Efficient estimation of ideal-observer performance in classification tasks involving high-dimensional complex backgrounds

Subok Park\textsuperscript{1,*} and Eric Clarkson\textsuperscript{2}
\textsuperscript{1}NIBIB/CDRH Laboratory for the Assessment of Medical Imaging Systems, DIAM, CDRH, FDA, Silver Spring, Maryland 20993, USA
\textsuperscript{2}Optical Sciences Center and Department of Radiology, University of Arizona, Tucson, Arizona 85724, USA

Abstract

The Bayesian ideal observer is optimal among all observers and sets an absolute upper bound for the performance of any observer in classification tasks [Van Trees, Detection, Estimation, and Modulation Theory, Part I (Academic, 1968)]. Therefore, the ideal observer should be used for objective image quality assessment whenever possible. However, computation of ideal-observer performance is difficult in practice because this observer requires the full description of unknown, statistical properties of high-dimensional, complex data arising in real life problems. Previously, Markov-chain Monte Carlo (MCMC) methods were developed by Kupinski et al. [J. Opt. Soc. Am. A 20, 430(2003)] and by Park et al. [J. Opt. Soc. Am. A 24, B136 (2007) and IEEE Trans. Med. Imaging 28, 657 (2009)] to estimate the performance of the ideal observer and the channelized ideal observer (CIO), respectively, in classification tasks involving non-Gaussian random backgrounds. However, both algorithms had the disadvantage of long computation times. We propose a fast MCMC for real-time estimation of the likelihood ratio for the CIO. Our simulation results show that our method has the potential to speed up ideal-observer performance in tasks involving complex data when efficient channels are used for the CIO.

1. INTRODUCTION

The Bayesian ideal observer provides a quantitative measure of the diagnostic performance of an imaging system, summarized by the area under the receiver operating characteristic curve (AUC) [1]. This observer is optimal among all observers, either human or model, and sets an absolute upper bound for observer performance in classification tasks [1]. For system optimization and design, making use of all available information in raw data is much desired when assessing medical image quality. Therefore, the ideal observer should be used for image quality assessment whenever possible, as suboptimal observers use only a part of the information available in the data. However, the Bayesian ideal observer requires the full description of the statistical properties of the data to optimally perform a classification task [1], but such information is often unknown for complex, realistic backgrounds found in clinical applications. This has been the barrier to the use of the ideal observer approach for system optimization in practice. Furthermore, the dimension of the integrals that need to be calculated for the observer is huge because of the use of high-dimensional data sets. For
estimating the test statistic of the ideal observer, the likelihood ratio, in binary classification
tasks involving non-Gaussian distributed lumpy backgrounds, Kupinski et al. [2] developed
a Markov-chain Monte Carlo (MCMC) method. Subsequently, other researchers adopted
and modified the Kupinski MCMC approach to explore tasks involving other types of
images from cardiac single-photon emission computed tomography (SPECT) and breast
computed tomography [3,4]. However, these MCMC methods still have the disadvantage of
long computation times due to the high dimensionality of the data. In addition, the methods
are limited to particular patient models investigated in these studies.

In an attempt to reduce the computation load and still approximate ideal-observer
performance, Park et al. [5] investigated a channelized ideal observer (CIO) in binary
classification tasks involving non-Gaussian lumpy backgrounds, where a channel matrix
using Laguerre–Gauss (LG) channels was applied to the data in order to reduce the
dimension of the data. In that work, they extended the Kupinski MCMC to incorporate the
conditional probability of channelized data into the acceptance probability for constructing a
Markov chain to estimate the likelihood ratio of the CIO while using the same proposal
density for sampling backgrounds as in the Kupinski MCMC. They found that the CIO with
LG channels gave a close approximation to ideal-observer performance. However, there still
was an absolute gap between the mean performance of the ideal observer and the CIO with
LG channels. Moreover, LG channels are limited to rotationally symmetric signals and
lumpy backgrounds [6].

To overcome this shortcoming of LG channels as efficient channels, in their subsequent
work, Park et al. [7] investigated the use of singular value decomposition (SVD) of a given
linear imaging system in choosing efficient channels for the ideal observer in the same tasks.
In that work, singular vectors most strongly associated with the signal-only image were
found to be highly efficient for the ideal observer. In addition, singular vectors most strongly
associated with the background were also efficient when a large number of channels were
used. These SVD channels are not only highly efficient for the ideal observer, but also are
not limited to types of backgrounds and signals. However, the SVD approach for finding
efficient channels has a limitation in that it requires the system to be linear and the system’s
response functions to be known.

In order to develop a more flexible but effective method of choosing efficient channels, Park
and her collaborators considered a partial least square (PLS) method, where the PLS weights
are estimated by maximizing the covariance between the data and the truth (either signal-
present or signal-absent) [8]. The PLS channels were again found to be highly efficient for
the ideal observer in the aforementioned tasks. Furthermore, the PLS approach did not
require as many channels as the SVD for the CIO to estimate ideal-observer performance.
The PLS approach is flexible in the sense that it requires neither the knowledge of the
system’s response functions nor the system to be linear.

The extended MCMC developed by Park et al. [5] that was used for estimating the CIO
provided a computational tool for investigating the aforementioned efficient channels.
However, it did not reduce the computational burden of calculating ideal-observer
performance because the extended MCMC still used the proposal density of the Kupinski
MCMC for sampling high-dimensional backgrounds. The use of a parameter vector of the
lumpy-background model for testing the Kupinski MCMC algorithm facilitated the design
of their symmetric proposal density, but has limitations in practice in that it is often
infeasible to find a parameter vector that represents complex, realistic background images
found in clinical applications. Moreover, even though the use of the background parameter
vector somewhat alleviated the high-dimensionality problem, it still was not sufficient to
reduce the computational cost effectively and make the method practical to be used in real-time computation of ideal-observer performance.

In the present paper, we propose a novel approach to estimate the CIO, which samples from a much-lower-dimensional channelized data space without any parametric background model rather than from either the high-dimensional image or background parameter space. We call this approach a channelized ideal-observer Markov-chain Monte Carlo (CIO-MCMC). The CIO-MCMC method is general in that it can be used for estimating ideal-observer performance in binary classification tasks involving any type of background and signal provided that efficient channels for such types of background and signal are used. In this paper, to show the validity of the method for estimating the CIO in comparison with the ideal observer, we will test the CIO-MCMC in binary classification tasks involving non-Gaussian lumpy backgrounds, where the Kupinski MCMC algorithm provides a way to estimate ideal-observer performance. We will compare the CIO-MCMC method to the Kupinski MCMC for the ideal observer [2] and the extended MCMC for the CIO [5]. We will show that the CIO-MCMC method has the potential to become a practical method for real-time computation of ideal-observer performance in tasks involving complex backgrounds.

2. BACKGROUND

A. Binary Classification and Data Channelization

For binary classification tasks, signal-present and signal-absent hypotheses are considered, given by

\[ H_1: g = b + s + n, \]  
\[ H_0: g = b + n, \]

where the vectors \( b \) and \( s \) represent the noiseless background and signal images, respectively, \( n \) represents measurement noise, and \( g \) represents the resulting data vector. All the vectors are \( M \)-dimensional.

For channelization of the data, an \( N_c \times M \) channel matrix, \( T \), is applied to the data \( g \). This matrix consists of \( N_c \) rows of channel vectors to generate an \( N_c \)-dimensional channelized data vector \( v \) via

\[ v = Tg, \]

In order to investigate a fast, efficient computational method for estimating ideal-observer performance, the high-dimensionality problem of calculating the ideal observer needs to be resolved. To this end, efficient channels generated by the methods of SVD and PLS [7,8] can be used in \( T \) for the CIO to approximate the ideal observer. For notational convenience, \( Tb = v_b \) and \( Ts = v_s \) will be used, respectively, for the channelized noiseless background and signal images. For the channelized noise, \( Tn = v_n \) will be used.

B. Bayesian Ideal and Channelized Ideal Observers

The test statistic of the ideal observer is the likelihood ratio, given by
where $p(g|H_j)$ are the probability density functions (PDFs) of image data $g$ conditional on the hypothesis $H_j$, $j = 0, 1$. Similarly, the CIO likelihood ratio is defined as the ideal observer that optimally detects the signal seen in the channelized data, given by [5]:

$$
\Lambda(v) = \frac{p(v|H_1)}{p(v|H_0)}.
$$

The likelihood ratio of the CIO can be rewritten as an integral, using a parameter vector $\theta$ for the object to be imaged [5]:

$$
\Lambda(v) = \int \frac{d\theta}{\Lambda_{BKE}(v|\theta)} p(\theta|v, H_0),
$$

where $\Lambda_{BKE}(v|\theta)$ represents the likelihood ratio of channelized data for the background known exactly (BKE) case. In Eq. (6), $\Lambda_{BKE}(v|\theta)$ is given by

$$
\Lambda_{BKE}(v|\theta) = \frac{p(v|\theta, H_1)}{p(v|\theta, H_0)}.
$$

and the posterior density is

$$
p(\theta|v, H_0) = \frac{p(v|\theta, H_0) p(\theta)}{\int d\theta' p(v|\theta', H_0) p(\theta')}.
$$

Then, the likelihood ratio $\Lambda(v)$ in Eq. (6) can be approximated by an MCMC approach, in particular the extension of the Kupinski MCMC proposed by Park et al. [5]:

$$
\Lambda(v) = \frac{1}{J - J_0} \sum_{i=J_0+1}^{J} \Lambda_{BKE}(v|\theta^{(i)}),
$$

where $\{\theta^{(i)}\}$ are sampled parameter vectors of the background $\theta$ and $\{\theta^{(i)}\}$ constitute a Markov chain, and $J_0$ indicates a burn-in to remove the initial unstable samples. The extended MCMC samples backgrounds to construct a Markov chain for estimating the likelihood ratio of the CIO.

The extended MCMC was useful in studying the properties of the CIO in comparison with the ideal observer [5,7,8]. However, it did not improve the computational time for estimating ideal-observer performance. In the next section, we will describe our CIO-MCMC method that speeds up the computation of the performance of the CIO and hence ideal-observer performance by use of efficient channels.
3. METHODS

A. Likelihood Ratio for the Channelized Ideal Observer

Our approach to efficient estimation of ideal-observer performance is to design an algorithm that samples much lower dimensional channelized backgrounds, \( v_b \), for estimating the CIO rather than high-dimensional backgrounds themselves, \( b \). Our approach will improve sampling efficiency and hence reduce the computation burden of calculating ideal-observer performance, while the CIO combined with efficient channels still approximates ideal-observer performance. To this end, the CIO likelihood ratio in Eq. (5) can be rewritten as an integral with respect to the channelized background, \( v_b \) [9]:

\[
\Lambda(v) = \int dv_b \Lambda_{\text{CBKE}}(v|v'_b) \text{pr}(v'_b|v, H_0),
\]

(10)

where \( \Lambda_{\text{CBKE}}(v|v'_b) \) represents the likelihood ratio of the CIO for the channelized background known exactly (CBKE) case, given by

\[
\Lambda_{\text{CBKE}}(v|v'_b) = \frac{\text{pr}(v|v'_b, H_1)}{\text{pr}(v|v'_b, H_0)},
\]

(11)

and \( \text{pr}(v'_b|v, H_0) \) represents the posterior density of \( v'_b \) given by

\[
\text{pr}(v'_b|v, H_0) = \frac{\text{pr}(v|v'_b, H_0) \text{pr}(v'_b)}{\int dv'_b \text{pr}(v|v'_b, H_0) \text{pr}(v'_b)}. \]

(12)

In Appendix A, we show that Eqs. (6) and (10) are equivalent.

In the expression for \( \Lambda_{\text{CBKE}}(v|v'_b) \) given in Eq. (11), \( \text{pr}(v|v'_b, H_j) \) values are calculated by plugging \( v'_b \) into the PDFs, \( \text{pr}(v|v'_b, H_j) \), of channelized data \( v \) under the hypothesis \( H_j \) conditional on the channelized background \( v_b \). These PDFs are governed by the statistics of channelized noise \( v_n \equiv Tn \), since \( n \) is the only source of uncertainty given the background \( b \), i.e., \( \text{pr}(v|v_b, H_j) = \text{pr}(v_n|H_j) \). Therefore, \( \Lambda_{\text{CBKE}}(v|v'_b) \) can be calculated with knowledge of the statistics of channelized noise.

B. Proposal Density and Acceptance Probability for CIO-MCMC

The likelihood \( \Lambda(v) \) of the CIO, given in Eq. (10), can be approximated by an MCMC approach [9]:

\[
\Lambda(v) = \frac{1}{J - J_0} \sum_{j=J_0+1}^J \Lambda_{\text{CBKE}}(v|v^{(i)}_b),
\]

(13)

where \( \{v^{(i)}_b\}_{i=1}^J \) constitute a Markov chain and \( J_0 \) indicates a burn-in to remove sample dependence on the given channelized background. The MCMC approach allows a user-
defined proposal density from which to draw channelized background vectors, which would converge to samples from the posterior density without sampling from it directly.

Given the channelized data $v$ and the $i$th sample $v_b^{(i)}$, for sampling $v_b^{(s)}$, the proposal density, $q(v_b^{(s)}|v_b^{(i)})$, is modeled to be a multivariate Gaussian proposal with width, $\sigma_c$, centered at the previously accepted sample, $v_b^{(i)}$. Then the acceptance probability, $\alpha_{\text{c}}$, to either accept or reject the sampled vector, $v_b^{(s)}$, is given by [10]

$$
\alpha_{\text{c}} = \min \left\{ 1, \frac{\text{pr}(v_b^{(s)}|v, H_0)q(v_b^{(s)}|v_b^{(i)})}{\text{pr}(v_b^{(i)}|v, H_0)q(v_b^{(i)}|v_b^{(s)})} \right\}.
$$

(14)

The symmetry of our proposal density allows a Metropolis approach [10]. That is, the symmetry of the proposal density results in cancellation of the $q(\cdot|\cdot)$ terms in Eq. (14), and hence the acceptance probability is reduced to

$$
\alpha_{\text{c,post}} = \min \left\{ 1, \frac{\text{pr}(v_b^{(s)}|v, H_0)}{\text{pr}(v_b^{(i)}|v, H_0)} \right\}.
$$

(15)

The posterior density of $v_b^{(i)}$, $\text{pr}(v_b^{(i)}|v, H_0)$, is a constant multiple of $\text{pr}(v|v_b^{(i)}, H_0)\text{pr}(v_b^{(i)})$, where $\text{pr}(v|v_b^{(i)})$ is a prior density on $v_b^{(i)}$ and $\text{pr}(v|v_b^{(i)}, H_0)$ is the conditional density of $v$ on the channelized background $v_b^{(i)}$ and $H_0$. Therefore, another formula for the acceptance probability, $\alpha_{\text{p}}$, can be written as

$$
\alpha_{\text{p,post}} = \min \left\{ 1, \frac{\text{pr}(v|v_b^{(s)}, H_0)\text{pr}(v_b^{(s)})}{\text{pr}(v|v_b^{(i)}, H_0)\text{pr}(v_b^{(i)})} \right\}.
$$

(16)

**C. Modeling of Posterior or Prior Densities for CIO-MCMC**

As discussed above, for estimating the CIO likelihood ratio given in Eq. (5), the Park extended MCMC [5] was used. The extended MCMC uses the same proposal density for sampling backgrounds as the Kupinski MCMC [2], but a different acceptance probability expression that incorporates the conditional probabilities of channelized data on the channelized version of the given background, $b$. We note that this vector, $b$, was the background used for generating the data vector, $g$, for which the likelihood ratios of the ideal observer and the CIO were estimated. By design of the Kupinski MCMC and the Park extension, which followed the definition of the Metropolis–Hastings algorithm, the accepted sample backgrounds for constructing a Markov chain to estimate the CIO likelihood ratio tend to be similar to the given background. In other words, backgrounds drawn by the Kupinski proposal density are likely to be rejected if they are too different from the given background or, equivalently, if they cause too low values of the posterior.

The aforementioned consideration leads to an assumption that the accepted sample vectors, $\{v_b^{(i)}\}$, for estimating the CIO likelihood ratio, $\Lambda(v)$, using the extended MCMC, follow a Gaussian distribution centering around the channelized version of the given background, $v_b$. Therefore, marginal distributions of channelized vectors of the background samples...
accepted by the extended MCMC were modeled to be Gaussian in this work. This assumption is also supported by our observation of marginal distributions of channelized backgrounds obtained by using the extended MCMC. Figure 1 illustrates that a marginal distribution of the channelized version of MCMC sampled backgrounds for the given background are usually a sum of either two or three different Gaussian distributions. The figure also indicates that the sampled marginal distribution lies well within an overall marginal distribution of the channelized version of many different randomly generated backgrounds with the same background parameters (i.e., the same mean number of lumps, $N^\text{\bar{}}$, in the lumpy-background case [13]).

In order to calculate the acceptance probability, either the posterior or prior density needs to be modeled, since these quantities are unknown. With the aforementioned assumption, we propose to model the posterior density $p(v^{\text{\bar{}}} | \mathbf{v})$ in Eq. (15) to be a multivariate Gaussian with mean $v_b$ and covariance $K^{\text{post}}_{v_b}$, where $v_b$ is the channelized version of the noiseless background $b$ for the given data $g$. The covariance, $K^{\text{post}}_{v_b}$, needs to be chosen so that sampling efficiency of the CIO-MCMC is good enough for a Markov chain to converge. We shall discuss this aspect later in Section 4. With this posterior model, the CIO likelihood ratio in Eq. (10) can be estimated via either a general Monte Carlo sampling directly from the posterior or the MCMC sampling from the proposal density. In this work, we focus on the MCMC simulation.

Another expression for calculating the acceptance probability, given in Eq. (16), requires knowledge of the prior on the channelized background, which is often unknown, as well as the conditional probability of channelized data on the channelized background, which is governed by the system’s noise statistics as discussed earlier. Channelized-background samples drawn by the proposal density can be far from the channelized background for the given image, which would lead to a low value of the conditional probability and hence an increased chance of rejecting these samples. Therefore, such samples may not be as appropriate for use in constructing a Markov chain to estimate the CIO likelihood ratio of $v$ as other samples that are closer to the given channelized background. However, since the prior also plays a significant role in the acceptance probability, if the prior is not modeled appropriately, bad samples still can be accepted, which may hurt the convergent property of the Markov chain. As an attempt to improve sampling efficiency, we can use the aforementioned observation in modeling the unknown prior density. That is, we propose to model the prior to be another Gaussian with covariance $K^{\text{prior}}_{v_b}$ and to be centered at the channelized background, $v_b$, of the given data, $g$. Note that we use “prior” with double quotation to indicate that the prior model we use in this work is not a true prior.

**D. Consistency Checks for CIO-MCMC**

To check the bias of the estimates of the CIO likelihood ratio, calculated by the CIO-MCMC, a number of things can be considered. One is to plot the moment-generating function for $\Lambda$ under $H_0$, $M_0(\beta)$, given by [11]:

$$
M_0(\beta)=\int_0^\infty d\Lambda \Lambda^\beta p(\Lambda|H_0)=\left\langle \Lambda^\beta \right\rangle_0.
$$

(17)

For the likelihood ratio, $\Lambda(v)$, this function has to be concave upward going through unity at $\beta = 0$ and $\beta = 1$. In addition, using the moment-generating function under $H_1$, it can be shown that $\left\langle \Lambda \right\rangle_1 - \text{var}(\Lambda)_0 = 1$. 


Lower and upper bounds for AUC estimates of the CIO can also be calculated by using the method proposed by Clarkson [12]. That is, using the following definition for the likelihood-generating function

\[ G(\beta) = \frac{\ln \left[M_0(\beta + \frac{1}{2})\right]}{(\beta^2 - \frac{1}{4})}, \]  

(18)

upper and lower bounds on the AUC of the CIO are given by

\[ 1 - \frac{1}{2} \exp \left[-\frac{1}{2} G(0)\right] \leq \text{AUC}_\Lambda \leq 1 - \frac{1}{2} \exp \left[-\frac{1}{2} G(0) - \sqrt{G(0) - \frac{1}{8} G''(0)}\right]. \]  

(19)

4. SIMULATION STUDY

A. Models for Imaging System, Object, Noise, and Channels

The lumpy-background model [13] was used for simulating 64 × 64 randomly varying background objects with the mean number of Gaussian lumps \( N = 5 \), lump (spatial) width of 7, and lump magnitude of 1. Since images are 64 × 64, the dimension of each image vector is \( M = 64^2 \). For the signal object, a Gaussian profile with (spatial) width 3 and magnitude 0.2 was used. A continuous-to-discrete linear imaging operator with Gaussian blur functions with (spatial) width 0.5 and magnitude 40 was applied to simulate noiseless signal and background images, \( s \) and \( b \). See [5] for the details of the imaging operator and image simulation. For the noise, \( n \), an independent, identically distributed (i.i.d.) multivariate Gaussian model with (amplitude) width \( \sigma_n = 20 \) was used. See Fig. 2 for the Gaussian signal, a signal-absent lumpy background, and its resulting images by use of the continuous-to-discrete operator and Gaussian and Poisson noise models.

The lumpy-background images generated by the afore-mentioned method do not always follow a Gaussian distribution. One way to check is to look at a histogram of an ensemble of lumpy-background images as shown in Fig. 3. This figure shows two histograms of 5000 lumpy-background images with Gaussian and Poisson noise realizations. Both histograms reveal that the lumpy-background images are non-Gaussian distributed, though at different degrees. Mean, minimum, and maximum of all pixel values used in the histogram for the Gaussian noise case are 12, −106, and 223, whereas they are respectively 12, 0, and 192 for the Poisson noise case. The difference in the lengths of the tails from the mean in each histogram reveals how much non-Gaussian behavior each data set has. Therefore, the background statistics in our simulation are not exactly Gaussian for the chosen level of Gaussian noise with respect to the background parameters even if a higher level of Gaussian noise can cause the background to be Gaussian distributed.

With use of the Gaussian noise statistics of \( n \), the conditional PDFs of channelized data given the channelized background are

\[
\text{pr}(v|v_b,H) = \frac{\exp \left(-\frac{1}{2\sigma_n^2} [v - v_b - jv_s]^T [TT^T]^{-1} [v - v_b - jv_s]\right)}{\sqrt{(2\pi)^N \det(\sigma_n^2 TT^T)}}, j = 0, 1.
\]  

(20)

Then, the likelihood ratio for the CBKE case, \( \Lambda_{\text{CBKE}}(v|v_b) \), is
where $T$ is the channel matrix consisting of $N_c$ rows of $M$-dimensional channel vectors.

For our choice of efficient channels, the 80 singular vectors most strongly associated with the signal, reproduced from [7], were used. See Fig. 4 for the first 100 of these singular vectors in a 2D profile. These channels were found to be highly efficient for the ideal observer in the same task used in our simulation [7].

**B. Proposal, Posterior, and Prior Densities for CIO-MCMC**

For the proposal density, $q(v_b^{(s)} | v_b^{(i)})$, to sample $v_b^{(s)}$ given the previously accepted sample $v_b^{(i)}$, we assumed that the proposal was an i.i.d. Gaussian with width $\sigma_c$ and centered at $v_b^{(i)}$. In particular, $\sigma_c = 19/10$, 1/32 were chosen, respectively, for the posterior and prior model cases of the CIO-MCMC.

For the posterior density, $p_r(v_b^{(i)} | H_j)$, we assumed a Gaussian posterior with covariance $K_{v_b}^{post}$ and mean $v_b$; i.e., the posterior was centered at the true channelized background $v_b$.

For calculating the covariance of the Gaussian posterior, we assumed $K_{v_b}^{post} = \sigma_b^2 I_{N_c}$, where each component of $\sigma_b$ was estimated as 10 times the sample standard deviation of each element of the channelized lumpy background using many different realizations of the noiseless lumpy-background image. In the simulation, 5000 noiseless lumpy backgrounds were used for this calculation. See Fig. 5 for the sample standard deviations of the channelized background used in this simulation.

For the Gaussian “prior” density, $p(v_b^{(i)})$, we again assumed that it was centered at the channelized version of the given background, $v_b$. For calculating the covariance of this prior, we assumed $K_{v_b}^{prior} = \sigma_c^2 I_{N_c}$, where $\sigma_c$ contained 1/6 times the same sample standard deviation of each element of the channelized background. Note that the particular values for $\sigma_c$ and $\sigma_b$ were chosen after investigating a number of different choices by use of the consistency checks discussed in Section 3.D.

In Table 1, we summarize a complete list of all the parameter values used for modeling the proposal, posterior, and prior densities. In Section 5, we show that both the posterior and “prior” models enable the CIO-MCMC to well approximate ideal-observer performance. In Appendix B, we discuss how the CIO-MCMC algorithm with our “prior” model actually works as another MCMC method with a different posterior density.

**C. Observer Performance, Variance Analysis, and Consistency Checks**

To generate a Markov chain for calculating one likelihood ratio, we used $J = 400,000$, $J_0 = 200,000$ to choose stable samples for both the posterior and the prior model cases. These numbers were determined after many simulations with different values of $J$ and $J_0$. To estimate the AUC of either of the ideal and MCMC-CIO observers, the calculation was done for the same set of 200 pairs of signal-absent and signal-present lumpy images. This calculation was repeated five times, using five different random-number seeds for generating five different Markov chains, resulting in five different AUC estimates. The arithmetic mean of the five AUC estimates of either of the observers was used as the AUC.
of either observer. To estimate the variance of the AUC estimates, a multiple-reader multiple-case variance analysis based on a probabilistic approach [14], in particular the one-shot method by Gallas [15], was employed.

As discussed in Subsection 3.D, we performed the following consistency checks to see whether the CIO likelihood ratio estimates calculated by the CIO-MCMC satisfy the properties of the likelihood ratio: (1) plotting $\Lambda_{CBKE}(v|v_b)$ as a function of $v_b$ to show the random nature of the samples used for estimating the CIO likelihood ratio, $\hat{\Lambda}(v)$; (2) $\Lambda^*(v)$ as a function of the number of samples, $J$, to show the convergent property of $\hat{\Lambda}(v)$; (3) $M_0(\beta)$ as a function $\beta$ to show the stability of our CIO-MCMC algorithms; (4) $\langle \Lambda \rangle_1 - \text{var}(\Lambda)_0$, using 10,000 pairs of signal-present and signal-absent estimates of the likelihood ratio, $\hat{\Lambda}(v|H_1)$ and $\hat{\Lambda}(v|H_0)$; and (5) calculating upper and lower bounds on the performance of the CIO.

D. Analytical Expressions for the CIO Likelihood Ratio

In our simulation, either the posterior of or prior on the channelized background was assumed to be Gaussian. In addition, the noise was assumed to follow an i.i.d. Gaussian distribution with zero mean and width $\sigma_n$. Use of these Gaussian models allows for analytical expressions for the CIO likelihood ratio via the convolution theorem. That is, a convolution of two Gaussians with means, $\mu_1, \mu_2$, and variances, $\sigma_1^2I, \sigma_2^2I$, is another Gaussian with mean $\mu_1 + \mu_2$ and variance $(\sigma_1^2 + \sigma_2^2)I$.

Under the assumption of the Gaussian posterior and noise statistics, $\Lambda_{CBKE}(v|v_b)$ in Eq. (10) can be replaced with Eq. (21), and $\text{pr}(v_b|v,H_0)$ in Eq. (10) with the Gaussian with mean $v_b$ and covariance $\sigma_n^2I_n$. Then, the integral given in Eq. (10) can be calculated analytically, resulting in the following expression for the CIO likelihood ratio:

$$
\Lambda(v) = \sqrt{\frac{N}{\pi}} \prod_{k=1}^{N} \sigma_{\upsilon,k} \exp \left( \sum_{k=1}^{N} \frac{v_{b,k} (u_k - u_{b,k})}{\sigma_{\upsilon,k}^2} + \sum_{k=1}^{N} \left( \frac{v_k (u_k - u_{b,k} - \frac{1}{2} v_{s,k})}{\sigma_n^2} \right) \right),
$$

(22)

where $u_k, v_{b,k}, v_{s,k}$, and $\sigma_{\upsilon,k}$ are, respectively, the $k$th elements of the vectors $v, v_b, v_s$, and $\sigma_{\upsilon}$. Note that the orthogonality of the SVD channels was used to derive this expression (PLS channels are also applicable, owing to their orthogonality).

Under the assumption of the Gaussian prior and noise statistics, the probability of channelized data $v$ under $H_j, j = 0, 1$, can be estimated via

$$
\text{pr}(v|H_j) = \int dv'_y \text{pr}(v'|v_y, H_j) \text{pr}(v'_y),
$$

(23)

which can be regarded as a convolution of two Gaussians (one with mean $j * v_s$ and variance $\sigma_n^2I_n$, and the other with mean $v_b$ and variance $\sigma_b^2I_b$). Via the convolution theorem, this results in $\text{pr}(v|H_j)$ being another Gaussian with mean $j * v_s + v_b$ and variance $(\sigma_n^2 + \sigma_b^2)I_n$. Thus, the resulting expression for the CIO likelihood ratio in this case is given by

$$
\Lambda(v) = \prod_{k=1}^{N} \exp \left( \frac{v_{s,k} (u_k - u_{b,k} - \frac{1}{2} v_{s,k})}{(\sigma_n^2 + \sigma_{\upsilon,k}^2)} \right).
$$

(24)
Table 2 summarizes these analytical expressions for the CIO likelihood ratio. These analytical expressions were used to compare against the CIO-MCMC method. They were also used to explore the relationship between the Gaussian widths and the AUC of the CIO. In our simulation, for computing the CIO likelihood ratio via these analytical expressions, another 1000 pairs of signal-present and signal-absent lumpy-background images, which were independent of the 200 pairs used for the CIO-MCMC and ideal observer, were used.

5. RESULTS AND DISCUSSION

A. Performance of CIO-MCMC Method

Figures 6 and 7, respectively, show the relationship of the width parameters of the Gaussian posterior and the prior to the performance of the CIO estimated by the analytical method. While the other consistency checks provided ways to locate and narrow down reasonable ranges of these width parameters, the analytical method, when available, can be used to study how the width parameters affect the outcome of the CIO performance and hence to confirm that the parameters chosen by the other checks were indeed appropriate. The chosen values of $C_\nu = 10$ and $1/6$ used for the posterior and “prior” CIO-MCMC simulations are consistent with the outcomes shown in these figures. We note that the target performance for the CIO was the performance of the true ideal observer, $0.91 \pm 0.02$, as shown in Table 3.

Table 3 summarizes the maximum CIO performance achieved by the extended MCMC [5], CIO-MCMC, and analytical expressions given in Eqs. (23) and (24) in comparison with ideal-observer performance estimated by the Kupinski MCMC. For the CIO results, 80 SVD channels associated with the signal were used. As indicated in Table 3, both the CIO-MCMC and the analytical CIO provide good approximations to the performance of the ideal observer and the CIO estimated by the extended MCMC and reproduced from [7]. In addition, we computed the AUC of the analytical CIO for the same 200 image pairs that were used for the CIO-MCMC and ideal observer. In this case, the AUCs of the analytical CIO for the posterior and prior models were $0.88 \pm 0.03$ and $0.92 \pm 0.03$ in the format of $\text{AUC} \pm 2 \text{STD}$, where $\text{AUC}$ and STD represent the sample mean AUC and the standard deviation. These are also good approximations to ideal-observer performance.

Figures 8 and 9 each present four empirical ROC curves generated using the 1000 estimates of the likelihood ratio for the ideal observer and the CIO by the extended MCMC, MCMC-CIO, and the analytical method. For Fig. 8, the Gaussian posterior was used in the CIO-MCMC and the analytical MCMC, whereas the Gaussian “prior” was used in the algorithm for Fig. 9.

These figures reveal how well, across all false positive rates, the CIO estimated by the CIO-MCMC and analytical CIO predicts the ideal observer. It appears that the posterior CIO-MCMC predicts the ideal observer equally well across all false positive rates, whereas in the range $(0.05,0.25)$ of false positive rates, the “prior” CIO-MCMC overestimates the ideal observer. The analytical CIO predicts the ideal observer well across all false positive rates for both posterior and prior cases.

For the reduction of computation times, the CIO-MCMC proved to be much faster than the Kupinski MCMC and the extended MCMC. For our simulation study, a computer cluster, which consisted of 56 dual-core dual-processor Opterons running at 2 GHz with 8 Gbytes of RAM, was used. To calculate the five AUC estimates of the CIO, using 80 independent cores on this cluster, the CIO-MCMC took only a few minutes, while the Kupinski MCMC and its extension, for the ideal observer and the CIO, respectively, required several hours.
B. Consistency Checks on the CIO-MCMC Method

Figures 10–13 present four different consistency checks for the posterior CIO-MCMC case, whereas Figs. 14–17 are for the “prior” CIO-MCMC case.

Figures 10 and 14 show estimates of $\Lambda_{CBKE}(v)$ as a function of $J$ for $4 \times 10^5$ states of a Markov chain for the 80 SVD channel cases used in Table 3. In each figure, the random nature of the $\Lambda_{CBKE}$ values indicates that the samples of $\Lambda_{CBKE}(v)$ were randomly chosen for estimating a particular likelihood-ratio estimate, $\hat{\Lambda}(v)$. Figures 11 and 15 each show that an estimate of $\hat{\Lambda}(v)$ converges well after a certain number of iterations, $J$.

Figures 12 and 16 each show a plot of the moment-generating function, $M_0(\beta)$, of the CIO likelihood ratio, $\hat{\Lambda}(v)$, for the 80 SVD channel cases used in Table 3. We observed that the estimates of the CIO likelihood ratio fairly well satisfy the properties of $M_0(\beta)$, such as the upward concavity for the CIO likelihood ratio discussed in Subsection 3.D. However, the property of $M_0(1) = 1$ was better satisfied with the “prior” CIO-MCMC than with the posterior CIO-MCMC, although even for the posterior case the value of $M_0(1)$ was still not blown up to such a large quantity, which was often seen in other density choices discussed in the next section. Figures 13 and 17 show $\langle \Lambda \rangle_1 - \text{var}(\Lambda)_0 = 1$ as a function of the number of samples, $J$, for estimating $\Lambda$. These figures reveal that the property of $\langle \Lambda \rangle_1 - \text{var}(\Lambda)_0 = 1$ is not well satisfied.

Finally, the bounds of the CIO AUC for the posterior and “prior” CIO-MCMC algorithms were (0.59,0.92) and (0.82,0.96), respectively, which well include the mean AUCs of 0.91 and 0.92, respectively. This calculation was done with the 1000 likelihood-ratio estimates for the five MCMC runs of the 200 pairs of images used in our simulation.

C. Other Choices for Posterior and Prior Densities

The “prior” choices suggested in this work may not be optimal, as the true prior density should be wide enough to sufficiently cover distributions of any random samples of the channelized background. Therefore, one of the more theoretically reasonable choices for the prior is to use the overall mean and covariance of the channelized background for many different realizations of the channelized background. We considered this option with 5000 different channelized noiseless background images. However, using the “prior” with the overall mean and covariance in the CIO-MCMC resulted in inappropriate selection of samples for constructing a Markov chain, which again led to nonconvergence of the Markov chain using millions of samples. Therefore, we did not include the results using this choice since CIO performance estimated by this choice was not close enough to the performance of the ideal observer.

Another choice is to use a flat prior. Then the priors in both the numerator and the denominator of the acceptance probability expression given in Eq. (16) would be canceled. We tried this approach, but did not get good results for reasons similar to those discussed above, and hence the results are not included in this paper.

We also tried independent but differently distributed (i.d.d.) Gaussian proposal models combined with either an i.d.d. Gaussian posterior or prior model. However, this approach did not yield a good approximation to the ideal observer AUC. In particular, samples of $\Lambda_{CBKE}(v)$ were biased, and the estimates of $\Lambda$ as a function of the number of samples did not converge. We observed the AUC of the CIO using this approach decrease with the number of samples, $J$. This indicates that the approach does not provide stable, efficient sampling for the CIO-MCMC.
6. CONCLUSIONS

We have proposed a fast Markov-chain Monte Carlo method, the CIO-MCMC, for efficiently calculating ideal-observer performance via a novel way of sampling channelized-background vectors rather than background images. In addition, in the case where the system’s noise statistics can be modeled to be Gaussian, our approach provides analytical expressions for the CIO likelihood ratio, which yield good approximations to ideal-observer performance. The analytical CIO, while linear, is a new observer that outperforms the best linear observer, the channelized Hotelling observer, and estimates the ideal observer.

The CIO-MCMC requires far less computer time (a few minutes as opposed to several hours of the Kupinski MCMC computation), thereby providing a potentially practical tool for real-time computation of ideal-observer performance. Furthermore, the analytical CIO provides another method for real-time computation of ideal-observer performance in the case of Gaussian system noise statistics.

Our approach uses noise statistics of the given imaging system in estimation of the CIO likelihood ratio. Therefore, with knowledge of noise statistics of an imaging system of interest, the CIO-MCMC has the potential to speed up the computation of ideal-observer performance in classification tasks involving complex or realistic background images in clinical applications.

7. FUTURE WORK

We noticed that the widths of the Gaussian proposal and prior densities, $\sigma_c$ and $\sigma_p$, used in the CIO-MCMC significantly affected how well the performance of the CIO, estimated by the CIO-MCMC, approximates the performance of the ideal observer and CIO, respectively, estimated by the Kupinski MCMC and its extension that sample background images. The consistency checks performed in this work provided reasonable pointers to where to look for locating a neighborhood of the proposal and posterior width parameters for the CIO-MCMC to yield good approximations to ideal-observer performance. However, for the CIO-MCMC to be used in tasks involving different types of images and non-Gaussian noise, further investigations are necessary to find a method of choosing appropriate width parameters that bring good sampling efficiency and approximations to ideal-observer performance.

With the models used for our MCMC simulation, the property of $\left\langle \Lambda \right\rangle_1 - \text{var}(\Lambda)_0 = 1$ was not well satisfied. This may be because of modeling multimode Gaussian marginal distributions with single-mode Gaussian models. Further investigations by use of different density models are necessary for improving the property of the Markov chain to satisfy the aforementioned condition as well as the other checks executed in our simulation.

We are also interested in applying our method to classification tasks involving different types of backgrounds such as in planar and volumetric x-ray breast imaging for estimating ideal-observer performance. Last, while our investigation using a Gaussian noise model can be useful in various imaging modalities, including x-ray imaging systems, further investigations using other noise models including Poisson noise would benefit other imaging applications such as nuclear medicine. Therefore, we are interested in expanding our method to cases with noise models other than Gaussian.

APPENDIX A: REPLACING $\theta$ WITH $v_b$ IN THE LIKELIHOOD RATIO

To show that Eqs. (6) and (10) are equivalent, we use a standard method of transforming density functions given by
where $N_v$ and $N_\theta$ are the dimensions of $v$ and $\theta$, respectively. Note that $\text{pr}(Tv|v,H_0) = \text{pr}(T\theta|v,H_0)$, since $T$ is a scalar matrix and $b$ is fully governed by the vector $\theta$. Then, by use of Eq. (A1), Eq. (10) leads to Eq. (6) as follows:

$$\Lambda(v) = \int dv_b \Lambda_{\text{HKL}}(v|v_b) \text{pr}(v_b|v,H_0),$$

(A3)

$$= \int dv_b \int d\theta \Lambda_{\text{HKL}}(v|v_b) \delta^{N_v}(v_b - T\theta) \text{pr}(\theta|v,H_0)$$

(A4)

$$= \int d\theta \Lambda_{\text{HKL}}(v|T\theta) \text{pr}(\theta|v,H_0)$$

(A5)

$$= \int d\theta \Lambda_{\text{HKL}}(v|\theta) \text{pr}(\theta|v,H_0).$$

(A6)

The last equality is due to the fact that $Tv$ is fully represented by the parameter vector $\theta$. The other way [from Eq. (6) to Eq. (10)] is done similarly by using Eq. (A2).

**APPENDIX B: “PRIOR” CIO-MCMC AS ANOTHER POSTERIOR CIO-MCMC**

In this subsection, we show that the “prior” CIO-MCMC is equivalent to another MCMC method with a posterior density model. For ease of notation, we define the matrices $A$ and $B$ via $A = \sigma_a^2 T T^+$ and $B = c a_2 I_N$.

For the posterior CIO-MCMC method, the Gaussian posterior model is given by

$$\text{pr}(v_b|v) = \text{Gauss}(v_b, B).$$

(B1)

For the “prior” CIO-MCMC method, the probability of $v$ conditional on $v_b$ is given by

$$\text{pr}(v|v_b) = \text{Gauss}(v_b, A),$$

(B2)

and the Gaussian prior model is given by

$$\text{pr}(v_b) = \text{Gauss}(v_b, B).$$

(B3)

With Eqs. (B2) and (B3), the resulting posterior, $\text{pr}(v_b|v)$, for the “prior” CIO-MCMC method is given by
where $C$ and $C'$ do not depend on $v_b^*$ and the quadratic quantity $Q$ is given by

$$Q = (v - v_b^*)^T A^{-1} (v - v_b^*) + (v_b^* - v_b)^T B^{-1} (v_b^* - v_b).$$  \tag{B5}

This expression, $Q$, expands into

$$v_b^* (A^{-1} + B^{-1}) v_b^* - v_b^* (A^{-1} v + B^{-1} v_b) - [v_b^* (A^{-1} v + B^{-1} v_b)]^T v + v_b^* v + v_b^* B^{-1} v_b - v_b^* A^{-1} v.$$ \tag{B6}

Now consider a similar quadratic expression

$$Q' = (v_b^* - x) (A^{-1} + B^{-1}) (v_b^* - x),$$ \tag{B7}

which expands into

$$v_b^* (A^{-1} + B^{-1}) v_b^* - v_b^* (A^{-1} + B^{-1}) x - [v_b^* (A^{-1} + B^{-1})] x + x^T (A^{-1} + B^{-1}) x.$$ \tag{B8}

If we set $x = (A^{-1} + B^{-1})^{-1} (A^{-1} v + B^{-1} v_b)$, then $Q$ and $Q'$ differ only by terms that do not depend on $v_b^*$. This means that the “prior” CIO-MCMC method is equivalent to an MCMC method using the posterior

$$\text{pr}(v_b^* | v) = \text{Gauss}(x, K),$$ \tag{B9}

where $K = (A^{-1} + B^{-1})^{-1}$. The mean of this posterior, $x$, is a matrix-weighted sum of the data $v$ and the true background $v_b$. This explains why the “prior” CIO-MCMC works fairly similarly to the posterior CIO-MCMC.

**Acknowledgments**

The authors thank Drs. Laura Thompson and Frank Samuelson at the FDA for their comments on the manuscript. This work is supported in part by the National Institute of Biomedical Imaging and Bioengineering at the National Institutes of Health intramural grant to the Center for Devices and Radiological Health, FDA.

**REFERENCES**


8. Witten J, Park S, Myers KJ. Using partial least square to compute efficient channels for the Bayesian ideal observer. Proc. SPIE 2009;7263:72630Q.


Fig. 1.
(Color online) Top, overall histogram of the fifth element, $v_5$, of the channelized background generated using 5000 random lumpy backgrounds of $N = 5$. Bottom, histogram of $v_5$ of a given $v_b$, generated using 30,808 sample backgrounds obtained via the Park extended MCMC. In this case, the true $v_5$ of the given channelized background, $v_b$, for generating the channelized data, $v$, lies in the leftmost histogram.
Fig. 2.
Top, Gaussian signal object and one realization of the signal-absent lumpy object. Bottom, images (scaled by using the minimum and maximum of all image pixel values in both images) of the lumpy object with additive Gaussian noise (left) and Poisson noise (right). For the lumpy object, the magnitude and the spatial width of the Gaussian lump are 1 and 7. For the Gaussian signal, the magnitude and the spatial width of the signal are 0.2 and 3. For Gaussian noise, the amplitude width $\sigma_n$ is 20.
Fig. 3. (Color online) Histograms of 5000 lumpy-background images (top) using the same parameters as in the simulation and (bottom) using the same parameters except a Poisson distribution with each image pixel being the mean was used to realize the noise. The difference in the lengths of the tails from the mean of each histogram reveals how much non-Gaussian behavior each data set has.
Fig. 4.
Fig. 5.
(Color online) Sample standard deviations of each of all the pixels in the channelized background $\psi_b$. 
Fig. 6.
Performance (AUC) of the analytical CIO with the Gaussian posterior model of width $\sigma_0 (= C_0 \text{ STD of } v_p)$ as a function of $C_0$ and the number of channels $N_c$. STD represents the sample standard deviation. In the detection task, 1000 pairs of signal-present and signal-absent were used. In this plot, $N_c = 5, 10, 20, 30, 40, 50, 60, 70, 80$. Note that the target performance for the CIO is the performance of the ideal observer, $0.91 \pm 0.02$, as shown in Table 3.
Fig. 7.
The performance (AUC) of the analytical CIO with the Gaussian “prior” model of width $\sigma_v$ ($=C_v$ STD of $v_b$) as a function of $C_v$ and the number of channels $N_c$. STD represents sample standard deviation. In the detection task, 1000 pairs of signal-present and signal-absent were used. In the plot, $N_c = 5, 10, 20, 30, 40, 50, 60, 70, 80$. Note that the target performance for the CIO is the performance of the ideal observer, 0.91±0.02, as shown in Table 3.
Fig. 8.
For the posterior CIO-MCMC, four empirical ROC curves obtained by using the likelihood ratio estimates for the ideal observer by the Kupinski MCMC and the CIO by the extended MCMC, CIO-MCMC, and analytical CIO. For all the likelihood ratios obtained by the MCMC algorithms, the same set of 200 pairs of images were used with five different Markov chains. For the likelihood ratios obtained by the analytical CIO method, another 1000 pairs of images were used. For the CIO, $N_c = 80$ was used.
For the “prior” CIO-MCMC, four empirical ROC curves obtained by using the likelihood ratio estimates for the ideal observer by the Kupinski MCMC and the CIO by the extended MCMC, CIO-MCMC, and the analytical CIO. For all the likelihood ratios by the MCMC algorithms, the same set of 200 pairs of images were used with five different Markov chains. For the likelihood ratios by the analytical CIO method, another 1000 pairs of images were used. For the CIO, $N_c = 80$ was used. Note that the ROC curve for CIO: MCMC-CIO appears above the ideal-observer ROC because of statistical fluctuation.
Fig. 10. (Color online) For the posterior CIO-MCMC, the plots show $\log(\Lambda_{CBKE})$ versus states of a Markov chain generated by the CIO-MCMC for the case of $N_c = 80$. Iteration number represents each state. The bottom plot shows the first 10,000 states in the top plot.
Fig. 11. (Color online) For the posterior CIO-MCMC, the plot shows $\hat{\Lambda}(\nu)$ as a function of the number, $J$, of $\Lambda_{CBKE}(\nu|\nu_b)$ samples for the case of $N_c = 80$. In this plot, $J = 200k$. 
Fig. 12.
(Color online) For the posterior CIO-MCMC, the plot shows the moment-generating function, $M_0(\beta)$, using 10,000 estimates of $\hat{\Lambda}(v|H_0)$, calculated by the CIO-MCMC and the analytical method for the case of $N_c = 80$. 

\[ M_0(\beta) = \sum_{v=0}^{\infty} \exp\left(\beta v - \lambda v\right) P(v) \]
Fig. 13.
(Color online) For the posterior CIO-MCMC, the plot shows $\langle \Lambda|H_1 \rangle - \text{var}(\Lambda|H_0)$ as a function of the number of samples, $J$, using 10,000 pairs of $\hat{\Lambda}(v|H_0)$ and $\hat{\Lambda}(v|H_1)$ estimated by the CIO-MCMC using the Gaussian posterior model for the case of $N_c = 80$. 
Fig. 14. (Color online) For the “prior CIO-MCMC,” the plots show $\log(\Lambda_{\text{CBKE}})$ versus states of a Markov chain generated by the CIO-MCMC for the case of $N_c = 80$ in Table 3. Iteration number represents each state. The bottom plot shows the first 1000 states in the top plot.
Fig. 15.
(Color online) For the “prior” CIO-MCMC, the plot shows $\hat{\Lambda}(v)$ as a function of the number, $J$, of $\Lambda_{\text{CBKE}}(v|y_b)$ samples for the case of $N_c = 80$ in Table 3. In this plot, $J = 200k$. 
Fig. 16. (Color online) For the “prior” CIO-MCMC, the plot shows the moment-generating function, $M_0(\beta)$, using 10,000 estimates of $\hat{\Lambda}(v|H_0)$, calculated by the CIO-MCMC and the analytical method for the case of $\langle \Lambda|H_1 \rangle - \text{var}(\Lambda|H_0)$. 
Fig. 17.

(Color online) For the “prior” CIO-MCMC, the plot shows $\langle \Lambda | H_1 \rangle - \text{var}(\Lambda | H_0)$ as a function of the number of samples, $J$, using 1000 pairs of $\hat{\Lambda}(v|H_0)$ and $\hat{\Lambda}(v|H_1)$ estimated by the CIO-MCMC using the Gaussian prior model for the case of $N_c = 80$. 

$\langle \Lambda | H_1 \rangle$ and $\text{var}(\Lambda | H_0)$
Table 1
Parameters for the Proposal, Posterior, and Prior Models

<table>
<thead>
<tr>
<th></th>
<th>Posterior CIO-MCMC</th>
<th>“Prior” CIO-MCMC</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma_{c}$</td>
<td>$\frac{19}{10}$</td>
<td>$\frac{1}{32}$</td>
</tr>
<tr>
<td>$\sigma_{v}$</td>
<td>10 STDs of $v_{p}$</td>
<td></td>
</tr>
<tr>
<td>$\sigma_{t}$</td>
<td>$\frac{1}{6}$</td>
<td>STDs of $v_{p}$</td>
</tr>
</tbody>
</table>

$^a$ STD represents the standard deviation.
Table 2
Analytical Expressions of the CIO Likelihood Ratio with Gaussian Noise

Analytical CIO with Gaussian posterior:

\[
\Lambda(v) = \sqrt{N_c} \prod_{k=1}^{N_c} \sigma_{v,k} \exp \left( \sum_{k=1}^{N_c} \frac{\nu_{v,k}(v_k - v_{b,k})}{\sigma_{v,k}^2} + \sum_{k=1}^{N_c} \frac{\nu_{s,k} \left( v_k - v_{b,k} - \frac{1}{2} v_{s,k} \right)}{\sigma_n^2} + \sum_{k=1}^{N_c} \left( \frac{\nu_{s,k} \sigma_{v,k}^2}{4\sigma_n^4} \right) \right)
\]

Analytical CIO with Gaussian “prior”:

\[
\Lambda(v) = \prod_{k=1}^{N_c} \exp \left( \frac{\nu_{s,k} (v_k - v_{b,k} - \frac{1}{2} v_{s,k})}{(\sigma_n^2 + \sigma_{v,k}^2)} \right)
\]
Table 3

Comparison of Observer Performance<sup>a</sup>

<table>
<thead>
<tr>
<th>Observer: Algorithm</th>
<th>AUC ± 2 STD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ideal observer: Kupinski MCMC</td>
<td>0.91 ± 0.02</td>
</tr>
<tr>
<td>CIO: extended MCMC</td>
<td>0.90 ± 0.02</td>
</tr>
<tr>
<td>CIO: posterior CIO-MCMC</td>
<td>0.91 ± 0.04</td>
</tr>
<tr>
<td>CIO: “prior” CIO-MCMC</td>
<td>0.92 ± 0.02</td>
</tr>
<tr>
<td>CIO: posterior analytical</td>
<td>0.90 ± 0.01</td>
</tr>
<tr>
<td>CIO: “prior” analytical</td>
<td>0.91 ± 0.01</td>
</tr>
</tbody>
</table>

<sup>a</sup>Each performance is in the format $\text{AUC} ± 2 \text{STD}$ where the standard deviation, STD, was estimated by the one-shot method [15].