

PROJETO E IMPLEMENTAÇÃO DE UM SISTEMA DE DETECÇÃO COM PROBABILIDADE DE FALSO ALARME AJUSTÁVEL

MARCOS VINÍCIUS SILVA OLIVEIRA*, BERNARDO SOTTO-MAIOR PERALVA*, LUCIANO MANHÃES DE ANDRADE FILHO*, AUGUSTO SANTIAGO CERQUEIRA*

*Laboratório de Processamento de Sinais e Telecomunicações
Programa de Pós-Graduação de Engenharia Elétrica
Universidade Federal de Juiz de Fora
Juiz de Fora, Minas Gerais, Brasil

Emails: marcos.oliveira@engenharia.ufjf.br, bernardo.peralva@ice.ufjf.br,
luciano.andrade@ufjf.edu.br, augusto.santiago@ufjf.edu.br

Abstract— This work proposes an implementation of a robust signal-to-noise detection system that ensures a constant level of contamination based on an adaptive threshold, even though the contamination level may change according to a set of constraints adopted. Traditional methods of signal-to-noise detection take into account a constant level of contamination to perform the classification of events. However, as the scenario may change, these methods can lead to higher false alarm rate since the level of contamination varies. By monitoring the noise variance in real-time, the proposed method implements a trigger system where its decision threshold is proportional to the noise standard deviation, therefore, it ensures a constant false alarm rate. The algorithm implemented in FPGA is presented as well as the simulation results for an optimum signal detector.

Keywords— Signal detection, adjustable probability of false alarm, scientific instrumentation and particle physics.

Resumo— Propõe-se neste trabalho a implementação de um sistema de detecção de sinais que, a partir de um limiar adaptativo, garante nível de contaminação constante, mesmo que o cenário se altere dentro de um conjunto de restrições adotadas. Os métodos tradicionais de detecção adotam um limiar constante para realizar a classificação dos eventos, no entanto, com a mudança de cenário, o nível de contaminação pode variar consideravelmente. Monitorando-se a variância do ruído de forma *online*, o método proposto implementa um sistema decisor cujo o limiar de decisão é proporcional ao desvio padrão do ruído, garantindo uma probabilidade de falso alarme constante. O algoritmo, implementado em FPGA, é apresentado, assim como os resultados de simulações para um detector ótimo de sinais.

Palavras-chave— Detecção de sinais, probabilidade de falso alarme ajustável, instrumentação científica e física de partículas.

1 Introdução

A evolução da instrumentação científica causa impacto em diversas áreas do conhecimento, proporcionando significativo desenvolvimento em diversos ambientes, como por exemplo, processos industriais, materiais na área da saúde, aplicações de bens de consumo, entre outros. Diversas áreas do conhecimento fazem uso da instrumentação científica para coletar, processar e tirar conclusões a respeito das grandezas de um determinado ambiente. Dentre elas, existe a física experimental que, a partir de uma extensa análise de dados, procura confirmar conclusões teóricas ou, em outros momentos, até descobrir efeitos ainda desconhecidos. Dentro deste vasto campo existe o estudo de física de partículas que procura detectar sinais presentes em um determinado experimento (The Atlas Collaboration, 2008). Normalmente estes sinais possuem baixa intensidade e surgem com rara ocorrência (de Andrade Filho, 2009).

Um dos principais desafios da instrumentação científica é separar os sinais de interesse dos demais, com o maior nível de certeza possível. Na busca deste objetivo, inúmeras aplicações científicas utilizam filtros que, em suma, são capazes

de distinguir sinal do ruído, quando condições necessárias são atendidas. Neste contexto, os Filtros Ótimos (Trees, 2001) são muito utilizados, uma vez que eles podem aumentar significativamente a SNR (do inglês *signal-to-noise ratio*) e beneficiar a detecção do sinal e a estimativa dos parâmetros em meio ao ruído. No caso destes sistemas, é desenvolvido um filtro baseado no conhecimento prévio das estatísticas do ruído e dos sinais envolvidos.

Em seguida, surge o processo de decisão, que utiliza um parâmetro chamado limiar de decisão para distinguir as amostras vinculadas ao banco de sinal, das amostras vinculadas ao banco de ruído. Durante este processo, é utilizado um simples critério que classifica como sinal, amostras que atinjam ou ultrapassem este limiar, e como ruído, amostras que não atinjam o mesmo. A Figura 1 descreve as etapas de um sistema de detecção de sinais contra ruído.

Apesar do sistema exposto na Figura 1 apresentar enorme aceitação por parte dos profissionais da área de processamento de sinais, em contextos onde a energia do ruído pode variar em função do tempo o sistema pode apresentar resultados indesejáveis. Estas incertezas refletem diretamente nos índices de eficácia do sistema discutido,

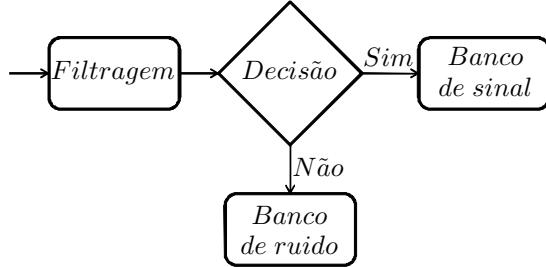


Figura 1: Representação do funcionamento do detector com decisor de limiar fixo.

entre eles, P_d e P_f , que denotam probabilidade de detecção e probabilidade de falso alarme, respectivamente. Pode-se perceber que para ambientes em que a energia do ruído é variante no tempo e heterogênea entre os canais, estes índices vão acompanhar tal diversidade e apresentarão flutuação em função destas propriedades estocásticas, sobre as quais o sistema não possui qualquer tipo de controle. Não possuindo controle perante as mesmas, resta somente intervir no processo de decisão que surge logo após o processo de filtragem.

Neste trabalho é proposto um detector que é capaz de redefinir seu limiar de decisão em função do desvio padrão do ruído de forma que se garanta constante a probabilidade de falso alarme. Na Seção 2, são apresentados os conceitos básicos de filtragem ótima. Na Seção 3, a teoria do método proposto neste trabalho é descrita em detalhes e na Seção 4, a implementação proposta do sistema é apresentada. Os resultados de simulações são mostrados na Seção 5. Finalmente, as conclusões são derivadas na Seção 6.

2 Detecção por Filtro Ótimo

A teoria de detecção de sinais é baseada no teste de hipóteses (Shanmugan, 1988). A partir do mesmo é realizado um teste em uma sequência de N amostras, correspondendo a um sinal temporal digitalizado. Neste trabalho, será abordado o ambiente onde o sinal de interesse é determinístico e o ruído aditivo, Gaussiano e branco. Neste cenário, deve-se escolher entre duas hipóteses: a hipótese H_0 , em que somente ruído eletrônico de média zero e distribuição Gaussiana n_i está presente, ou a hipótese H_1 , em que o sinal determinístico procurado s_i está adicionado ao ruído, conforme indicado na seguinte equação:

$$\begin{aligned} H_0 : \quad r_i &= n_i \\ H_1 : \quad r_i &= s_i + n_i \end{aligned} \quad (1)$$

em que r_i denota a sequência de entrada do processo de filtragem com amostras compreendidas no intervalo $1 \leq i \leq N$.

No intuito de decidir entre uma das duas hipóteses anteriores, utiliza-se a razão de semelhança, $\Lambda(\mathbf{r})$, que maximiza a eficiência de detecção para

o teste de hipóteses supracitado, como mostrado abaixo:

$$\Lambda(\mathbf{r}) = \frac{f_{\mathbf{R}|H_1}(\mathbf{r}|H_1)}{f_{\mathbf{R}|H_0}(\mathbf{r}|H_0)} \stackrel{H_1}{\gtrless} \stackrel{H_0}{\lessdot} \gamma \quad (2)$$

em que, $\mathbf{r} = [r_1, r_2, \dots, r_N]$ e os parâmetros $f_{\mathbf{R}|H_1}$ e $f_{\mathbf{R}|H_0}$ são as funções densidade de probabilidade (fdp) do sinal recebido \mathbf{R} dado que H_1 e H_0 ocorreram, respectivamente. Já o parâmetro γ está relacionado com a probabilidade de falso alarme (P_f) requerida. Desta forma a Equação (2) pode ser interpretada da seguinte maneira: se a razão $\Lambda(\mathbf{r})$ for maior que γ , escolhe-se por H_1 , senão, H_0 é escolhido. Supondo-se que a sequência i com distribuição Gaussiana e com suas amostras descorrelacionadas entre si, pode-se obter $f_{\mathbf{R}|H_1}$ e $f_{\mathbf{R}|H_0}$ a partir do produtório das distribuições individuais de cada amostra, conforme descrito na equação seguinte:

$$\Lambda(\mathbf{r}) = \frac{\prod_{i=1}^{i=N} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(r_i - s_i)^2}{2\sigma^2}}}{\prod_{i=1}^{i=N} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{r_i^2}{2\sigma^2}}} \stackrel{H_1}{\gtrless} \stackrel{H_0}{\lessdot} \gamma \quad (3)$$

Aplicando-se logaritmo natural nos dois termos da igualdade e efetuando algumas manipulações algébricas, obtém-se a seguinte equação:

$$\sum_{i=1}^{N} r_i s_i \stackrel{H_1}{\gtrless} \stackrel{H_0}{\lessdot} \gamma' \quad (4)$$

em que γ' corresponde ao limiar de decisão resultante que deverá ser avaliado na saída do filtro.

Com esse desenvolvimento, conclui-se que o resultado do filtro apresentado, pode ser representado simplesmente pelo produto interno entre a sequência adquirida pelo detector r_i e o sinal determinístico procurado s_i . Este filtro é conhecido como Filtro-Casado (Shanmugan, 1988).

3 Detecção com Limiar de Decisão Adaptativo

Nota-se que o projeto do limiar γ' , apresentado na Seção 2, está relacionado ao valor de $\sigma_{n'}$ e a relação entre P_d e P_f desejado. Em que, $\sigma_{n'} \triangleq \sqrt{E\{n'^2_i\}}$, representa o desvio padrão do ruído após a passagem pelo filtro casado e n'^i , representa as amostras de ruído após a passagem pelo mesmo filtro. No cenário em que $\sigma_{n'}$ varie com o tempo, haverão diferentes combinações para P_d e P_f .

Em algumas aplicações, se deseja que P_d , P_f ou a relação de maior eficiência entre P_d e P_f , seja controlável, ou seja, que a partir da observação de algum parâmetro do sistema, se firme o controle de alguma das variáveis citadas anteriormente. Este trabalho irá se atter em manter fixo o nível de contaminação da detecção, fixando P_f .

O diagrama de processos do sistema proposto com limiar de detecção adaptativo é ilustrado na Figura 2. Nota-se que existe a inclusão de um processo chamado *Estimação de $\sigma_{n'}$* , visto que o detector com decisão adaptativa necessita deste parâmetro para firmar o controle perante a variável P_f .

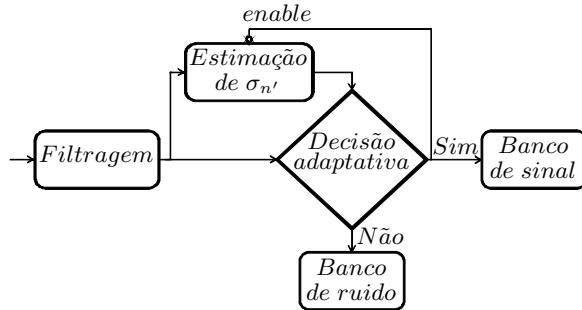


Figura 2: Representação do funcionamento do detector com decisior adaptativo.

Nota-se que P_f possui uma simples relação com $\sigma_{n'}$, de forma que um novo limiar adaptativo γ_{adapt} pode ser dado pela seguinte relação:

$$\gamma_{adapt} = k\sigma_{n'} \quad (5)$$

Dessa forma a Equação (4) pode ser reescrita conforme descrito abaixo:

$$\sum_{i=1}^N r_i s_i \frac{H_1}{H_0} \gtrless \gamma_{adapt} \quad (6)$$

Para a escolha de k deve-se fixar a P_f desejada. No intuito de relacionar k e P_f , inicialmente deve-se conhecer a Função Distribuição de Probabilidade Acumulativa (Papoulis, 1991) $F_{n'}(x)$ do ruído n'_i , que é representada pela expressão:

$$F_{n'}(x) = \frac{1}{\sqrt{2\pi\sigma_{n'}^2}} \int_{-\infty}^x e^{-\frac{\xi^2}{2\sigma_{n'}^2}} d\xi \quad (7)$$

A fim de se obter $F_{n'}(x)$ generalizada para variados valores de $\sigma_{n'}^2$, determina-se as Equações (8) e (9), em que $F(x)$ representa a Função Distribuição de Probabilidade normalizada (Papoulis, 1991).

$$F_{n'}(x) = F\left(\frac{x}{\sigma_{n'}}\right) \quad (8)$$

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{\xi^2}{2}} d\xi \quad (9)$$

Desta forma, é possível definir P_f para detector com decisior adaptativo (Trees, 2001), a partir da relação compreendida a seguir:

$$P_f(\gamma_{adapt}) = F_{n'}(-\gamma_{adapt}) \quad (10)$$

A partir das equações 5, 8 e 10, deduz-se a seguinte expressão:

$$\begin{aligned} P_f(\gamma_{adapt}) &= F\left(\frac{-\gamma_{adapt}}{\sigma_{n'}}\right) \\ &= F\left(\frac{-k\sigma_{n'}}{\sigma_{n'}}\right) \\ &= F(-k) \end{aligned} \quad (11)$$

Neste sentido, para uma determinada probabilidade de falso alarme desejada P_{fd} deve ser escolhido um k que obedeça a relação definida pelas expressões que se seguem.

$$\begin{aligned} F(-k) &= P_{fd} \\ 1 - F(k) &= P_{fd} \\ F(k) &= 1 - P_{fd} \end{aligned} \quad (12)$$

A fim de ilustrar o comportamento da função $F(k)$, é ilustrado na Figura 3 a relação entre P_{fd} e k , para $-4 \leq k \leq 4$.

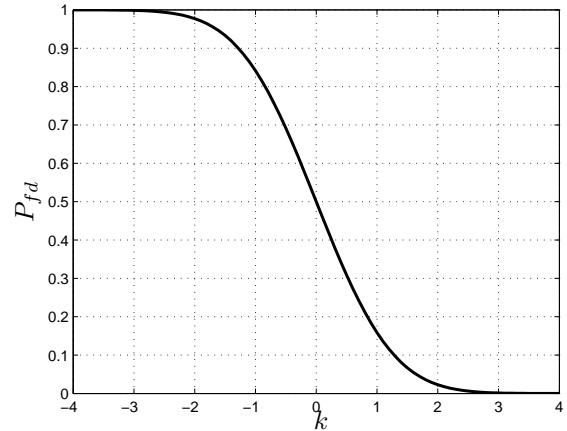


Figura 3: Relação entre P_{fd} e k correspondente.

4 Implementação

O detector proposto na Seção 3 foi implementado em um chip FPGA (Brown, 2000) do fabricante Altera, família Cyclone III e código de identificação EP3C40F484C6. O circuito sintetizado é mostrado no diagrama da Figura 2. Para o decisior foi adotado $k = 4$ (detecção em $4\sigma_{n'}$), de forma que conciliasse o par de P_d e P_f desejado e simplicidade de implementação, visto que a multiplicação por um número potência de 2 pode ser implementado facilmente a partir de um simples deslocamento de bit.

No processo de estimação de $\sigma_{n'}$, foi adotado um estimador conforme Figura 4, em que o parâmetro *trigger* está relacionado com o resultado da detecção. Desta forma, ele está ativo sempre quando um evento é classificado como sinal. No caso de eventos classificados como ruído, o parâmetro *enable* se mantém ativado, habilitando o incremento no acumulador.

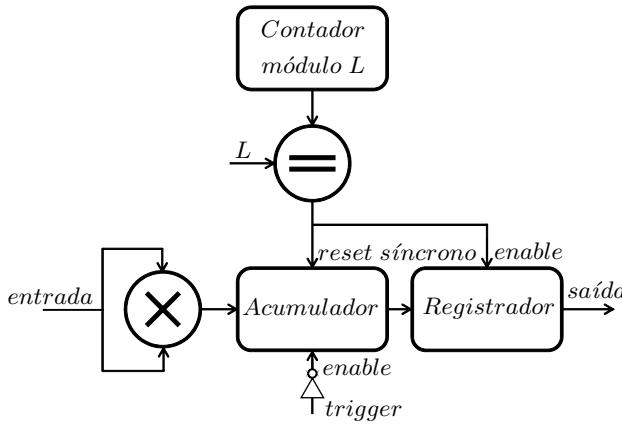


Figura 4: Circuito implementado para a estimativa de $\sigma_{n'}$.

Verifica-se, a partir da Figura 4, que o circuito de estimativa de $\sigma_{n'}$ faz uso de um multiplicador, calculando o quadrado do sinal de entrada, e um acumulador, em conjunto com um contador e um comparador configurados de forma que o acumulador reinicie a cada L eventos. Com isso, o estimador oferece na saída uma nova estimativa de $\sigma_{n'}^2$, multiplicada por L , a cada L eventos, em que o parâmetro *reset* da Figura 4 é ativado quando o limite L de eventos é alcançado.

Nota-se ainda a ausência do circuito de raiz quadrada e de divisão por L para a estimativa de $\sigma_{n'}$. Isto se deve pela estratégia de simplificação adotada, uma vez que o resultado na saída será utilizado apenas para realizar a comparação definida no processo de decisão. Neste sentido, verifica-se que é possível manipular o outro termo da comparação afim de se obter o mesmo resultado esperado. A manipulação mencionada pode ser numericamente representada pela Equação (13), para $L = 2^{10}$.

$$\begin{aligned} \sum_{i=1}^N r_i s_i &> 4\sigma_{n'} \\ \left(\sum_{i=1}^N r_i s_i \right)^2 &> 16\sigma_{n'}^2 \\ 2^6 \left(\sum_{i=1}^N r_i s_i \right)^2 &> 2^{10}\sigma_{n'}^2 \end{aligned} \quad (13)$$

Verifica-se que multiplicando o termo referente ao quadrado da saída do filtro por 2^6 , e comparando com $2^{10}\sigma_{n'}^2$, se equivale a comparação definida pela Equação (6), para $k = 4$, evitando os circuitos de divisão e radiciação antes necessários.

5 Simulação

Para realizar a simulação, foi adotada uma sequência de $N = 30$ amostras do sinal determinístico normalizado mostrado na Figura 5. Trata-se de uma forma de sinal típico, normalmente encontrado em experimentos de física de altas energias.

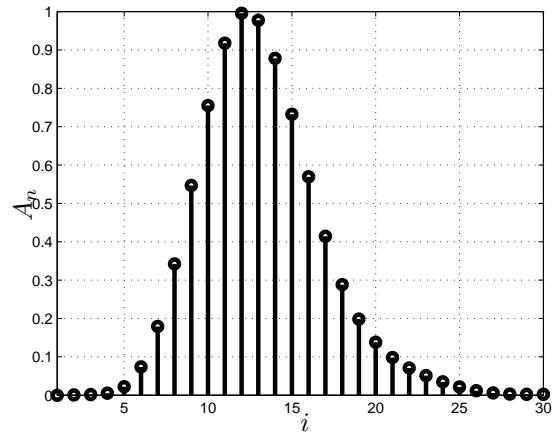


Figura 5: Sinal determinístico s_i , em que a amplitude normalizada A_n esta representada no eixo da ordenada e o índice de amostra i na abcissa.

O ruído possui característica estocástica Gaussiana, branca e de média zero, ou seja, $E\{n_k\} = 0$. O desvio padrão do ruído foi modelado de forma variante no tempo, possuindo intervalo $1 \leq \sigma_n \leq 3$, esta variação possuiu forma orientada pela imagem de $\sin(\omega)$, para $0 \leq \omega \leq \pi$, conforme Figura 6.

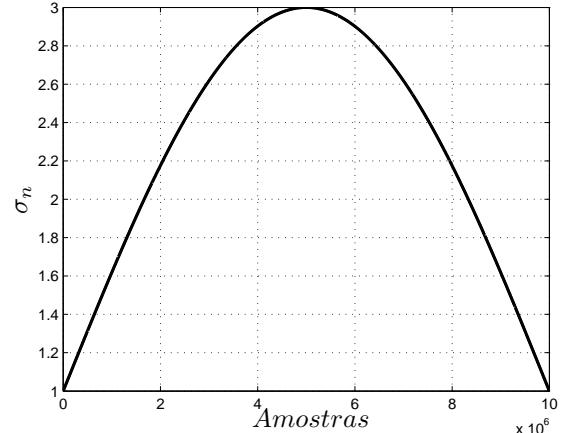


Figura 6: Ilustração da variação temporal sofrida pelo σ_n , em que σ_n esta representado no eixo da ordenada e o índice de amostra k na abcissa.

Vale ressaltar que o intervalo de atualização desta estimativa deve ser projetado de forma que o desvio padrão real permaneça estacionário para uma determinada janela de estimativa de $\sigma_{n'}$. A partir da observação do cenário proposto, verificou-se que um intervalo $L = 1024$ amostras é suficiente para realizar a estimativa do parâmetro supracitado.

Após a simulação com o total de 8×10^5 eventos, distribuídos igualmente para sinal e ruído, foi possível gerar as curvas ROC (do inglês *Receiver Operating Curve*) (Trees, 2001) para cada tipo de

detector. O resultado pode ser conferido na Figura 7. Nota-se que não existe uma diferença considerável de desempenho entre o detector de limiar adaptativo e limiar fixo, uma vez que, está sendo utilizado um filtro ótimo. Conforme visto na Figura 7, não se esperava incremento de desempenho, por parte do detector de decisão adaptativo, mas o que se pode perceber é uma suave perda de desempenho do mesmo visto a dificuldade em estimar $\sigma_{n'}$.

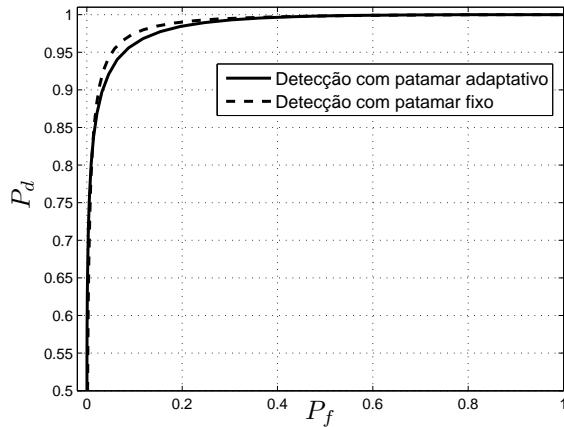


Figura 7: Curva de desempenho dos detectores com decisão adaptativa e não adaptativa, em que P_d esta representado no eixo da ordenada e P_f no eixo da abcissa.

No entanto, a motivação de se utilizar a decisão adaptativa confere ao controle da variável P_f , para visualizar tal comportamento, é necessário buscar valores de γ' e γ_{adapt} , que levem a um mesmo conjunto de P_d e P_f resultante. Devido a ligeira variação de desempenho entre os dois detectores é impossível encontrar pares idênticos de P_d e P_f resultante. No entanto, conforme pode ser observado na Tabela 1 foi encontrado pares P_d e P_f resultantes aproximadamente iguais. Neste sentido o comportamento de P_f para ambos os detectores pode ser visualizado na Figura 8.

Tabela 1: Valores adotados de P_f e seus consequentes P_d , γ' e k .

Tipo de decisão	P_f	P_d	γ'	k
Adaptativa	39,94%	99,66%	—	0,25
Fixa	39,97%	99,73%	1.36	—

6 Conclusão

Neste trabalho foi discutido uma das motivações para adotar um limiar de decisão adaptativo na classificação de eventos em sistemas de detecção, destacando o controle da probabilidade de

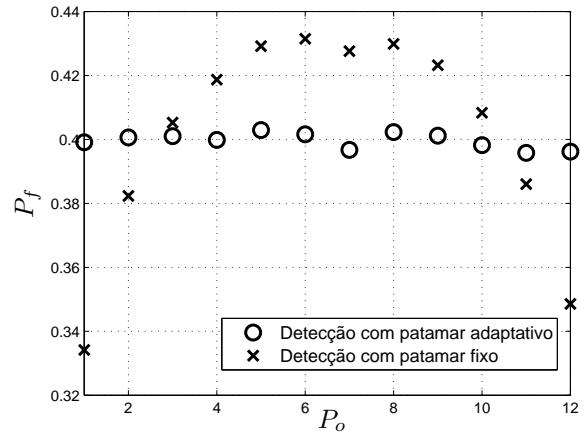


Figura 8: Comportamento da variável P_f em 12 pontos de observação P_o diferentes.

falso alarme ou nível de contaminação em função do desvio padrão do ruído resultante na saída do filtro.

Foi ilustrada a efetividade deste controle e também foi abordado a equivalência dos índices P_f e P_d resultantes para detectores com limiar fixo e adaptativo. Como resultado, a versão adaptativa oferece como motivação a não flutuação de P_f , ao contrário da versão não adaptativa, que demonstra intensa flutuação de P_f em função da mudança das características estocásticas do ruído, conforme mostrado nas Figuras 6 e 8.

Conclui-se também que a parametrização do decisor pode ser facilmente efetuada, para cenário com qualquer σ_n , uma vez que, para a configuração do detector, seja necessário apenas a escolha do P_{fd} . Uma vez escolhido o P_{fd} , resta apenas encontrar o parâmetro k correspondente de acordo com a Equação (12) ou através de consulta ao gráfico da Figura 3. A partir dos dados simulados foi possível verificar uma variação da P_f em até 6%, para o detector com limiar não adaptativo nos pontos de observação adotados. Enquanto no detector com limiar adaptativo verificou variação de P_f sempre menor que 0,5%.

Aplicações que desejam nível de contaminação constante podem fazer uso da estratégia abordada neste trabalho, e a partir da mesma, assegurar um nível constante para P_f , mesmo que o ruído presente possua desvio padrão variante no tempo.

Agradecimentos

À Coordenação de Aperfeiçoamento Pessoal de Nível Superior (CAPES), ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e à Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) pelo suporte financeiro.

Referências

Brown, S. (2000). *Fundamentals of Digital Logic with VHDL Design*, McGraw Hill.

de Andrade Filho, L. M. (2009). *Detecção e Reconstrução de Raios Cósmicos Usando Calorimetria de Altas Energias*, Tese de doutorado, Universidade Federal do Rio de Janeiro UFRJ, Rio de Janeiro, Brasil.

Papoulis, S. (1991). *Probability, Random Variable and Stochastic Processes*, McGraw Hill.

Shanmugan, K. S. (1988). *Detection, Estimation and Data Analysis*, John Wiley and Sons.

The Atlas Collaboration (2008). The atlas collaboration at cern large hadron collider, *Journal of Instrumentation*.

Trees, H. L. V. (2001). *Detection, Estimation and Modulation Theory*, Wiley-Interscience.