat{\boldsymbol{f}}$ (for observers applied on the reconstructed image):

e,obs,lin

$$t(\boldsymbol{y}) = \boldsymbol{w}' \hat{\boldsymbol{f}}(\boldsymbol{y}). \tag{21.5.1}$$

We list here some linear observers that are often encountered in the objective image quality literature.

21.5.1 The Hotelling observer

The **Hotelling Observer** (HO) [12] assumes knowledge of the first- and second-order statistics of the target signal $\boldsymbol{f}_s$, the background $\boldsymbol{f}_b$, and the noise $\boldsymbol{\varepsilon}$. This results in the **ideal linear discriminant**, that is the one that achieves maximum SNR:

e,obs,hot

$$\boldsymbol{w}_{\mathrm{HO}} = \boldsymbol{S}_{\hat{\boldsymbol{f}}}^{\dagger} \left(\mathsf{E}\left[\hat{\boldsymbol{f}} | H_1\right] - \mathsf{E}\left[\hat{\boldsymbol{f}} | H_0\right] \right), \tag{21.5.2}$$

where

$$\boldsymbol{S}_{\hat{\boldsymbol{f}}} = p_1 \,\mathsf{Cov}\left\{\hat{\boldsymbol{f}} | H_1\right\} + p_0 \,\mathsf{Cov}\left\{\hat{\boldsymbol{f}} | H_0\right\}$$

is the unconditional covariance of $\hat{\boldsymbol{f}}$, equivalent to the **intra-class scatter matrix** in the pattern classification literature, the superscript "†" denotes a **pseudo-inverse**, and $\mathsf{E}[\cdot|H_i]$ and $\mathsf{Cov}\{\cdot|H_i\}$ denote the mean and covariance respectively of some random vector under hypothesis H_i. No linear transformation of the data can improve the SNR of this ideal linear observer, just like no transformation of the data can improve the performance of the ideal observer.

21.5.2 The prewhitening observer

The **PreWhitening Observer** (PW observer) assumes knowledge of the second-order statistics of the background $\boldsymbol{f}_b$ and noise $\boldsymbol{\varepsilon}$ but not necessarily those of the target signal $\boldsymbol{f}_s$. The corresponding template is

e,obs,pw

$$\boldsymbol{w}_{\mathrm{PW}} = \mathsf{Cov}\left\{\hat{\boldsymbol{f}} | H_0\right\}^{\dagger} \left(\mathsf{E}\left[\hat{\boldsymbol{f}} | H_1\right] - \mathsf{E}\left[\hat{\boldsymbol{f}} | H_0\right] \right). \tag{21.5.3}$$

If the reconstruction method is linear and the target signal $\boldsymbol{f}_s$ is deterministic, then the ideal linear discriminant (21.5.2) reduces to the PW discriminant.

21.5.3 The non-prewhitening observer

The **Non-PreWhitening Observer** (NPW observer) assumes knowledge only of the first-order statistics of the target signal $\boldsymbol{f}_s$, the background $\boldsymbol{f}_b$, and the noise $\boldsymbol{\varepsilon}$. The corresponding template is

e,obs,npw

$$\boldsymbol{w}_{\mathrm{NPW}} = \mathsf{E}\left[\hat{\boldsymbol{f}} | H_1\right] - \mathsf{E}\left[\hat{\boldsymbol{f}} | H_0\right]. \tag{21.5.4}$$

If the reconstruction method is linear, the target signal $\boldsymbol{f}_s$ is deterministic and the background $\boldsymbol{f}_b$ and noise $\boldsymbol{\varepsilon}$ are white and independent, then the ideal linear discriminant (21.5.2) reduces to the NPW discriminant.

21.5.4 The region-of-interest observer

The **Region-of-Interest observer** (ROI observer) assumes knowledge only of the first-order statistics of the target signal $\boldsymbol{f}_s$. The corresponding template is that of a simple **matched filter**:

e,obs,roi

$$\boldsymbol{w}_{\mathrm{ROI}} = \mathsf{E}[\boldsymbol{f}_s | H_1]. \tag{21.5.5}$$

21.5.5 Channelized linear observers

Channelized observers include a set of frequency-selective channels in an attempt either to construct an efficient basis for the approximation of the ideal observer [13], or to model the frequency selectivity that is believed to characterize human visual perception [10]. Here we are interested primarily in the latter type of channel, since suboptimal observers are the focus for the purpose of image reconstruction optimization. However, our analysis applies to either channel flavor.

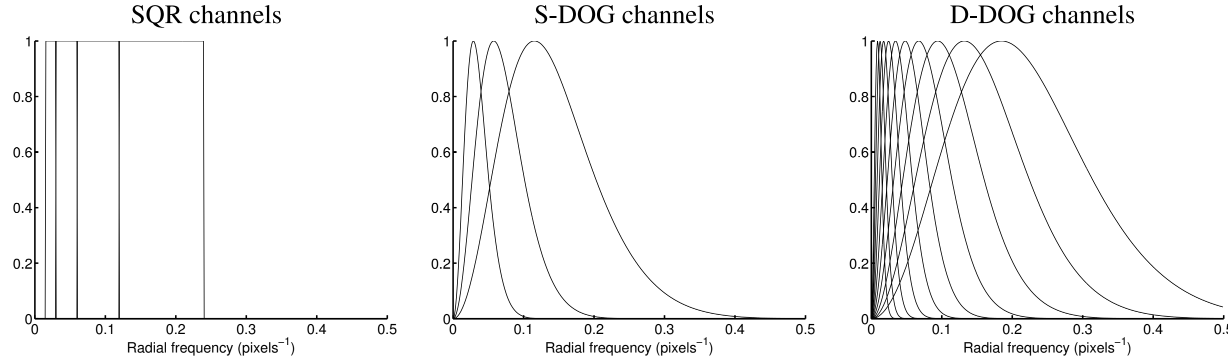

Figure 21.5.1: Profiles of CHO channel frequency responses for the three channel sets investigated in [17]: Square (SQR), Sparse Difference-of-Gaussians (S-DOG), and Dense Difference-of-Gaussians (D-DOG). fig:channels

Conceptually, channelized observers first pass the reconstructed image $\hat{\boldsymbol{f}}$ through a set of M bandpass filters. The new feature vector $\hat{\boldsymbol{c}} \in \mathbb{C}^M$ is formed from the values of the filter outputs at the known location of the target signal center and can include additive noise:

$$\hat{\boldsymbol{c}}(\boldsymbol{y}) = \boldsymbol{\mathcal{C}}' \hat{\boldsymbol{f}}(\boldsymbol{y}) + \varepsilon_{\text{int}},$$

where $\boldsymbol{\mathcal{C}}$ is an $1 \times M$ collection of operators. The mth of these operators applies the impulse response of the mth bandpass filter and samples the output at the center of the target signal. Typically this filtering step is not invertible and it greatly reduces the dimensionality of the detection problem (*e.g.*, $M = 4$ in [14]). Furthermore, the bandpass filters involved in $\boldsymbol{\mathcal{C}}$ typically correspond to distinct frequency bands, in which case the covariance of $\hat{\boldsymbol{c}}$ can be assumed to be nonsingular. Nevertheless, we will use pseudo-inverses in the interest of generality. The **internal noise** vector $\boldsymbol{\varepsilon}_{\text{int}}$ models inherent uncertainty in the observer's decisions and is zero-mean Gaussian with covariance matrix $\boldsymbol{\Pi}_{\text{int}}$.

A generic channelized linear observer forms its test statistic t_{ch} by applying a template $\boldsymbol{w} \in \mathbb{C}^M$ to the output of the filter bank: e,obs,ch

$$t_{\text{ch}}(\boldsymbol{y}) = \boldsymbol{w}' \hat{\boldsymbol{c}}(\boldsymbol{y}). \tag{21.5.6}$$

21.5.5.1 The channelized Hotelling observer

The **Channelized Hotelling Observer** (CHO) has been shown to be particularly successful in predicting human observer performance in several detection tasks [14–17]. It applies the ideal linear discriminant with respect to the output $\hat{\boldsymbol{c}}$ of the M-channel filter-bank. This corresponds to the template e,obs,cho

$$\boldsymbol{w}_{\text{CHO}} = \boldsymbol{S}_{\hat{\boldsymbol{c}}}^{\dagger} \Big(\mathsf{E}[\hat{\boldsymbol{c}}|H_1] - \mathsf{E}[\hat{\boldsymbol{c}}|H_0] \Big), \tag{21.5.7}$$

where

$$\boldsymbol{S}_{\hat{\boldsymbol{c}}} = p_1 \,\mathsf{Cov}\{\hat{\boldsymbol{c}}|H_1\} + p_0 \,\mathsf{Cov}\{\hat{\boldsymbol{c}}|H_0\}$$

is the unconditional covariance matrix of $\hat{\boldsymbol{c}}$. Examples of channel sets that have been used with the CHO are the Square (SQR), Sparse Difference-of-Gaussians (S-DOG), and Dense Difference-of-Gaussians (D-DOG) channels defined in [17]. Profiles through the frequency responses of these channels are shown in figure 21.5.1.

s,det,gau 21.6 Optimal reconstruction for signal detection (s,det,gau)

To facilitate analysis, we consider here linear reconstructors. Many common tomographic reconstruction techniques can be approximated as linear, except maybe when enforcing a nonnegativity constraint [18]. Here we will assume that the target signal appears on a background that is sufficiently high to render any such nonnegativity constraint inactive around the signal location. We denote a generic linear reconstructor by an operator $\boldsymbol{\mathcal{Z}}$. The reconstructed image is then given by e,det,recon,linear

$$\hat{\boldsymbol{f}} = \hat{\boldsymbol{f}}(\boldsymbol{y}) = \boldsymbol{\mathcal{Z}} \boldsymbol{y}. \tag{21.6.1}$$

We may view the reconstruction $\hat{\boldsymbol{f}}$ either as a vector in a Hilbert space, in which case $\boldsymbol{\mathcal{Z}}$ is a general linear mapping from $\mathbb{R}^{n_{\text{d}}}$ to that Hilbert space, or as a discrete representation in $\mathbb{R}^{n_{\text{p}}}$, in which case $\boldsymbol{\mathcal{Z}}$ is a matrix in $\mathbb{R}^{n_{\text{p}} \times n_{\text{d}}}$. (We could also think of $\boldsymbol{\mathcal{Z}} = \boldsymbol{\mathcal{I}}$, the identity operator, for detection directly from the raw measurements.)

Our goal is to optimize the reconstructor $\boldsymbol{\mathcal{Z}}$ with respect to the performance of various observers of interest in the detection of $\boldsymbol{f}_s$.

21.6.1 Assumptions of Gaussianity

In subsequent sections, we will assume Gaussian-distributed test statistics and focus on maximizing the SNR, in which case the AUC is also maximized. For linear observers and linear reconstructors in particular, the test statistic is a weighted sum of the measurements, so it can be approximated as Gaussian by the central limit theorem. Furthermore, since we focus on observers that are applied to the reconstructed image rather than the raw data, we are mainly interested in the statistics of the reconstruction rather than the observation. The statistics of an image reconstructed from Poisson data through a penalized-likelihood method can be described approximately by a Gaussian distribution [19].

We denote the expectations of the background $\boldsymbol{f}_b$ and the signal $\boldsymbol{f}_s$ by $\bar{\boldsymbol{f}}_b$ and $\bar{\boldsymbol{f}}_s$ respectively. We denote their covariances by $\boldsymbol{\mathcal{K}}_b$ and $\boldsymbol{\mathcal{K}}_s$ respectively. In the general case where both background and signal are random, $\boldsymbol{\mathcal{K}}_b$ and $\boldsymbol{\mathcal{K}}_s$ are positive definite. In the special case known as the **signal known exactly** (SKE) detection task, we have $\boldsymbol{\mathcal{K}}_s = \mathbf{0}$ and thus a deterministic signal $\boldsymbol{f}_s = \bar{\boldsymbol{f}}_s$. Similarly, in the **background known exactly** (BKE) task, we have $\boldsymbol{\mathcal{K}}_b = \mathbf{0}$ and thus a deterministic background $\boldsymbol{f}_b = \bar{\boldsymbol{f}}_b$. In all cases, we assume that $\boldsymbol{\mathcal{K}}_b$ and $\boldsymbol{\mathcal{K}}_s$ are known to the observer. We also take the a priori probabilities of the hypotheses H_1 and H_0 to be known and equal to p_1 and $p_0 = 1 - p_1$ respectively.

Finally, we will assume that the measurement noise ε is zero-mean with a known covariance matrix that is independent of whether the target signal is present or not. These assumptions are less restrictive than may first appear. In the case of emission tomography, the measurements $\boldsymbol{y}$ are independent and Poisson-distributed conditional on the object $\boldsymbol{f}$. The conditional mean and covariance of the measurement vector are, respectively,

$$\begin{aligned} \mathsf{E}[\boldsymbol{Y}|\boldsymbol{f}] &= \boldsymbol{\mathcal{A}}\boldsymbol{f} + \boldsymbol{r} \\ \mathsf{Cov}\{\boldsymbol{Y}|\boldsymbol{f}\} &= \mathrm{diag}\{\boldsymbol{\mathcal{A}}\boldsymbol{f} + \boldsymbol{r}\}, \end{aligned}$$

for some vector $\boldsymbol{r} \in \mathbb{R}^{n_\mathrm{d}}$ that represents scatter and/or random coincidences and is assumed to be deterministic and known. For hypothesis $H_i, i = 0, 1$, the data moments are given by:

e,det,m,hi

$$\mathsf{E}[\boldsymbol{Y}|H_i] = \mathsf{E}\left[\mathsf{E}[\boldsymbol{Y}|\boldsymbol{f}, H_i]\,\middle|H_i\right] = \boldsymbol{\mathcal{A}}\,\mathsf{E}[\boldsymbol{f}|H_i] + \boldsymbol{r} \tag{21.6.2}$$

e,det,cov,hi

$$\begin{aligned} \mathsf{Cov}\{\boldsymbol{Y}|H_i\} &= \mathsf{E}\left[\mathsf{Cov}\{\boldsymbol{Y}|\boldsymbol{f}, H_i\}\,\middle|H_i\right] + \mathsf{Cov}\left\{\mathsf{E}[\boldsymbol{Y}|\boldsymbol{f}, H_i]\,\middle|H_i\right\} \\ &= \mathrm{diag}\{\boldsymbol{\mathcal{A}}\,\mathsf{E}[\boldsymbol{f}|H_i] + \boldsymbol{r}\} + \boldsymbol{\mathcal{A}}\,\mathsf{Cov}\{\boldsymbol{f}|H_i\}\,\boldsymbol{\mathcal{A}}'. \end{aligned} \tag{21.6.3}$$

We assume that $\boldsymbol{f}_b$ and $\boldsymbol{f}_s$ are independent. Applying (21.6.2) and (21.6.3) to each of the two hypotheses in (21.2.1) yields

e,det,m,h0,h1

$$\mathsf{E}[\boldsymbol{Y}|H_1] = \mathsf{E}[\boldsymbol{Y}|H_0] + \boldsymbol{\mathcal{A}}\bar{\boldsymbol{f}}_s \tag{21.6.4}$$

e,det,cov,h0,h1

$$\begin{aligned} \mathsf{Cov}\{\boldsymbol{Y}|H_1\} &= \mathsf{Cov}\{\boldsymbol{Y}|H_0\} + \mathrm{diag}\{\boldsymbol{\mathcal{A}}\bar{\boldsymbol{f}}_s\} + \boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{K}}_s\boldsymbol{\mathcal{A}}' \\ &\approx \mathsf{Cov}\{\boldsymbol{Y}|H_0\} + \boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{K}}_s\boldsymbol{\mathcal{A}}'. \end{aligned} \tag{21.6.5}$$

The approximation in the last step is reasonable, since it is safe to assume that the counts from the target signal will be few compared to the counts from the background of the organ. From (21.6.4) and (21.6.5), the problem can be viewed as one with additive noise that is independent of the target signal $\boldsymbol{f}_s$.

In general, problems with Poisson measurement statistics and nonlinear estimators can be analyzed using local shift invariance approximations [19–21] and linearizations [18]. These local approximations and linearizations yield expressions that correspond to the case of Gaussian noise and linear estimators, provided one uses appropriate (background signal dependent) covariance matrices.

s,det,ideal

21.6.2 Ideal observer (s,det,ideal)

Assuming Gaussian reconstruction statistics and a linear reconstructor $\boldsymbol{\mathcal{Z}}$, the distribution of $\hat{\boldsymbol{f}}$ under each hypothesis in (21.2.1) is

$$\begin{aligned} H_0 &: \hat{\boldsymbol{f}} \sim \mathcal{N}(\boldsymbol{\mathcal{Z}}\boldsymbol{\mathcal{A}}\bar{\boldsymbol{f}}_b, \boldsymbol{\mathcal{Z}}\boldsymbol{\Pi}_0\boldsymbol{\mathcal{Z}}') \\ H_1 &: \hat{\boldsymbol{f}} \sim \mathcal{N}(\boldsymbol{\mathcal{Z}}\boldsymbol{\mathcal{A}}(\bar{\boldsymbol{f}}_b + \bar{\boldsymbol{f}}_s), \boldsymbol{\mathcal{Z}}[\boldsymbol{\Pi}_0 + \boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{K}}_s\boldsymbol{\mathcal{A}}']\boldsymbol{\mathcal{Z}}'), \end{aligned}$$

where $\boldsymbol{\Pi}_0 \triangleq \boldsymbol{\Pi} + \boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{K}}_b\boldsymbol{\mathcal{A}}'$ is the covariance of the data under the signal-absent hypothesis and $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\mathcal{K}})$ denotes the Gaussian probability law with mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\mathcal{K}}$. We use "$'$" to denote the adjoint of an operator or equivalently the complex transpose of a matrix.

Substituting these Gaussian likelihoods into the likelihood ratio (21.4.2), yields the discriminant function of the ideal observer that uses $\hat{\boldsymbol{f}}$ as a feature vector:

$$t_\circ(\hat{\boldsymbol{f}}) = \frac{1}{2}\hat{\boldsymbol{f}}'\left(\mathsf{Cov}\left\{\hat{\boldsymbol{f}}|H_0\right\}^\dagger - \mathsf{Cov}\left\{\hat{\boldsymbol{f}}|H_1\right\}^\dagger\right)\hat{\boldsymbol{f}} + \left(\mathsf{Cov}\left\{\hat{\boldsymbol{f}}|H_1\right\}^\dagger \mathsf{E}\left[\hat{\boldsymbol{f}}|H_1\right] - \mathsf{Cov}\left\{\hat{\boldsymbol{f}}|H_0\right\}^\dagger \mathsf{E}\left[\hat{\boldsymbol{f}}|H_0\right]\right)'\hat{\boldsymbol{f}}. \tag{21.6.6}$$

e,det,tstat,ideal,quad

In general, this test statistic is quadratic in $\hat{\boldsymbol{f}}$ and thus not Gaussian-distributed (except in the SKE case, when the observation covariance is equal under both hypotheses and the quadratic term vanishes). Therefore, one would have to optimize the AUC directly rather than work with the SNR. However, the ideal observer is generally not used to evaluate reconstruction methods, since its performance is invariant to any nonsingular data transformation [4, §10.1]. The observer models commonly used in the literature to evaluate image reconstruction methods are linear.

21.6.2.1 Ideal observer for the SKE/BKE task

In the SKE/BKE case we have

$$\mathsf{Cov}\left\{\hat{\boldsymbol{f}}|H_1\right\} = \mathsf{Cov}\left\{\hat{\boldsymbol{f}}|H_0\right\} = \boldsymbol{\mathcal{Z}}\boldsymbol{\Pi}\boldsymbol{\mathcal{Z}}'$$

and so the quadratic term in (21.6.6) is eliminated and the ideal discriminant simplifies to a linear form:

e,det,tstat,ideal

$$t_\circ(\hat{\boldsymbol{f}}) = (\boldsymbol{\mathcal{Z}}\boldsymbol{\mathcal{A}}\boldsymbol{f}_s)'(\boldsymbol{\mathcal{Z}}\boldsymbol{\Pi}\boldsymbol{\mathcal{Z}}')^\dagger \hat{\boldsymbol{f}}. \tag{21.6.7}$$

This is simply a classical **matched filter** applied to the reconstruction.

The distribution of the test statistic $t = t_\circ(\boldsymbol{y})$ under each hypothesis is

$$H_0 : t \sim \mathcal{N}(\boldsymbol{f}_s'\boldsymbol{\mathcal{A}}'\boldsymbol{\mathcal{Z}}'(\boldsymbol{\mathcal{Z}}\boldsymbol{\Pi}\boldsymbol{\mathcal{Z}}')^\dagger\boldsymbol{\mathcal{Z}}\boldsymbol{\mathcal{A}}\boldsymbol{f}_b, \boldsymbol{f}_s'\boldsymbol{\mathcal{A}}'\boldsymbol{\mathcal{Z}}'(\boldsymbol{\mathcal{Z}}\boldsymbol{\Pi}\boldsymbol{\mathcal{Z}}')^\dagger\boldsymbol{\mathcal{Z}}\boldsymbol{\mathcal{A}}\boldsymbol{f}_s)$$
$$H_1 : t \sim \mathcal{N}(\boldsymbol{f}_s'\boldsymbol{\mathcal{A}}'\boldsymbol{\mathcal{Z}}'(\boldsymbol{\mathcal{Z}}\boldsymbol{\Pi}\boldsymbol{\mathcal{Z}}')^\dagger\boldsymbol{\mathcal{Z}}\boldsymbol{\mathcal{A}}(\boldsymbol{f}_b+\boldsymbol{f}_s), \boldsymbol{f}_s'\boldsymbol{\mathcal{A}}'\boldsymbol{\mathcal{Z}}'(\boldsymbol{\mathcal{Z}}\boldsymbol{\Pi}\boldsymbol{\mathcal{Z}}')^\dagger\boldsymbol{\mathcal{Z}}\boldsymbol{\mathcal{A}}\boldsymbol{f}_s).$$

By substituting the moments above into (21.3.1), we conclude that the SNR of the SKE/BKE ideal observer is

e,det,snr,ideal

$$\mathrm{SNR}_\circ^2 = \boldsymbol{f}_s'\boldsymbol{\mathcal{A}}'\boldsymbol{\mathcal{Z}}'(\boldsymbol{\mathcal{Z}}\boldsymbol{\Pi}\boldsymbol{\mathcal{Z}}')^\dagger\boldsymbol{\mathcal{Z}}\boldsymbol{\mathcal{A}}\boldsymbol{f}_s = t_\circ(\hat{\boldsymbol{f}}_s), \tag{21.6.8}$$

where $\hat{\boldsymbol{f}}_s \triangleq \boldsymbol{\mathcal{Z}}\boldsymbol{\mathcal{A}}\boldsymbol{f}_s$ is the reconstructed sinogram of the target signal.

For an invertible $\boldsymbol{\mathcal{Z}}$, the SNR in (21.6.8) becomes $\mathrm{SNR}_\circ^2 = \boldsymbol{f}_s'\boldsymbol{\mathcal{A}}'\boldsymbol{\Pi}^{-1}\boldsymbol{\mathcal{A}}\boldsymbol{f}_s$, which corresponds to the SNR for direct detection from the sinogram.

s,det,linear

21.6.3 Generic linear observers (s,det,linear)

Linear observer models facilitate analysis and they have been found to approximate the suboptimality of human observers [22, 23]. For a linear discriminant of the form (21.5.1), the distribution of $t = t(\boldsymbol{y})$ under each hypothesis is

$$H_0 : t \sim \mathcal{N}(\boldsymbol{w}'\boldsymbol{\mathcal{Z}}\boldsymbol{\mathcal{A}}\bar{\boldsymbol{f}}_b, \boldsymbol{w}'\boldsymbol{\mathcal{Z}}\boldsymbol{\Pi}_0\boldsymbol{\mathcal{Z}}'\boldsymbol{w})$$
$$H_1 : t \sim \mathcal{N}(\boldsymbol{w}'\boldsymbol{\mathcal{Z}}\boldsymbol{\mathcal{A}}(\bar{\boldsymbol{f}}_b+\bar{\boldsymbol{f}}_s), \boldsymbol{w}'\boldsymbol{\mathcal{Z}}[\boldsymbol{\Pi}_0 + \boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{K}}_s\boldsymbol{\mathcal{A}}']\boldsymbol{\mathcal{Z}}'\boldsymbol{w}).$$

Substituting these moments into (21.3.1) yields the SNR of this general linear observer:

e,det,snr,linear

$$\mathrm{SNR}_{\mathrm{lin}}^2 = \frac{(\boldsymbol{w}'\boldsymbol{\mathcal{Z}}\,\bar{\boldsymbol{y}})^2}{\boldsymbol{w}'\boldsymbol{\mathcal{Z}}\check{\boldsymbol{\Pi}}\boldsymbol{\mathcal{Z}}'\boldsymbol{w}} = \frac{\boldsymbol{w}'\boldsymbol{\mathcal{Z}}(\bar{\boldsymbol{y}}\,\bar{\boldsymbol{y}}')\boldsymbol{\mathcal{Z}}'\boldsymbol{w}}{\boldsymbol{w}'\boldsymbol{\mathcal{Z}}\check{\boldsymbol{\Pi}}\boldsymbol{\mathcal{Z}}'\boldsymbol{w}}, \tag{21.6.9}$$

where $\bar{\boldsymbol{y}} \triangleq \boldsymbol{\mathcal{A}}\bar{\boldsymbol{f}}_s$ is the expected sinogram of the target signal,

e,det,linear,bPich

$$\check{\boldsymbol{\Pi}} \triangleq \boldsymbol{\Pi} + \boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{K}}_f\boldsymbol{\mathcal{A}}' \tag{21.6.10}$$

is the unconditional covariance of the data, and $\boldsymbol{\mathcal{K}}_f \triangleq \boldsymbol{\mathcal{K}}_b + p_1\boldsymbol{\mathcal{K}}_s$ is the unconditional covariance of the object. Since $\boldsymbol{\Pi}$, $\boldsymbol{\mathcal{K}}_b$ and $\boldsymbol{\mathcal{K}}_s$ are positive definite, the ratio in (21.6.9) will be well-defined provided $\boldsymbol{\mathcal{Z}}'\boldsymbol{w}$ is nonzero.

The left-hand side of (21.6.9) has the form of a generalized Rayleigh quotient. This form is maximized with respect to $\boldsymbol{\mathcal{Z}}'\boldsymbol{w}$ when (*e.g.*, see [1, p.120])

e,det,Zw

$$\boldsymbol{\mathcal{Z}}'\boldsymbol{w} \propto \check{\boldsymbol{\Pi}}^{-1}\,\bar{\boldsymbol{y}}\,. \tag{21.6.11}$$

Substituting (21.6.11) into (21.6.9) gives the upper bound of the SNR for any linear observer and any linear reconstructor:

$$\mathrm{SNR}_{\mathrm{lin}}^2 \le \bar{\boldsymbol{y}}'\,\check{\boldsymbol{\Pi}}^{-1}\,\bar{\boldsymbol{y}} = \bar{\boldsymbol{f}}_s'\check{\boldsymbol{\mathcal{F}}}\bar{\boldsymbol{f}}_s \triangleq \mathrm{SNR}_{\mathrm{lin}_\circ}^2,$$

where we define

e,det,Fch

e,det,Fch2

$$\check{\boldsymbol{\mathcal{F}}} \triangleq \boldsymbol{\mathcal{A}}'\check{\boldsymbol{\Pi}}^{-1}\boldsymbol{\mathcal{A}} = \boldsymbol{\mathcal{A}}'(\boldsymbol{\Pi} + \boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{K}}_f\boldsymbol{\mathcal{A}}')^{-1}\boldsymbol{\mathcal{A}} \tag{21.6.12}$$
$$= (\boldsymbol{\mathcal{I}} + \boldsymbol{\mathcal{F}}\boldsymbol{\mathcal{K}}_f)^{-1}\boldsymbol{\mathcal{F}} = \boldsymbol{\mathcal{F}}(\boldsymbol{\mathcal{I}} + \boldsymbol{\mathcal{K}}_f\boldsymbol{\mathcal{F}})^{-1}, \tag{21.6.13}$$

where $\boldsymbol{\mathcal{F}} \triangleq \boldsymbol{\mathcal{A}}'\boldsymbol{\Pi}^{-1}\boldsymbol{\mathcal{A}}$, and $\boldsymbol{\mathcal{I}}$ is the identity operator. From (21.6.12) we derive (21.6.13) using the identity $\boldsymbol{\mathcal{A}}(\boldsymbol{\mathcal{I}} + \boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{A}})^{-1} = (\boldsymbol{\mathcal{I}} + \boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{B}})^{-1}\boldsymbol{\mathcal{A}}$.

A simple combination that satisfies (21.6.11) is $\boldsymbol{\mathcal{Z}} = \boldsymbol{\mathcal{I}}$ (which is not a reconstruction method) and $\boldsymbol{w} = \check{\boldsymbol{\Pi}}^{-1}\,\bar{\boldsymbol{y}}$, which corresponds to the Hotelling observer (see section 21.6.5) for detection in the sinogram rather than the reconstruction domain. However, even when we restrict attention to observers that are applied to reconstructed images, usually there are still many ways to satisfy (21.6.11), as the analysis that follows indicates.

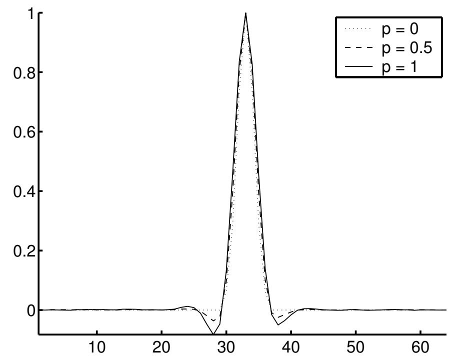

fig:template

Figure 21.6.1: Profiles through the center of (normalized) Fisher observer templates $\boldsymbol{w}_p$ for $p = 0, 0.5, 1$.

s,det,fisher 21.6.4 Fisher observers and Fisher reconstructors (s,det,fisher)

As shown in more detail below, several of the mathematical observers that have been proposed in the literature can achieve the optimal SNR when paired with simple reconstructors that correspond to some power of $\check{\mathcal{F}}$ applied to a backprojection of the data. For lack of a better term, we refer to the family of **Fisher observers**, whose templates have the following form:

e,det,w,fisher

$$\boldsymbol{w}_p \triangleq \check{\mathcal{F}}^p \bar{\boldsymbol{f}}_s \tag{21.6.14}$$

for some $p \in \mathbb{R}$. We allow p to be negative, in which case we interpret $\check{\mathcal{F}}^p = (\check{\mathcal{F}}^\dagger)^{-p}$.

Are there (linear) reconstruction methods that provide optimal detection performance when combined with a Fisher observer? The answer is "yes," and we refer to the corresponding reconstruction methods as **Fisher reconstructors,** defined by

e,det,B,fisher

$$\mathcal{Z}_q \triangleq \check{\mathcal{F}}^q \mathcal{A}' \check{\boldsymbol{\Pi}}^{-1} = \check{\mathcal{F}}^q \mathcal{A}' (\boldsymbol{\Pi} + \mathcal{A} \mathcal{K}_f \mathcal{A}')^{-1} \tag{21.6.15}$$

for some $q \in \mathbb{R}$. The corresponding estimator is

$$\hat{\boldsymbol{f}} = \mathcal{Z}_q \boldsymbol{y} = \check{\mathcal{F}}^q \mathcal{A}' \check{\boldsymbol{\Pi}}^{-1} \boldsymbol{y} = \check{\mathcal{F}}^q \mathcal{A}' (\boldsymbol{\Pi} + \mathcal{A} \mathcal{K}_f \mathcal{A}')^{-1} \boldsymbol{y},$$

which is a kind of weighted backprojection with a (perhaps somewhat unusual) postfilter. (For $q < 0$, this postfilter is something like a deconvolver.) For example, in the case $q = -1$ for the SKE/BKE task ($\check{\mathcal{F}} = \mathcal{F}$ and $\check{\boldsymbol{\Pi}} = \boldsymbol{\Pi}$), the Fisher reconstructor is

e,det,wls

$$\hat{\boldsymbol{f}} = \mathcal{F}^\dagger \mathcal{A}' \boldsymbol{\Pi}^{-1} \boldsymbol{y} \triangleq \hat{\boldsymbol{f}}_{\text{WLS}}, \tag{21.6.16}$$

which is the unregularized **weighted least-squares** (**WLS**) estimator.

For the family of reconstructors (21.6.15) and the family of observers (21.6.14) we have

$$\mathcal{Z}_q' \boldsymbol{w}_p = \check{\boldsymbol{\Pi}}^{-1} \mathcal{A} \check{\mathcal{F}}^q \check{\mathcal{F}}^p \bar{\boldsymbol{f}}_s.$$

By comparing the above with (21.6.11) we see that the choice $p = -q$ leads to the optimal SNR, *i.e.*, *for any Fisher observer, there is a corresponding Fisher reconstructor that achieves the ideal SNR*. Furthermore, these Fisher reconstructors appear to be largely devoid of regularization[1], so it appears that **regularization is not essential** for a large family of linear observers for the detection task at hand. Fig. 21.6.1 shows template profiles for some of these observers. The profile shape for $p = 0.5$ especially is reminiscent of those estimated from human observers (*e.g.*, see [24]). Next we explore some specific observer examples.

s,det,ho 21.6.5 Hotelling observer (s,det,ho)

Substituting the moments of $\hat{\boldsymbol{f}}$ in the Hotelling template (21.5.2) yields

e,obs,hot,Z

$$\boldsymbol{w}_{\text{HO}} = (\mathcal{Z} \check{\boldsymbol{\Pi}} \mathcal{Z}')^\dagger \mathcal{Z} \bar{\boldsymbol{y}} \tag{21.6.17}$$

and thus

$$\mathcal{Z}' \boldsymbol{w}_{\text{HO}} = \mathcal{Z}' (\mathcal{Z} \check{\boldsymbol{\Pi}} \mathcal{Z}')^\dagger \mathcal{Z} \bar{\boldsymbol{y}} = \check{\boldsymbol{\Pi}}^{-1/2} \mathcal{P}_{\check{\boldsymbol{\Pi}}^{1/2} \mathcal{Z}'} (\check{\boldsymbol{\Pi}}^{-1/2} \bar{\boldsymbol{y}}),$$

where $\mathcal{P}_{\check{\boldsymbol{\Pi}}^{1/2} \mathcal{Z}'}(\cdot)$ denotes the orthogonal projection of a vector onto $\mathcal{R}_{\check{\boldsymbol{\Pi}}^{1/2} \mathcal{Z}'}$, the range space of $\check{\boldsymbol{\Pi}}^{1/2} \mathcal{Z}'$. By comparing the above with (21.6.11) we find that the optimal SNR is achieved when $\check{\boldsymbol{\Pi}}^{-1/2} \bar{\boldsymbol{y}} \in \mathcal{R}_{\check{\boldsymbol{\Pi}}^{1/2} \mathcal{Z}'}$. There are a multitude of choices of $\mathcal{Z}$ that satisfy this mild condition. For example, for a $\mathcal{Z}_q$ of the Fisher reconstructor family (21.6.15), the HO template in (21.6.17) becomes $\boldsymbol{w}_{\text{HO}} = \check{\mathcal{F}}^{-q} \bar{\boldsymbol{f}}_s$, which corresponds to the Fisher observer

[1] If $q > -1$, then one could construe the Fisher reconstructor as being marginally regularized since it entails somewhat "less deconvolution" than the WLS reconstructor. However, this type of "regularization" does not improve the condition number in the case of singular $\check{\mathcal{F}}$, and it is unlike most regularization methods described in the literature.

in (21.6.14) with $p = -q$. In other words, any Fisher reconstructor achieves the ideal SNR for the Hotelling observer applied in the image domain. This is consistent with the fact that linear transformations of the data do not affect the performance of the optimal linear observer [4, §10.2] (except when the transformation operator does not have a right inverse, in which case performance degrades).

21.6.6 PW observer

Substituting the moments of $\hat{\boldsymbol{f}}$ in the PW template (21.5.3) yields

$$\boldsymbol{w}_{\mathrm{PW}} = (\boldsymbol{\mathcal{Z}} \boldsymbol{\Pi}_0 \boldsymbol{\mathcal{Z}}')^{\dagger} \boldsymbol{\mathcal{Z}} \,\bar{\boldsymbol{y}}$$

and thus by (21.6.11) the $\boldsymbol{\mathcal{Z}}$ that achieves the optimal SNR must satisfy

$$\boldsymbol{\mathcal{Z}}' (\boldsymbol{\mathcal{Z}} \boldsymbol{\Pi}_0 \boldsymbol{\mathcal{Z}}')^{\dagger} \boldsymbol{\mathcal{Z}} \,\bar{\boldsymbol{y}} \propto \check{\boldsymbol{\Pi}}^{-1} \,\bar{\boldsymbol{y}},$$

or equivalently

$$\mathcal{P}_{\boldsymbol{\Pi}_0^{1/2} \boldsymbol{\mathcal{Z}}'} (\boldsymbol{\Pi}_0^{-1/2} \,\bar{\boldsymbol{y}}) \propto \boldsymbol{\Pi}_0^{1/2} \check{\boldsymbol{\Pi}}^{-1} \,\bar{\boldsymbol{y}},$$

which in turn implies that $\boldsymbol{\Pi}_0^{-1/2} \,\bar{\boldsymbol{y}} - c \boldsymbol{\Pi}_0^{1/2} \check{\boldsymbol{\Pi}}^{-1} \,\bar{\boldsymbol{y}}$ must be orthogonal to $\mathcal{R}_{\boldsymbol{\Pi}_0^{1/2} \boldsymbol{\mathcal{Z}}'}$ for any constant c. This finally leads to the requirement that

$$\boldsymbol{\mathcal{Z}} \,\bar{\boldsymbol{y}} \propto \boldsymbol{\mathcal{Z}} \boldsymbol{\Pi}_0 \check{\boldsymbol{\Pi}}^{-1} \,\bar{\boldsymbol{y}},$$

i.e., that the sinograms $\bar{\boldsymbol{y}}$ and $\boldsymbol{\Pi}_0 \check{\boldsymbol{\Pi}}^{-1} \,\bar{\boldsymbol{y}}$ yield the same reconstructed image but for a scaling constant. Thus, in general there are no linear reconstructors that can achieve the optimal SNR when paired with the PW observer.

An exception to this is the SKE case, where $\check{\boldsymbol{\Pi}} = \boldsymbol{\Pi}_0$ and the PW observer is the same as the HO, so there are infinitely many linear reconstructors that achieve optimal SNR. The minimal dependence of the SNR on $\boldsymbol{\mathcal{Z}}$ for the HO is consistent with the observation of Qi et al. that performance of the PW observer in the SKE task is independent of smoothing method in the MAP case [25].

21.6.7 NPW observer

Substituting the moments of $\hat{\boldsymbol{f}}$ in the NPW template (21.5.4) yields

e,obs,npw,Z

$$\boldsymbol{w}_{\mathrm{NPW}} = \boldsymbol{\mathcal{Z}} \,\bar{\boldsymbol{y}} \tag{21.6.18}$$

and thus by (21.6.11) the optimal reconstructor must satisfy

e,det,npw,Zopt

$$\boldsymbol{\mathcal{Z}}' \boldsymbol{\mathcal{Z}} \,\bar{\boldsymbol{y}} \propto \check{\boldsymbol{\Pi}}^{-1} \,\bar{\boldsymbol{y}}\,. \tag{21.6.19}$$

For a $\boldsymbol{\mathcal{Z}}_q$ of the Fisher reconstructor family (21.6.15), the NPW template becomes $\boldsymbol{w}_{\mathrm{NPW}} = \check{\boldsymbol{\mathcal{F}}}^{q+1} \,\bar{\boldsymbol{y}}$, which corresponds to the Fisher observer in (21.6.14) with $p = q + 1$. Since $q = -1/2$ satisfies $p = -q$, the optimal SNR is achieved by the somewhat unusual Fisher reconstructor

e,det,B,npw

$$\boldsymbol{\mathcal{Z}} = \check{\boldsymbol{\mathcal{F}}}^{-1/2} \boldsymbol{\mathcal{A}}' \check{\boldsymbol{\Pi}}^{-1}, \tag{21.6.20}$$

where $\check{\boldsymbol{\mathcal{F}}}^{-1/2} = (\check{\boldsymbol{\mathcal{F}}}^{\dagger})^{1/2}$ as defined earlier. Whether there are other solutions that satisfy (21.6.19) is an *open problem*. (There is also the choice of $\boldsymbol{\mathcal{Z}} = \check{\boldsymbol{\Pi}}^{-1/2}$, which is not a reconstruction method. It is equivalent to the Hotelling observer for sinogram-based detection.)

The reconstruction method (21.6.20) corresponds to the estimator

$$\hat{\boldsymbol{f}} = \check{\boldsymbol{\mathcal{F}}}^{-1/2} \boldsymbol{\mathcal{A}}' \check{\boldsymbol{\Pi}}^{-1} \boldsymbol{y},$$

or for the SKE/BKE task

$$\hat{\boldsymbol{f}} = \boldsymbol{\mathcal{F}}^{-1/2} \boldsymbol{\mathcal{A}}' \boldsymbol{\Pi}^{-1} \boldsymbol{y} = \boldsymbol{\mathcal{F}}^{1/2} \hat{\boldsymbol{f}}_{\mathrm{WLS}},$$

which is a type of "post-filtered" WLS estimate, with an unusual shift-variant post-filter. This estimator is very impractical for two reasons. Firstly, even if $\boldsymbol{\mathcal{A}}$ happens to have full rank, $\boldsymbol{\mathcal{F}}$ is usually very ill conditioned, so computing the WLS solution $\hat{\boldsymbol{f}}_{\mathrm{WLS}}$ will require a multitude of iterations for any practical iterative algorithm. Secondly, the shift-variant post-filter $\boldsymbol{\mathcal{F}}^{1/2}$ would be very complicated to implement.

21.6.8 ROI observer

From (21.5.5) the template of the ROI observer is given by

$$\boldsymbol{w}_{\mathrm{ROI}} = \bar{\boldsymbol{f}}_s.$$

Then by (21.6.11) the optimal reconstructor must satisfy

$$\boldsymbol{\mathcal{Z}}' \bar{\boldsymbol{f}}_s \propto \check{\boldsymbol{\Pi}}^{-1} \bar{\boldsymbol{y}}.$$

Since the ROI template corresponds to the Fisher observer with $p = 0$, the optimal SNR is achieved by the Fisher reconstructor with $q = 0$, *i.e.*, e,det,B,roi

$$\boldsymbol{\mathcal{Z}} = \boldsymbol{\mathcal{A}}' \check{\boldsymbol{\Pi}}^{-1}. \tag{21.6.21}$$

Curiously, in this case

$$\hat{\boldsymbol{f}} = \boldsymbol{\mathcal{Z}} \boldsymbol{y} = \boldsymbol{\mathcal{A}}' \check{\boldsymbol{\Pi}}^{-1} \boldsymbol{y},$$

which for the SKE/BKE task reduces to $\hat{\boldsymbol{f}} = \boldsymbol{\mathcal{A}}' \boldsymbol{\Pi}^{-1} \boldsymbol{y}$. This is a very blurry estimate of $\boldsymbol{f}$, being simply unfiltered backprojection. Yet for the ROI observer it is optimal, and no amount of deconvolution will improve the SNR for this detection task, which is an indication that the task is too simple.

The optimality of $\boldsymbol{\mathcal{Z}} = \boldsymbol{\mathcal{A}}' \boldsymbol{\Pi}^{-1}$ is consistent with the demonstration in Qi et al. of the ROI observer (for a MAP (aka PWLS) reconstructor with $\boldsymbol{\mathcal{R}} = \beta \boldsymbol{\mathcal{I}}$ in (21.6.29) below) approaching the PW observer's performance as $\beta \to \infty$ [25].

21.6.9 Summary of Fisher observers and reconstructors

For three of the specific observer examples considered above, at least one reconstructor of the Fisher family (21.6.15) was found to achieve the highest SNR possible for linear observers. The following table summarizes these examples.

Observer	q	Best estimator	Interpretation
Hotelling	$\mathbb{R}$	$\check{\boldsymbol{\mathcal{F}}}^q \boldsymbol{\mathcal{A}}' \check{\boldsymbol{\Pi}}^{-1} \boldsymbol{y}$	any Fisher reconstructor (*e.g.*, WLS)
NPW	-1/2	$\check{\boldsymbol{\mathcal{F}}}^{-1/2} \boldsymbol{\mathcal{A}}' \check{\boldsymbol{\Pi}}^{-1} \boldsymbol{y}$	partly deconvolved backprojection
ROI	0	$\boldsymbol{\mathcal{A}}' \check{\boldsymbol{\Pi}}^{-1} \boldsymbol{y}$	backprojection

Thus the optimizing $\boldsymbol{\mathcal{Z}}$ for any of these observers need not include any form of regularization, even in the case that the system operator $\boldsymbol{\mathcal{A}}$ is a matrix with less than full column rank. We conclude on theoretical grounds that regularization is not absolutely essential for the task of detecting a statistically varying signal on a statistically varying background for any of the observers considered above.

Furthermore, there is a strong dependence of the optimal reconstruction method on the type of observer considered. This implies that there is no universally optimal reconstruction method, even for the simple detection task considered here, so it seems essential to consider observer models whose performance correlates well with human observers. As we mentioned above, the apparent premise of human-observer studies in the literature is that humans do not perform as well as the ideal observer. Therefore, the fact that there exist "simple" reconstruction methods that allow the observers considered above to achieve the ideal linear-detection SNR (which is also the overall ideal SNR for SKE tasks) suggests that these observers, the tasks, or both are somehow inappropriate.

s,det,cho 21.6.10 Channelized linear observers (s,det,cho)

For a channelized linear discriminant of the form (21.5.6) and a linear reconstruction method $\boldsymbol{\mathcal{Z}}$, the distribution of the test statistic $t = t_{\mathrm{ch}}(\boldsymbol{y})$ is

$$\begin{aligned} H_0 &: t_{\mathrm{ch}} \sim \mathcal{N}(\boldsymbol{w}' \boldsymbol{\mathcal{C}}' \boldsymbol{\mathcal{Z}} \boldsymbol{\mathcal{A}} \bar{\boldsymbol{f}}_b, \boldsymbol{w}' \boldsymbol{\mathcal{C}}' \boldsymbol{\mathcal{Z}} \boldsymbol{\Pi}_0 \boldsymbol{\mathcal{Z}}' \boldsymbol{\mathcal{C}} \boldsymbol{w} + \boldsymbol{w}' \boldsymbol{\Pi}_{\mathrm{int}} \boldsymbol{w}) \\ H_1 &: t_{\mathrm{ch}} \sim \mathcal{N}(\boldsymbol{w}' \boldsymbol{\mathcal{C}}' \boldsymbol{\mathcal{Z}} \boldsymbol{\mathcal{A}} (\bar{\boldsymbol{f}}_b + \bar{\boldsymbol{f}}_s), \boldsymbol{w}' \boldsymbol{\mathcal{C}}' \boldsymbol{\mathcal{Z}} [\boldsymbol{\Pi}_0 + \boldsymbol{\mathcal{A}} \boldsymbol{\mathcal{K}}_s \boldsymbol{\mathcal{A}}'] \boldsymbol{\mathcal{Z}}' \boldsymbol{\mathcal{C}} \boldsymbol{w} + \boldsymbol{w}' \boldsymbol{\Pi}_{\mathrm{int}} \boldsymbol{w}). \end{aligned}$$

Combining the above with (21.3.1) yields the SNR of the channelized observer:

e,det,snr,chan

$$\mathrm{SNR}^2_{\mathrm{ch}} = \frac{(\boldsymbol{w}' \boldsymbol{\mathcal{C}}' \boldsymbol{\mathcal{Z}} \bar{\boldsymbol{y}})^2}{\boldsymbol{w}' \boldsymbol{\mathcal{C}}' \boldsymbol{\mathcal{Z}} \check{\boldsymbol{\Pi}} \boldsymbol{\mathcal{Z}}' \boldsymbol{\mathcal{C}} \boldsymbol{w} + \boldsymbol{w}' \boldsymbol{\Pi}_{\mathrm{int}} \boldsymbol{w}}. \tag{21.6.22}$$

In the absence of internal noise ($\boldsymbol{\Pi}_{\mathrm{int}} = \boldsymbol{0}$), the SNR is maximized when $\boldsymbol{\mathcal{Z}}' \boldsymbol{\mathcal{C}} \boldsymbol{w} \propto \check{\boldsymbol{\Pi}}^{-1} \bar{\boldsymbol{y}}$, similarly to the non-channelized version in section 21.6.3. The presence of internal noise will always decrease the SNR. The problem with optimizing (21.6.22) with $\boldsymbol{\Pi}_{\mathrm{int}} \neq \boldsymbol{0}$ over an unconstrained $\boldsymbol{\mathcal{Z}}$ is that a $\boldsymbol{\mathcal{Z}}$ of infinitely large norm will be optimal. Thus one would need some constraint on $\boldsymbol{\mathcal{Z}}$ to optimize (21.6.22) in the presence of internal noise. This, however, is beyond the scope of our analysis.

21.6.11 Channelized Hotelling observer

Substituting the statistics of the M-channel output $\hat{\boldsymbol{c}}$ in the CHO template (21.5.7) yields

$$\boldsymbol{w}_{\text{CHO}} = (\mathcal{C}'\mathcal{Z}\check{\Pi}\mathcal{Z}'\mathcal{C} + \Pi_{\text{int}})^{\dagger}\mathcal{C}'\mathcal{Z}\,\bar{\boldsymbol{y}}\,.$$

Combining this template with (21.6.22), we find that the CHO observer has the following SNR:

e,det,snr,cho

$$\text{SNR}^2_{\text{CHO}}(\bar{\boldsymbol{f}}_s, \mathcal{Z}) = \bar{\boldsymbol{y}}'\,\mathcal{Z}'\mathcal{C}(\mathcal{C}'\mathcal{Z}\check{\Pi}\mathcal{Z}'\mathcal{C} + \Pi_{\text{int}})^{\dagger}\mathcal{C}'\mathcal{Z}\,\bar{\boldsymbol{y}}\,. \tag{21.6.23}$$

In the absence of internal noise, the usual SNR inequality holds:

$$\begin{aligned}\text{SNR}^2_{\text{CHO}}(\bar{\boldsymbol{f}}_s, \mathcal{Z}) &= (\check{\Pi}^{-1/2}\,\bar{\boldsymbol{y}})'\mathcal{P}_{\check{\Pi}^{1/2}\mathcal{Z}'\mathcal{C}}(\check{\Pi}^{-1/2}\,\bar{\boldsymbol{y}}) \\ &\leq (\check{\Pi}^{-1/2}\,\bar{\boldsymbol{y}})'\check{\Pi}^{-1/2}\,\bar{\boldsymbol{y}} = \bar{\boldsymbol{f}}_s'\check{\mathcal{F}}\bar{\boldsymbol{f}}_s = \text{SNR}^2_{\text{lin}_\circ}.\end{aligned}$$

The ideal SNR is achieved when $\check{\Pi}^{-1/2}\,\bar{\boldsymbol{y}} \in \mathcal{R}_{\check{\Pi}^{1/2}\mathcal{Z}'\mathcal{C}}$. It is easy to check that this is satisfied by any reconstructor of the form

$$\mathcal{Z} = \mathcal{W}\mathcal{C}(\mathcal{C}'\mathcal{W}\mathcal{C})^{-1}\mathcal{G}'\mathcal{A}'\check{\Pi}^{-1},$$

where $\mathcal{W}$ is any image-domain weighting operator and $\mathcal{G}$ is a mapping from $\mathbb{C}^M$ to object space that satisfies $\bar{\boldsymbol{f}}_s = \mathcal{G}\,\boldsymbol{u}$ for some $\boldsymbol{u} \in \mathbb{C}^M$. An obvious way to satisfy this requirement is to choose $\mathcal{G}$ so that one of its "columns" is proportional to $\bar{\boldsymbol{f}}_s$. This rather unconventional family of reconstructors does not produce what we usually consider to be reconstructed images. Furthermore, in the presence of internal noise the SNR will always be strictly less than the ideal. Nevertheless, we can still gain insight by using local Fourier-domain approximations to examine whether more conventional reconstructors can be combined with the CHO to achieve SNR values close to the ideal. In the following sections we examine one unregularized and one regularized example.

21.6.11.1 Local Fourier analysis of CHO performance

In the following we will use **local shift invariance** approximations. Circulant approximations of $\mathcal{A}'\Pi^{-1}\mathcal{A}$ have proven to be useful and accurate when this operator is approximately locally shift-invariant [19–21, 25–28]. Here we adopt the more general, angle-dependent analysis followed in [29, 30]. Since a practical implementation would employ DFT's, we use here a discretized version of the frequency response derived in Chapter 4. Thus we adopt the following approximation around the location of the target signal:

e,det,F,approx

$$\mathcal{F} = \mathcal{A}'\Pi^{-1}\mathcal{A} \approx \mathcal{U}^{-1}\,\Lambda\,\mathcal{U}, \tag{21.6.24}$$

where $\mathcal{U}$ is here a continuous-to-discrete Fourier operator and $\Lambda = \text{diag}\{\lambda_k\}$. The λ_k, $k = 1, ..., n_{\text{p}}$ are a discretized version of the frequency response $\lambda(\rho, \Phi)$ around the location of the target signal, as given in (4.4.7). Since $\mathcal{F}$ is symmetric positive-semidefinite, we force the λ_k's to be real and nonnegative by discarding imaginary parts and setting negatives to zero.

Similarly, we take $\mathcal{K}_f \approx \mathcal{U}^{-1}\boldsymbol{N}\,\mathcal{U}$, where $\boldsymbol{N} = \text{diag}\{\nu_k\}$. The ν_k's contain the frequency response of $\mathcal{K}_f$ (*i.e.*, the object power spectrum) local to the target signal position. Using these approximations, we can start from (21.6.13) to derive the following form for the Fisher information matrix $\check{\mathcal{F}}$:

e,det,Fch,approx

$$\check{\mathcal{F}} = \mathcal{F}(\mathcal{I} + \mathcal{K}_f\mathcal{F})^{-1} \approx \mathcal{U}^{-1}\,\Lambda(\boldsymbol{I} + \boldsymbol{N}\,\Lambda)^{-1}\mathcal{U} = \mathcal{U}^{-1}\check{\Lambda}\,\mathcal{U}, \tag{21.6.25}$$

where $\check{\Lambda} \triangleq \text{diag}\{\check{\lambda}_k\}$ and $\check{\lambda}_k \triangleq \frac{\lambda_k}{1+\nu_k\lambda_k}$, which reduces to $\check{\Lambda} = \Lambda$ in the SKE/BKE case.

As in [31–33], we also use the fact that $\mathcal{C}$ is a collection of filters to get its frequency-domain representation. Let $\boldsymbol{t}^m \in \mathbb{C}^{n_{\text{p}}}$ denote the frequency response of the mth bandpass filter. Then the mth operator in $\mathcal{C}$ has the form $\mathcal{U}^{-1}\,\text{diag}\{\boldsymbol{t}^m\}\,\mathcal{U}\boldsymbol{e}_0$, where $\boldsymbol{e}_0$ is an impulse centered at the same location as $\bar{\boldsymbol{f}}_s$. Without loss of generality, we can choose the target signal center to correspond to the "0" coordinate for the DFT matrix, in which case $\mathcal{U}\boldsymbol{e}_0 = \frac{1}{\sqrt{n_{\text{p}}}}\boldsymbol{1}$, where $\boldsymbol{1}$ is the vector of n_{p} ones. Thus we have

e,det,channels

$$\mathcal{C} = \mathcal{U}^{-1}\boldsymbol{T}, \qquad \boldsymbol{T} = \frac{1}{\sqrt{n_{\text{p}}}}\left[\boldsymbol{t}^1\ \ldots\ \boldsymbol{t}^M\right]. \tag{21.6.26}$$

s,det,cho,fisher 21.6.11.2 CHO and Fisher reconstructors (s,det,cho,fisher)

For a Fisher reconstructor $\mathcal{Z}_q$ of the form (21.6.15), we can use (21.6.25) and (21.6.26) to obtain the approximations

$$\begin{aligned}\mathcal{C}'\mathcal{Z}_q\,\bar{\boldsymbol{y}} &= \mathcal{C}'\check{\mathcal{F}}^{1+q}\bar{\boldsymbol{f}}_s \approx \boldsymbol{T}'\check{\Lambda}^{1+q}\boldsymbol{X} \\ \mathcal{C}'\mathcal{Z}_q\check{\Pi}\mathcal{Z}_q'\mathcal{C} &= \mathcal{C}'\check{\mathcal{F}}^{1+2q}\mathcal{C} \approx \boldsymbol{T}'\check{\Lambda}^{1+2q}\boldsymbol{T},\end{aligned}$$

where we assume that the mean target signal $\bar{\boldsymbol{f}}_s$ is spatially localized and where $\boldsymbol{X} = \mathcal{U}\bar{\boldsymbol{f}}_s$ is its spectrum. The accuracy of the second approximation will depend on how localized in space the channel responses are. However, it is exact for $q = -1/2$, so it may be reasonable for q near $-1/2$ (including $q = -1$ and $q = 0$, which are the two cases of greatest practical interest). Approximation error plots for some values of q were presented in [32].

Substituting the above approximations into (21.6.23), we get the following approximation for the SNR of the CHO when combined with a Fisher reconstructor:

$$\mathrm{SNR}^2_{\text{CHO,F}} \approx \boldsymbol{X}' \check{\boldsymbol{\Lambda}}^{1+q} \boldsymbol{T} (\boldsymbol{T}' \check{\boldsymbol{\Lambda}}^{1+2q} \boldsymbol{T} + \boldsymbol{\Pi}_{\text{int}})^{\dagger} \boldsymbol{T}' \check{\boldsymbol{\Lambda}}^{1+q} \boldsymbol{X}.$$

21.6.11.3 Special case: Disjoint passbands

The CHO filters are sometimes assumed to be bandpass filters with disjoint passbands. In that case, the vectors t_m have disjoint nonzero entries and the $M \times M$ matrix $\boldsymbol{T}' \check{\boldsymbol{\Lambda}}^{1+2q} \boldsymbol{T}$ is diagonal, so we have

$$[\boldsymbol{T}' \check{\boldsymbol{\Lambda}}^{1+q} \boldsymbol{X}]_m = \frac{1}{\sqrt{n_{\mathrm{p}}}} \sum_{k \in \mathcal{T}_m} (t_k^m)^* \check{\lambda}_k^{1+q} X_k$$
$$[\boldsymbol{T}' \check{\boldsymbol{\Lambda}}^{1+2q} \boldsymbol{T}]_{mm} = \frac{1}{n_{\mathrm{p}}} \sum_{k \in \mathcal{T}_m} |t_k^m|^2 \check{\lambda}_k^{1+2q},$$

where $\mathcal{T}_m = \{k : t_k^m \neq 0\}$ is the passband of the mth filter. The SNR approximation then simplifies to

$$\mathrm{SNR}^2_{\text{CHO,F}} \approx \mathrm{SNR}_1^2 \triangleq \sum_{m=1}^M \frac{\frac{1}{n_{\mathrm{p}}} \left| \sum_{k \in \mathcal{T}_m} X_k (t_k^m)^* \check{\lambda}_k^{1+q} \right|^2}{\frac{1}{n_{\mathrm{p}}} \sum_{k \in \mathcal{T}_m} |t_k^m|^2 \check{\lambda}_k^{1+2q} + \sigma^2_{\text{int}}}, \qquad (21.6.27)$$ e,det,cho,snr1

where we have assumed, following convention, that the components of $\boldsymbol{\varepsilon}_{\text{int}}$ are i.i.d. and σ^2_{int} is their variance. (This is equivalent to noise being added to the test statistic after the template is applied [34].) We examine now whether there are conditions under which maximum SNR can be achieved for $\sigma^2_{\text{int}} = 0$.

s,det,cho,fisher,opt 21.6.11.4 Achievability of the optimal SNR

To determine an upper bound on SNR_1, define

$$u_k^m \triangleq X_k \check{\lambda}_k^{1/2} 1_{\{k \in \mathcal{T}_m\}}$$
$$v_k^m \triangleq t_k^m \check{\lambda}_k^{q+1/2}.$$

Then rewriting (21.6.27) for $\sigma^2_{\text{int}} = 0$ and using Cauchy's inequality yields

$$\mathrm{SNR}_1^2 = \sum_{m=1}^M \frac{|\langle \boldsymbol{u}^m, \boldsymbol{v}^m \rangle|^2}{\|\boldsymbol{v}^m\|^2} \leq \sum_{m=1}^M \|\boldsymbol{u}^m\|^2 = \sum_{k \in \mathcal{T}} |X_k|^2 \check{\lambda}_k \leq \sum_k |X_k|^2 \check{\lambda}_k \approx \bar{\boldsymbol{f}}_s' \check{\boldsymbol{\mathcal{F}}} \bar{\boldsymbol{f}}_s = \mathrm{SNR}^2_{\mathrm{lin}_\circ}, \qquad (21.6.28)$$ e,det,fisher,cho,bound1

where $\mathcal{T} = \cup_{m=1}^M \mathcal{T}_m$ denotes the combined passbands of all the channels. If the combined passbands do not contain all of the signal energy, then the SNR will be strictly less than the ideal SNR. This suboptimality is expected due to the dimensionality decrease caused by the channels.

When can the upper bound in (21.6.28) be achieved? Suppose that each channel filter has a flat passband, *i.e.*, $t_k^m = 1_{\{k \in \mathcal{T}_m\}}$. Then there are two obvious cases where this SNR achieves the upper bound in (21.6.28), as can be verified by substitution or by using the requirement $\boldsymbol{u}^m \propto \boldsymbol{v}^m \ \forall m$.

- If the X_k's are constant over each passband, then $q = 0$ will be optimal.
- If the $\check{\lambda}_k$'s are also constant over each passband, then any $q \in \mathbb{R}$ will be optimal.

In practice, it may be unlikely that either the $\check{\lambda}_k$'s or the X_k's are *exactly* uniform over each channel's passband, but if the passbands are reasonably narrow, then it is likely that these spectra will be *approximately* uniform over each passband. So to within the accuracy of the approximations considered above, one or more of these unregularized reconstructors will nearly achieve the highest SNR obtainable for the given CHO channels. Once again, in the absence of internal observer noise, regularization does not seem to play a crucial role, even for the CHO.

s,det,cho,pwls 21.6.11.5 CHO and PWLS reconstructors (s,det,cho,pwls)

The preceding analysis has shown several cases in which one or more unregularized reconstructors lead to the optimal SNR in the detection task at hand. We next examine the penalized weighted least-squares (PWLS) family of regularized reconstructors, to explore how closely one can approach the optimal SNR with a practical reconstruction method.

An unconstrained PWLS estimator has the following form:

e,det,pwls,w

$$\hat{\boldsymbol{f}}(\boldsymbol{y}) = \arg\min_{\boldsymbol{x}} \{(\boldsymbol{y} - \boldsymbol{\mathcal{A}}\boldsymbol{x})'\boldsymbol{W}(\boldsymbol{y} - \boldsymbol{\mathcal{A}}\boldsymbol{x}) + \boldsymbol{x}'\boldsymbol{\mathcal{R}}\boldsymbol{x}\} = (\boldsymbol{\mathcal{A}}'\boldsymbol{W}\boldsymbol{\mathcal{A}} + \boldsymbol{\mathcal{R}})^{-1}\boldsymbol{\mathcal{A}}'\boldsymbol{W}\boldsymbol{y} \tag{21.6.29}$$

for some regularization operator $\boldsymbol{\mathcal{R}}$ that we would like to design to optimize detectability. The usual weighting for the PWLS estimator is the one corresponding to MAP estimation, *i.e.*, $\boldsymbol{W} = (\mathsf{Cov}\{\boldsymbol{y}|\boldsymbol{f}_b = \bar{\boldsymbol{f}}_b\})^{-1} = \boldsymbol{\Pi}^{-1}$. In this case, (21.6.29) becomes

e,det,pwls

$$\hat{\boldsymbol{f}}(\boldsymbol{y}) = (\boldsymbol{\mathcal{F}} + \boldsymbol{\mathcal{R}})^{-1}\boldsymbol{\mathcal{A}}'\boldsymbol{\Pi}^{-1}\boldsymbol{y}, \tag{21.6.30}$$

which corresponds to the choice $\boldsymbol{\mathcal{Z}} = (\boldsymbol{\mathcal{F}} + \boldsymbol{\mathcal{R}})^{-1}\boldsymbol{\mathcal{A}}'\boldsymbol{\Pi}^{-1}$. We assume throughout that the regularization operator $\boldsymbol{\mathcal{R}}$ is chosen such that $\boldsymbol{\mathcal{F}} + \boldsymbol{\mathcal{R}}$ is positive definite.

To analyze CHO performance with PWLS reconstruction, we assume that both $\boldsymbol{\mathcal{F}}$ and $\boldsymbol{\mathcal{R}}$ are diagonalized locally by a common operator (the Fourier operator $\boldsymbol{\mathcal{U}}$). Specifically, following (21.6.24) we assume that

e,det,diag,pwls

$$\boldsymbol{\mathcal{F}} \approx \boldsymbol{\mathcal{U}}^{-1}\,\boldsymbol{\Lambda}\,\boldsymbol{\mathcal{U}} \text{ and } \boldsymbol{\mathcal{R}} \approx \boldsymbol{\mathcal{U}}^{-1}\Omega\,\boldsymbol{\mathcal{U}}, \tag{21.6.31}$$

where $\boldsymbol{\Lambda} = \mathsf{diag}\{\lambda_k\}$ as defined in the previous section, $\Omega = \mathsf{diag}\{\omega_k\}$ and the ω_k, $k = 1, \ldots, n_{\mathrm{p}}$ are the frequency response of the regularizer local to the target signal. Both the λ_k's and the ω_k's are real and nonnegative. The assumption of simultaneous diagonalization of $\boldsymbol{\mathcal{F}}$ and $\boldsymbol{\mathcal{R}}$ is reasonable. Similar approximations were used by other researchers who have analyzed observer performance with penalized-likelihood reconstruction [27, 31, 33].

Substituting the PWLS reconstructor (21.6.30) in the SNR of the CHO (21.6.23) yields

e,det,snr,pwls

$$\begin{aligned}\mathrm{SNR}^2_{\text{CHO,PWLS}} = {} & \bar{\boldsymbol{f}}_s'\boldsymbol{\mathcal{F}}(\boldsymbol{\mathcal{F}} + \boldsymbol{\mathcal{R}})^{-1}\boldsymbol{\mathcal{C}}[\boldsymbol{\mathcal{C}}'(\boldsymbol{\mathcal{F}} + \boldsymbol{\mathcal{R}})^{-1}(\boldsymbol{\mathcal{F}} + \boldsymbol{\mathcal{F}}\boldsymbol{\mathcal{K}}_f\boldsymbol{\mathcal{F}})(\boldsymbol{\mathcal{F}} + \boldsymbol{\mathcal{R}})^{-1}\boldsymbol{\mathcal{C}} + \boldsymbol{\Pi}_{\text{int}}]^{\dagger} \\ & \boldsymbol{\mathcal{C}}'(\boldsymbol{\mathcal{F}} + \boldsymbol{\mathcal{R}})^{-1}\boldsymbol{\mathcal{F}}\bar{\boldsymbol{f}}_s. \end{aligned} \tag{21.6.32}$$

Following section 21.6.11.2, we use the local Fourier decompositions in (21.6.31) to obtain the following approximations:

e,det,cho,pwls,TDT

$$\begin{aligned}\boldsymbol{\mathcal{C}}'(\boldsymbol{\mathcal{F}} + \boldsymbol{\mathcal{R}})^{-1}\boldsymbol{\mathcal{F}}\bar{\boldsymbol{f}}_s &\approx \boldsymbol{T}'(\boldsymbol{\Lambda} + \Omega)^{-1}\,\boldsymbol{\Lambda}\,\boldsymbol{X} \\ \boldsymbol{\mathcal{C}}'(\boldsymbol{\mathcal{F}} + \boldsymbol{\mathcal{R}})^{-1}(\boldsymbol{\mathcal{F}} + \boldsymbol{\mathcal{F}}\boldsymbol{\mathcal{K}}_f\boldsymbol{\mathcal{F}})(\boldsymbol{\mathcal{F}} + \boldsymbol{\mathcal{R}})^{-1}\boldsymbol{\mathcal{C}} &\approx \boldsymbol{T}'(\boldsymbol{\Lambda} + \Omega)^{-1}(\boldsymbol{\Lambda} + \boldsymbol{\Lambda}^2\,\boldsymbol{N})(\boldsymbol{\Lambda} + \Omega)^{-1}\boldsymbol{T}. \end{aligned} \tag{21.6.33}$$

Substituting the above into (21.6.32) yields the following approximation for the SNR of the CHO when combined with a PWLS reconstructor:

e,det,snr,pwls,local

$$\mathrm{SNR}^2_{\text{CHO,PWLS}} \approx \boldsymbol{X}'\,\boldsymbol{\Lambda}(\boldsymbol{\Lambda} + \Omega)^{-1}\,\boldsymbol{T}(\boldsymbol{T}'(\boldsymbol{\Lambda} + \Omega)^{-1}(\boldsymbol{\Lambda} + \boldsymbol{\Lambda}^2\,\boldsymbol{N})(\boldsymbol{\Lambda} + \Omega)^{-1}\boldsymbol{T} + \boldsymbol{\Pi}_{\text{int}})^{\dagger}\boldsymbol{T}'(\boldsymbol{\Lambda} + \Omega)^{-1}\,\boldsymbol{\Lambda}\,\boldsymbol{X}. \tag{21.6.34}$$

21.6.11.6 Special case: Disjoint passbands

When the CHO channels are bandpass filters with disjoint frequency responses, the $M \times M$ matrix in (21.6.33) above becomes diagonal. We assume that its diagonal elements are nonzero, which implies that the system has some nonzero λ_k for each passband. (If not, the noninformative passband could be eliminated.) Combining the approximations above yields the following approximate expression for the SNR in (21.6.32):

e,det,snr,pwls,approx

$$\mathrm{SNR}^2_{\text{CHO,PWLS}} \approx \mathrm{SNR}^2_2 \triangleq \sum_{m=1}^{M} \frac{\frac{1}{n_{\mathrm{p}}}\left|\sum_{k\in\mathcal{T}_m} X_k (t_k^m)^* \frac{\lambda_k}{\lambda_k + \omega_k}\right|^2}{\frac{1}{n_{\mathrm{p}}}\sum_{k\in\mathcal{T}_m} |t_k^m|^2 \frac{\lambda_k^2}{\breve{\lambda}_k(\lambda_k + \omega_k)^2} + \sigma^2_{\text{int}}}, \tag{21.6.35}$$

where $\mathcal{T}_m = \{k : t_k^m \neq 0\}$ is the passband of the mth filter.

s,det,cho,pwls,opt

21.6.11.7 Achievability of the optimal SNR

To determine an upper bound on the SNR in (21.6.35) in the absence of internal noise, define vectors $\boldsymbol{u}^m$, $\boldsymbol{v}^m$ with elements

$$\begin{aligned} u_k^m &\triangleq X_k \breve{\lambda}_k^{1/2} 1_{\{k\in\mathcal{T}_m\}}, \\ v_k^m &\triangleq t_k^m \lambda_k / \breve{\lambda}_k^{1/2}(\lambda_k + \omega_k), \end{aligned}$$

respectively. Then for $\sigma^2_{\text{int}} = 0$, the SNR in (21.6.35) simplifies to

e,det,pwls,cho,bound1

$$\mathrm{SNR}^2_2 = \sum_{m=1}^{M} \frac{|\langle \boldsymbol{u}^m,\, \boldsymbol{v}^m\rangle|^2}{\|\boldsymbol{v}^m\|^2} \leq \sum_{m=1}^{M} \|\boldsymbol{u}^m\|^2 = \sum_{k\in\mathcal{T}} |X_k|^2 \breve{\lambda}_k \tag{21.6.36}$$

e,det,pwls,cho,bound2

$$\leq \sum_{k} |X_k|^2 \breve{\lambda}_k \approx \bar{\boldsymbol{f}}_s'\breve{\boldsymbol{\mathcal{F}}}\bar{\boldsymbol{f}}_s = \mathrm{SNR}^2_{\text{lin}_\circ}, \tag{21.6.37}$$

where the combined passband of all the filters is given by $\mathcal{T} = \cup_{m=1}^{M} \mathcal{T}_m$. If the combined passband does not contain all of the signal energy, then the SNR will be strictly less than the ideal SNR.

The intermediate upper bound in (21.6.36) is achieved if $\boldsymbol{u}^m \propto \boldsymbol{v}^m \ \forall m$. Suppose that each channel filter is an ideal bandpass filter over some frequency band, *i.e.*, $t_k^m = 1_{\{k \in \mathcal{T}_m\}}$. Then, for $\check{\lambda}_k \neq 0$, $X_k \neq 0$, the intermediate upper bound in (21.6.36) is achieved for example when

e,det,pwls,cho,omopt

$$\omega_k = \alpha \frac{\lambda_k}{X_k \check{\lambda}_k} - \lambda_k, \tag{21.6.38}$$

where the constant $\alpha \neq 0$ can be chosen arbitrarily. Using $\alpha \triangleq 2 \max_k (X_k \check{\lambda}_k)$ would keep the ω_k's positive. To within approximation (21.6.31), the local frequency response in (21.6.38) corresponds to the following positive-semidefinite regularizer:

$$\boldsymbol{\mathcal{R}} = \alpha(\boldsymbol{\mathcal{I}} + \boldsymbol{\mathcal{F}} \boldsymbol{\mathcal{K}}_f)(\boldsymbol{\mathcal{U}}^{-1} \, \mathsf{diag}\{1/X_k\} \boldsymbol{\mathcal{U}}) - \boldsymbol{\mathcal{F}}.$$

This $\boldsymbol{\mathcal{R}}$ will usually have a high-pass characteristic, so it could be construed as a regularization matrix, but it is quite different from standard forms of regularization studied in the literature. With this $\boldsymbol{\mathcal{R}}$, the corresponding "PWLS" reconstruction yields

$$\hat{\boldsymbol{f}} = \frac{1}{\alpha} (\boldsymbol{\mathcal{U}}^{-1} \, \mathsf{diag}\{X_k\} \boldsymbol{\mathcal{U}}) \boldsymbol{\mathcal{A}}' \check{\boldsymbol{\Pi}}^{-1} \boldsymbol{y},$$

which is simply a weighted backprojection followed by application of a "matched" filter (convolution with the mean signal shape). This agrees with our conclusion from the previous section that regularization is not essential even for the CHO observer, when the passbands of the CHO channels are flat and there is no internal noise.

Similarly to the previous section, a degenerate case occurs when the channel passbands are flat and the mean signal spectrum $\{X_k\}$, system spectrum $\{\lambda_k\}$, and object power spectrum $\{\nu_k\}$ are all constant over each channel's passband. Then the first upper bound in (21.6.36) is achieved for *any* choice of regularization $\{\omega_k\}$ that is also constant over each passband, including $\omega_k = 0$. So apparently, in the absence of internal noise, the choice of regularization may be important only if there is significant within-passband variation of the mean signal spectrum, the system spectrum, the object power spectrum, and/or the channel response itself.

21.6.11.8 Example

We now present a practical example of how the regularizer's frequency response Ω affects the SNR of the CHO with overlapping or non-overlapping passbands in the presence of internal noise. We consider the case where $\boldsymbol{\mathcal{A}}$ corresponds to a 2-D PET system model with the characteristics of a CTI ECAT 931 scanner (matrix size 128×128, pixel size 4.7mm, 192 radial samples with 3.1mm spacing, 160 projection angles over $180°$). We assume that the target signal $\boldsymbol{f}_s$ has a known Gaussian shape with FWHM 2 pixels and amplitude 0.1, the background $\boldsymbol{f}_b$ has a Gaussian autocorrelation function with FWHM 4 pixels and standard deviation 0.05, and the mean background $\bar{\boldsymbol{f}}_b$ is the anthropomorphic phantom shown in figure 21.6.2. We determine imaging noise variance by assuming a total of 5×10^5 counts.

We consider the non-overlapping square channels with $M = 4$ (SQR) and the overlapping difference-of-Gaussians channels with $M = 3$ (S-DOG) and $M = 10$ (D-DOG), as defined by Abbey et al. [17] and shown in figure 21.5.1. Fig. 21.6.3 shows plots of the SNR for PWLS with uniform regularization within a first-order neighborhood and various values of the regularization parameter. The SNR is plotted for the three channel sets mentioned above and internal noise variance $\sigma_{\text{int}}^2 = 0.005$. All the SNR values in these plots are normalized with respect to the SNR upper bound in (21.6.37).

The sharp SNR drop for very large amounts of regularization, as seen in figure 21.6.3, occurs only when internal noise is present. Thus the SNR is somewhat sensitive to the choice of regularization parameter in the presence of internal noise. This is in agreement with what has been reported by Qi [33] and implies that observer noise is an important factor to consider when optimizing regularization methods with respect to detectability. However, no similar drop occurs for very small amounts of regularization, *i.e.*, the peak SNR achieved by PWLS with the optimal regularization parameter is not much higher than the SNR achieved by unregularized WLS. Varying the amount of imaging noise and/or background variability does not eliminate this effect. Thus, even when the CHO comes with internal noise and overlapping channel passbands, regularization does not appear to be essential in this detection task for the system considered here.

Fig. 21.6.3 also compares the exact SNR, computed from (21.6.32), to the approximate SNR, computed from (21.6.34). We compute here the approximate SNR using the angle-dependent local certainties $\kappa_{\varphi_j}, j = 1, \ldots, n_\varphi$, from (4.4.3), as we did in Chapter 5 with penalty design. The agreement between the exact and approximate SNR values reinforces the results that we obtained analytically in sections 21.6.11.4 and 21.6.11.7 using the approximate expressions.

21.7 Discussion (s,det,discuss)

s,det,discuss

Our analysis shows that, for the task of known-location signal detection, there are unregularized reconstruction methods that can achieve the optimal SNR for several linear observer models, including the CHO in the absence of internal fig_obj_prof

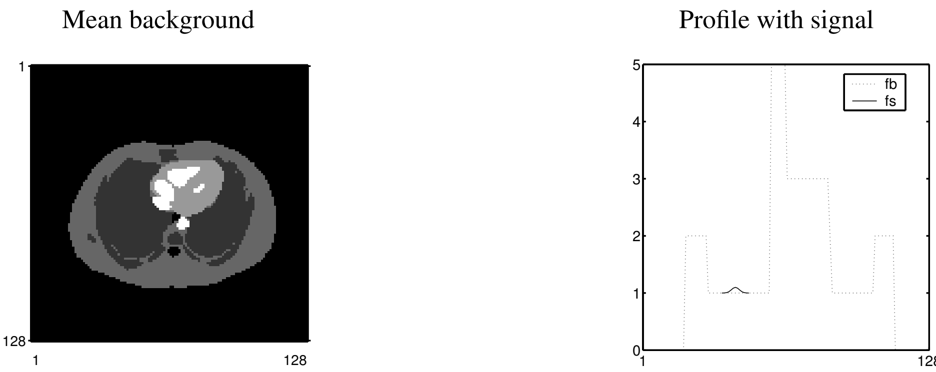

Figure 21.6.2: Mean background and profile through the mean background with the target signal superimposed. fig:object

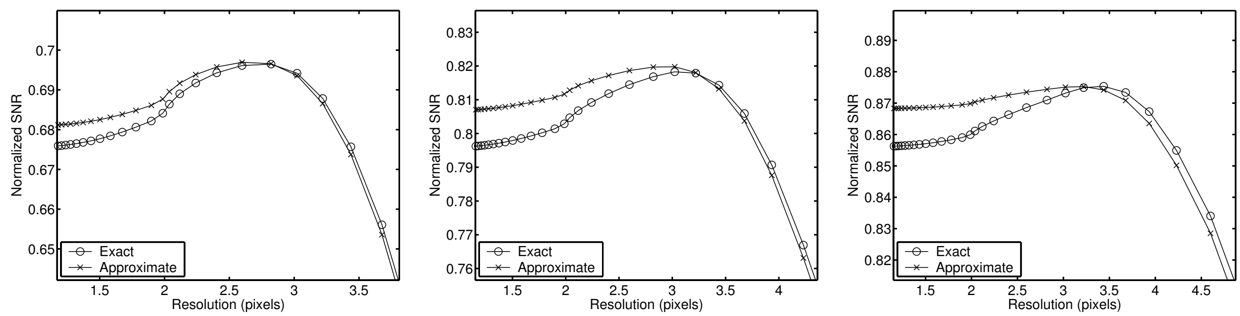

Figure 21.6.3: Exact and approximate SNR of CHO versus PWLS reconstruction resolution for three different channel sets. All SNR values are normalized with respect to the upper bound (ideal SNR for the non-channelized and internal-noise-free observer) in (21.6.37). fig:exact

observer noise. The presence of internal observer noise makes the SNR somewhat more sensitive to the choice of regularization parameter, in the sense that the SNR drops for large amounts of regularization, as shown by the PET example considered above. However, even in that case, optimizing regularized reconstruction does not lead to a significant improvement of SNR performance in comparison to unregularized reconstruction.

The relatively small significance of regularization throughout our analysis indicates that detection tasks where the target signal location is known exactly are most probably not suitable for optimizing regularized reconstruction methods. Apparently, resolution is not an essential image quality as far as known-location detectability is concerned. This indicates the importance of introducing location uncertainty in the analysis of image reconstruction methods with respect to detectability, a direction that recent work is in the process of exploring [3, 35, 36].

Another interesting area is signal detection in multimodality imaging [37].

s,det,prob

21.8 Problems (s,det,prob)

p,det,loc,glrt

Problem 21.1 *In the case where the signal shape is known but the signal location is unknown, a possible model is*
$H_0 : f(\vec{x}) = f_b(\vec{x})$
$H_1 : f(\vec{x}) = f_b(\vec{x}) + s(\vec{x} - \vec{\theta})$
where $\vec{\theta}$ denotes the unknown spatial location of the known signal $s(\vec{x})$. Consider regularized nonparametric image reconstruction of $f(\vec{x})$, and assume that $\hat{f}(\vec{x})$ is a ***gaussian random field****, with mean and covariance function that we can approximate by local shift invariance analysis. Then one can estimate $\vec{\theta}$ by maximum likelihood (**ML**) from $\hat{f}(\vec{x})$, and perform a* ***generalized likelihood ratio test*** *(**GLRT**) to decide between H_0 and H_1. The asymptotic performance of that GLRT test can be analyzed as in [38, p. 239], and involves chi-squared and noncentral χ^2 distributions. Use this sketch to investigate the performance as a function of the regularization parameter β.*
This is an alternative approach to the analysis in [2, 3, 39]. The details are an open problem. *(Solve?)*

21.9 Bibliography

duda:01
[1] R. O. Duda, P. E. Hart, and D. G. Stork. *Pattern classification*. New York: Wiley, 2001 (cit. on pp. 21.2, 21.7).

yendiki:07:aoo
[2] A. Yendiki and J. A. Fessler. "Analysis of observer performance in unknown-location tasks for tomographic image reconstruction." In: *J. Opt. Soc. Am. A* 24.12 (Dec. 2007). Special issue on Image Quality., B99–109. DOI: 10.1364/JOSAA.24.000B99 (cit. on pp. 21.2, 21.15).

yendiki:04:aoo
[3] A. Yendiki and J. A. Fessler. "Analysis of observer performance in detecting signals with location uncertainty for regularized tomographic image reconstruction." In: *Proc. IEEE Nuc. Sci. Symp. Med. Im. Conf.* Vol. 4. 2004, 2620–4. DOI: 10.1109/NSSMIC.2004.1462789 (cit. on pp. 21.2, 21.15).

fukunaga:90
[4] K. Fukunaga. *Introduction to statistical pattern recognition*. Boston: Academic, 1990 (cit. on pp. 21.2, 21.6, 21.9).

barrett:03
[5] H. H. Barrett and K. J. Myers. *Foundations of image science*. New York: Wiley, 2003 (cit. on pp. 21.2, 21.3, 21.4).

vantrees:68
[6] H. L. Van Trees. *Detection, estimation, and modulation theory*. New York: Wiley, 1968 (cit. on p. 21.3).

burgess:81:eoh
[7] A. E. Burgess et al. "Efficiency of human visual signal discrimination." In: *Science* 214.4516 (Oct. 1981), 93–94. URL: http://www.jstor.org/stable/1687274 (cit. on p. 21.4).

loo:84:aco
[8] L. N. D. Loo, K. Doi, and C. E. Metz. "A comparison of physical image comparison indices and observer performance in the radiographic detection of nylon beads." In: *Phys. Med. Biol.* 29.7 (July 1984), 837–57. DOI: 10.1088/0031-9155/29/7/007 (cit. on p. 21.4).

burgess:88:vsd
[9] A. E. Burgess and B. Colborne. "Visual signal detection. IV. Observer inconsistency." In: *J. Opt. Soc. Am. A* 5.4 (Apr. 1988), 617–. DOI: 10.1364/JOSAA.5.000617 (cit. on p. 21.4).

myers:87:aoa
[10] K. J. Myers and H. H. Barrett. "Addition of a channel mechanism to the ideal-observer model." In: *J. Opt. Soc. Am. A* 4.12 (Dec. 1987), 2447–57. DOI: 10.1364/JOSAA.4.002447 (cit. on p. 21.4).

myers:00:iom
[11] K. J. Myers. "Ideal observer models of visual signal detection." In: *Handbook of Medical Imaging, Volume 1. Physics and Psychophysics*. Ed. by J. Beutel, H. L. Kundel, and R. L. Van Metter. Bellingham: SPIE, 2000, pp. 558–92 (cit. on p. 21.4).

hotelling:1931:tgo
[12] H. Hotelling. "The generalization of Student's ratio." In: *Ann. Math. Stat.* 2.3 (1931), 360–78 (cit. on p. 21.4).

gallas:03:vtu
[13] B. D. Gallas and H. H. Barrett. "Validating the use of channels to estimate the ideal linear observer," in: *J. Opt. Soc. Am. A* 20.9 (Sept. 2003), 1725–38. DOI: 10.1364/JOSAA.20.001725 (cit. on p. 21.4).

gifford:99:aco
[14] H. C. Gifford, R. G. Wells, and M. A. King. "A comparison of human observer LROC and numerical observer ROC for tumor detection in SPECT images." In: *IEEE Trans. Nuc. Sci.* 46.4 (Aug. 1999), 1032–7. DOI: 10.1109/23.790820 (cit. on p. 21.5).

rolland:92:eor
[15] J. P. Rolland and H. H. Barrett. "Effect of random background inhomogeneity on observer detection performance." In: *J. Opt. Soc. Am. A* 9.5 (May 1992), 649–58. DOI: 10.1364/JOSAA.9.000649 (cit. on p. 21.5).

abbey:95:lir
[16] C. K. Abbey and H. H. Barrett. "Linear iterative reconstruction algorithms: study of observer performance." In: *Information Processing in Medical Im.* Ed. by Y. Bizais, C. Barillot, and R. D. Paola. Dordrect: Kluwer, 1995, pp. 65–76 (cit. on p. 21.5).

abbey:01:ham
[17] C. K. Abbey and H. H. Barrett. "Human- and model-observer performance in ramp-spectrum noise: effects of regularization and object variability." In: *J. Opt. Soc. Am. A* 18.3 (Mar. 2001), 473–88. DOI: 10.1364/JOSAA.18.000473 (cit. on pp. 21.5, 21.14).

fessler:96:mav
[18] J. A. Fessler. "Mean and variance of implicitly defined biased estimators (such as penalized maximum likelihood): Applications to tomography." In: *IEEE Trans. Im. Proc.* 5.3 (Mar. 1996), 493–506. DOI: 10.1109/83.491322 (cit. on pp. 21.5, 21.6).

qi:00:ran
[19] J. Qi and R. M. Leahy. "Resolution and noise properties of MAP reconstruction for fully 3D PET." In: *IEEE Trans. Med. Imag.* 19.5 (May 2000), 493–506. DOI: 10.1109/42.870259 (cit. on pp. 21.6, 21.11).

fessler:96:srp
[20] J. A. Fessler and W. L. Rogers. "Spatial resolution properties of penalized-likelihood image reconstruction methods: Space-invariant tomographs." In: *IEEE Trans. Im. Proc.* 5.9 (Sept. 1996), 1346–58. DOI: 10.1109/83.535846 (cit. on pp. 21.6, 21.11).

fessler:99:cgp
[21] J. A. Fessler and S. D. Booth. "Conjugate-gradient preconditioning methods for shift-variant PET image reconstruction." In: *IEEE Trans. Im. Proc.* 8.5 (May 1999), 688–99. DOI: 10.1109/83.760336 (cit. on pp. 21.6, 21.11).

barrett:91:lda
[22] H. H. Barrett et al. "Linear discriminants and image quality." In: *Information Processing in Medical Im.* LNCS 511. 1991, 458–73. DOI: 10.1007/BFb0033773 (cit. on p. 21.7).

barrett:94:npo
[23] H. H. Barrett, D. W. Wilson, and B. M. W. Tsui. "Noise properties of the EM algorithm: I. Theory." In: *Phys. Med. Biol.* 39.5 (May 1994), 833–46. DOI: 10.1088/0031-9155/39/5/004 (cit. on p. 21.7).

abbey:02:ose
[24] C. K. Abbey and M. P. Eckstein. "Optimal shifted estimates of human-observer templates in two-alternative forced-choice experiments." In: *IEEE Trans. Med. Imag.* 21.5 (May 2002), 429–40. DOI: 10.1109/TMI.2002.1009379 (cit. on p. 21.8).

qi:01:tso
[25] J. Qi and R. H. Huesman. "Theoretical study of lesion detectability of MAP reconstruction using computer observers." In: *IEEE Trans. Med. Imag.* 20.8 (Aug. 2001), 815–22. DOI: 10.1109/42.938249 (cit. on pp. 21.9, 21.10, 21.11).

qi:99:ats
[26] J. Qi and R. M. Leahy. "A theoretical study of the contrast recovery and variance of MAP reconstructions from PET data." In: *IEEE Trans. Med. Imag.* 18.4 (Apr. 1999), 293–305. DOI: 10.1109/42.768839 (cit. on p. 21.11).

bonetto:00:caf
[27] P. Bonetto, J. Qi, and R. M. Leahy. "Covariance approximation for fast and accurate computation of channelized Hotelling observer statistics." In: *IEEE Trans. Nuc. Sci.* 47.4 (Aug. 2000), 1567–72. DOI: 10.1109/23.873017 (cit. on pp. 21.11, 21.13).

stayman:00:rfu
[28] J. W. Stayman and J. A. Fessler. "Regularization for uniform spatial resolution properties in penalized-likelihood image reconstruction." In: *IEEE Trans. Med. Imag.* 19.6 (June 2000), 601–15. DOI: 10.1109/42.870666 (cit. on p. 21.11).

fessler:03:aat
[29] J. A. Fessler. "Analytical approach to regularization design for isotropic spatial resolution." In: *Proc. IEEE Nuc. Sci. Symp. Med. Im. Conf.* Vol. 3. 2003, 2022–6. DOI: 10.1109/NSSMIC.2003.1352277 (cit. on p. 21.11).

yendiki:04:aat
[30] A. Yendiki and J. A. Fessler. "Analytical approach to channelized Hotelling observer performance for regularized tomographic image reconstruction." In: *Proc. IEEE Intl. Symp. Biomed. Imag.* 2004, 360–3. DOI: 10.1109/ISBI.2004.1398549 (cit. on p. 21.11).

xing:03:rco
[31] Y. Xing, I-T. Hsiao, and G. Gindi. "Rapid calculation of detectability in Bayesian single photon emission computed tomography." In: *Phys. Med. Biol.* 48.22 (Nov. 2003), 3755–74. DOI: 10.1088/0031-9155/48/22/009 (cit. on pp. 21.11, 21.13).

fessler:02:cho
[32] J. A. Fessler and A. Yendiki. "Channelized Hotelling observer performance for penalized-likelihood image reconstruction." In: *Proc. IEEE Nuc. Sci. Symp. Med. Im. Conf.* Vol. 2. 2002, 1040–4. DOI: 10.1109/NSSMIC.2002.1239500 (cit. on pp. 21.11, 21.12).

qi:04:aol
[33] J. Qi. "Analysis of lesion detectability in Bayesian emission reconstruction with nonstationary object variability." In: *IEEE Trans. Med. Imag.* 23.3 (Mar. 2004), 321–9. DOI: 10.1109/TMI.2004.824239 (cit. on pp. 21.11, 21.13, 21.14).

abbey:00:mvd
[34] C. K. Abbey and F. O. Bochud. "Modeling visual detection tasks in correlated image noise with linear model observers." In: *Handbook of Medical Imaging, Volume 1. Physics and Psychophysics*. Ed. by J. Beutel, H. L. Kundel, and R. L. Van Metter. Bellingham: SPIE, 2000, pp. 629–54 (cit. on p. 21.12).

khurd:04:lmo
[35] P. K. Khurd and G. R. Gindi. "LROC model observers for emission tomographic reconstruction." In: *Proc. SPIE 5372 Medical Imaging 2004: Im. Perc., Obs. Perf., Tech. Assessment*. 2004, 509–20. DOI: 10.1117/12.536442 (cit. on p. 21.15).

qi:04:fat
[36] J. Qi and R. H. Huesman. "Fast approach to evaluate MAP reconstruction for lesion detection and localization." In: *Proc. SPIE 5372 Medical Imaging 2004: Im. Perc., Obs. Perf., Tech. Assessment*. 2004, 273–82. DOI: 10.1117/12.535916 (cit. on p. 21.15).

clarkson:08:atb
[37] E. Clarkson et al. "A task-based approach to adaptive and multimodality imaging." In: *Proc. IEEE* 96.3 (Mar. 2008), 500–511. DOI: 10.1109/JPROC.2007.913553 (cit. on p. 21.15).

kay:98
[38] S. M. Kay. *Fundamentals of statistical signal processing: Detection theory*. New York: Prentice-Hall, 1998 (cit. on p. 21.15).

yendiki:06:aou
[39] A. Yendiki and J. A. Fessler. "Analysis of unknown-location signal detectability for regularized tomographic image reconstruction." In: *Proc. IEEE Intl. Symp. Biomed. Imag.* 2006, 279–83. DOI: 10.1109/ISBI.2006.1624907 (cit. on p. 21.15).