

Amostragem em redes de difusão: análise teórica de um algoritmo baseado na máxima correntropia

Vitor H. Duque, Daniel G. Tiglea e Magno T. M. Silva

Resumo— Neste artigo, apresenta-se uma análise teórica do algoritmo RAS-dMCC (*robust adaptive sampling diffusion maximum correntropy criterion*) para amostragem adaptativa em redes de difusão na presença de ruído impulsivo Bernoulli-Gauss. São derivadas expressões analíticas para limiares associados ao parâmetro de penalização da amostragem, permitindo orientar sua seleção. Também são obtidas aproximações para a distribuição do erro de nós não amostrados e para o número médio de nós amostrados em regime permanente. Os resultados teóricos apresentam boa concordância com simulações computacionais, validando as aproximações propostas e fornecendo informações úteis para o ajuste do algoritmo.

Palavras-Chave— processamento distribuído de sinais, redes de difusão adaptativas, amostragem, critério da máxima correntropia, ruído impulsivo.

Abstract— In this paper, we present a theoretical analysis of the RAS-dMCC (*robust adaptive sampling diffusion maximum correntropy criterion*) algorithm for adaptive sampling in diffusion networks under Bernoulli-Gaussian impulsive noise. Analytical expressions are derived for thresholds associated with the sampling penalty parameter, providing guidelines for its selection. Approximations are also obtained for the error distribution of unsampled nodes and the average number of sampled nodes in steady state. The theoretical results show good agreement with computer simulations, validating the proposed approximations and providing useful insights for tuning the algorithm.

Keywords— Distributed signal processing, adaptive diffusion networks, sampling, maximum correntropy criterion, impulsive noise.

I. INTRODUÇÃO

As redes de difusão adaptativas têm se consolidado como ferramentas relevantes para o processamento distribuído de sinais em aplicações envolvendo aprendizagem cooperativa e processamento em tempo real. Nessas redes, múltiplos agentes cooperam entre si por meio da troca de informações locais, permitindo que um objetivo global de estimação ou aprendizagem seja alcançado de forma distribuída [1]–[3].

Em muitos cenários práticos, o custo associado à aquisição, processamento e transmissão de dados pode ser elevado, especialmente em aplicações envolvendo dispositivos com restrições energéticas ou computacionais. Nesse contexto, diversas estratégias têm sido propostas para reduzir o custo de operação dessas redes, dentre as quais destacam-se os mecanismos de censura e amostragem adaptativa [4]. Em particular, algoritmos de amostragem adaptativa procuram reduzir o número de nós que realizam a etapa de adaptação ao longo das iterações, mantendo a aprendizagem ativa apenas quando necessário.

Vitor H. Duque, Daniel G. Tiglea e Magno T. M. Silva, Escola Politécnica, Universidade de São Paulo, São Paulo, SP, emails: {vitorduque1, daniel.tiglea, magno.silva}@usp.br. Este trabalho foi financiado pela CAPES (código de financiamento 001), pela FAPESP (2025/05859-7, 2026/01471-7) e pelo CNPq (301743/2026-6 e 404081/2023-1).

Isso permite reduzir significativamente o custo computacional e energético sem degradar de maneira relevante o desempenho.

Recentemente, foi proposto em [5] o algoritmo RAS-dMCC (*robust adaptive sampling diffusion maximum correntropy criterion*), que combina a amostragem adaptativa com o critério da máxima correntropia. A motivação dessa solução decorre do fato de que estratégias de amostragem baseadas no erro quadrático médio (*mean square error* – MSE) tendem a ser sensíveis à ocorrência de *outliers*, o que pode comprometer seu desempenho em ambientes com ruído impulsivo. Ao incorporar o MCC ao mecanismo de adaptação e à estratégia de amostragem, o algoritmo RAS-dMCC mostrou-se significativamente mais robusto à presença desse tipo de ruído, mantendo a convergência adequada e reduzindo o número médio de nós amostrados em regime permanente [5].

A Fig. 1 ilustra o comportamento típico dos algoritmos *adaptive sampling diffusion least-mean-squares* (AS-dLMS) [4] e RAS-dMCC [5] em ambientes com ruído impulsivo. Observa-se que o mecanismo de amostragem baseado no MSE, utilizado pelo AS-dLMS, tende a manter um número elevado de nós amostrados devido à sensibilidade do erro quadrático a *outliers*. Além disso, o algoritmo apresenta uma degradação significativa de desempenho com alto valor de desvio quadrático médio da rede (*network mean-square deviation* – NMSD). Em contrapartida, o algoritmo RAS-dMCC mantém o erro baixo em regime permanente, reduzindo substancialmente o número de nós amostrados. Essa diferença motiva a análise teórica do RAS-dMCC desenvolvida neste trabalho.

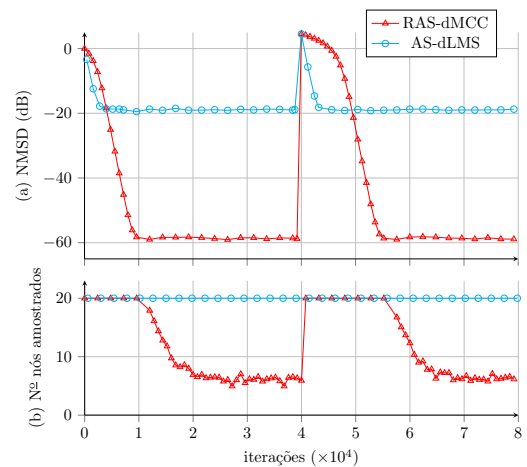


Fig. 1. Comparação entre o AS-dLMS e o RAS-dMCC na presença de ruído impulsivo: (a) NMSD e (b) número de nós amostrados ao longo das iterações. Considerou-se uma mudança abrupta no ambiente na metade de cada realização. Demais detalhes do cenário de simulação estão na Seção IV.

Apesar dos resultados promissores obtidos em [5] e re-

visitados na Fig. 1, ainda não há uma análise teórica que permita compreender o comportamento do algoritmo e orientar a seleção de seus parâmetros. Em particular, a escolha do parâmetro de penalização da amostragem influencia o número médio de nós amostrados em regime permanente, bem como o equilíbrio entre desempenho de estimação e custo computacional. Diante disso, este trabalho apresenta uma análise teórica do algoritmo RAS-dMCC considerando ambientes com ruído impulsivo modelado por um processo Bernoulli-Gauss.

O artigo está organizado da seguinte forma. Na Seção II, apresenta-se a formulação do problema e revisita-se o algoritmo RAS-dMCC. Na Seção III, são desenvolvidas as análises teóricas acerca da seleção dos parâmetros e do comportamento estatístico do mecanismo de amostragem. Resultados de simulação são apresentados na Seção IV. Finalmente, as conclusões são apresentadas na Seção V.

Notação: Letras minúsculas representam escalares e letras minúsculas em negrito representam vetores. Além disso, $(\cdot)^T$ denota transposição, $\setminus$ a diferença entre conjuntos, $P(\cdot)$ probabilidade, $\mathbb{E}\{\cdot\}$ esperança matemática, $\text{Var}(\cdot)$ a variância e $\|\cdot\|$ a norma euclidiana. Além disso, f denota função densidade de probabilidade (fdp), $\mathcal{F}$ função de distribuição acumulada (fda), $N(0, \sigma^2)$ distribuição normal com média nula e variância σ^2 e χ_k^2 distribuição qui-quadrado com k graus de liberdade.

II. FORMULAÇÃO DO PROBLEMA

Seja uma rede com V nós conectados por uma topologia pré-definida. Para cada nó k , o conjunto de nós com os quais ele pode se comunicar, incluindo ele mesmo, é denominado sua vizinhança e denotado por $\mathcal{N}_k$. A cada instante de tempo, cada nó k tem acesso a um sinal de entrada $u_k(n)$ e a um sinal desejado $d_k(n)$, modelado por $d_k(n) = \mathbf{u}_k^T(n) \mathbf{w}_o + v_k(n)$, em que $\mathbf{u}_k^T(n) = [u_k(n) \cdots u_k(n - M + 1)]^T$ é um vetor regressor formado pelas últimas M amostras de $u_k(n)$, $\mathbf{w}_o$ é um vetor coluna de dimensão M que representa um sistema desconhecido, que se supõe ser comum a todos os nós, e $v_k(n)$ é um ruído de medição independente de todos os outros sinais e independente e identicamente distribuído (iid) [1].

O objetivo comum a todos os nós da rede é estimar o vetor $\mathbf{w}_o$ de maneira distribuída, resolvendo-se $\min_{\mathbf{w}} \sum_{k=1}^V J_e(e_k(n))$, em que $J_e(e_k(n))$ é uma função custo local do erro de estimação $e_k(n) = d_k(n) - \mathbf{u}_k^T(n) \mathbf{w}$. Diversos algoritmos foram propostos para esse problema. O funcionamento dessas soluções consiste em duas etapas: a de *adaptação* e a de *combinação*. Quando realizadas nessa ordem, tem-se uma abordagem denominada *adapt-then-combine* (ATC). Empregando-se o método do gradiente estocástico para a minimização da função custo baseada na máxima correntropia $J_e(e_k(n)) = 1 - \mathbb{E}\{\kappa_\sigma(e_k(n))\}$ e o esquema ATC, foi proposto o algoritmo RAS-dMCC dado por [5]

$$\psi_k(n+1) = \mathbf{w}_k(n) + \bar{s}_k(n) \mu_k \mathbf{u}_k(n) e_k(n) \kappa_\sigma(e_k(n)), \quad (1a)$$

$$\mathbf{w}_k(n+1) = \sum_{i \in \mathcal{N}_k} c_{ik} \psi_i(n+1), \quad (1b)$$

em que ψ_k e $\mathbf{w}_k$ são respectivamente as estimativas local e combinada do vetor $\mathbf{w}_o$ no nó k , $\bar{s}_k(n) \in \{0, 1\}$ é uma variável que determina se o nó k é amostrado na iteração n ou não, μ_k é um passo de adaptação, $\kappa_\sigma(e_k(n)) = \exp(-e_k^2(n)/(2\sigma^2))$ e $e_k(n) = d_k(n) - \mathbf{u}_k^T(n) \mathbf{w}_k(n)$ é o erro de estimação. Por fim,

c_{ik} são pesos de combinação, que devem satisfazer a $c_{ik} \geq 0$, $\sum_{i \in \mathcal{N}_k} c_{ik} = 1$ e $c_{ik} = 0$ se $i \notin \mathcal{N}_k$ [1]. Há diversas estratégias para a seleção desses pesos, sendo a Metropolis uma das mais utilizadas e por isso é considerada nas simulações deste trabalho. Como métrica de desempenho, considera-se o NMSD dado por $\text{NMSD}(n) = (1/V) \sum_{k=1}^V \mathbb{E}\{\|\mathbf{w}_o - \mathbf{w}_k(n)\|^2\}$.

Para o ruído de medição, adota-se o modelo de Bernoulli-Gauss, dado por

$$v_k(n) = v_{b,k}(n) + b_k(n) v_{\text{imp},k}(n), \quad (2)$$

em que $v_{b,k} \sim N(0, \sigma_{b,k}^2)$ e $v_{\text{imp},k} \sim N(0, \sigma_{\text{imp},k}^2)$. Por fim, b_k tem uma distribuição de Bernoulli tal que $P(b_k(n) = 1) = p_k$ [6], [7]. A fdp de v_k nestas condições é dada por [8]

$$f_{v_k}(v_k) = (1 - p_k) N(0, \sigma_{b,k}^2) + p_k N(0, \sigma_{t,k}^2) \\ = \frac{1 - p_k}{\sqrt{2\pi\sigma_{b,k}^2}} \exp\left(\frac{-v_k^2}{2\sigma_{b,k}^2}\right) + \frac{p_k}{\sqrt{2\pi\sigma_{t,k}^2}} \exp\left(\frac{-v_k^2}{2\sigma_{t,k}^2}\right), \quad (3)$$

em que $\sigma_{t,k}^2 \triangleq \sigma_{b,k}^2 + \sigma_{\text{imp},k}^2$.

A ideia do mecanismo de amostragem do RAS-dMCC é que, se $\bar{s}_k(n) = 0$, nenhum dos termos na segunda parcela do lado direito de (1a) precisa ser calculado. Caso contrário, o termo de correção é calculado normalmente. Busca-se então controlar $\bar{s}_k(n)$ de modo a reduzir o custo computacional sem impactar negativamente o desempenho em termos de NMSD. Para isso, em vez de se adaptar $\bar{s}_k(n)$ diretamente, introduz-se uma variável auxiliar $\alpha_k(n)$ tal que

$$\bar{s}_k(n) = \begin{cases} 1, & \text{se } \alpha_k(n) \geq 0 \\ 0, & \text{caso contrário} \end{cases}. \quad (4)$$

Então, adapta-se $\alpha_k(n)$ ao longo das iterações. Para isso, define-se uma função custo, dada por [4], [5]

$$J_{\alpha_k}(n) = \phi_k(n) \beta \bar{s}_k(n) + [1 - \phi_k(n)] \sum_{i \in \mathcal{N}_k} c_{ik} \hat{J}_e(e_i(n)), \quad (5)$$

em que β é um parâmetro que o projetista do filtro deve escolher e $\hat{J}_e(e_k(n)) = 1 - \kappa_\sigma(e_k(n))$ é uma aproximação instantânea de $J_e(e_k(n))$. Por fim, $\phi_k(n)$ é dada por

$$\phi_k(n) = \phi[\alpha_k(n)] \triangleq \frac{\text{sgm}[\alpha_k(n)] - \text{sgm}[-\alpha^+]}{\text{sgm}[\alpha^+] - \text{sgm}[-\alpha^+]}, \quad (6)$$

em que $\text{sgm}[x] = [1 + \exp(-x)]^{-1}$ é uma função sigmoide e α^+ é o maior valor que $\alpha_k(n)$ pode assumir. Tipicamente, adota-se $\alpha^+ = 4$ [4], [5].

Adotando-se uma abordagem do tipo gradiente descendente para minimizar $J_{\alpha_k}(n)$, obtém-se após algumas simplificações uma equação de atualização para $\alpha_k(n)$, dada por [4], [5]

$$\alpha_k(n+1) = \alpha_k(n) + \mu_s \phi'_k(n) \left[\sum_{i \in \mathcal{N}_k} c_{ik} \hat{J}_e(\varepsilon_i(n)) - \beta \bar{s}_k(n) \right], \quad (7)$$

em que μ_s é um passo de adaptação e $\varepsilon_i(n) = \bar{s}_i(n) e_i(n) + [1 - \bar{s}_i(n)] \varepsilon_i(n-1)$ é o último erro de estimação disponível no nó i , já que seu valor não é calculado quando o nó não é amostrado. Por fim, $\phi'_k(n)$ é dada por [4]

$$\phi'_k(n) = \frac{d\phi[\alpha_k(n)]}{d\alpha_k(n)} = \frac{\text{sgm}[\alpha_k(n)] \{1 - \text{sgm}[\alpha_k(n)]\}}{\text{sgm}[\alpha^+] - \text{sgm}[-\alpha^+]}. \quad (8)$$

O bom comportamento do algoritmo depende de uma adequada seleção dos seus parâmetros, o que é abordado a seguir.

III. ANÁLISE TEÓRICA

A análise teórica é dividida em três partes. Na primeira, busca-se a seleção adequada do parâmetro β , que influencia o número de nós amostrados por iteração em regime permanente, pois controla a penalização da amostragem dos nós. Na segunda, busca-se um modelo para a distribuição do erro que leva o nó ao estado não amostrado. Na terceira, analisa-se o número esperado de nós amostrados em regime estacionário.

A. Seleção do parâmetro β

O nó k deixa de ser amostrado quando $\alpha_k(n) < 0$, o que, na média, requer

$$\mathbb{E}\{\alpha_k(n+1)\} < \mathbb{E}\{\alpha_k(n)\}. \quad (9)$$

Como $\mu_s \phi'_k(n) > 0$, conclui-se de (7) que (9) equivale a

$$\mathbb{E}\left\{\sum_{i \in N_k} c_{ik} \hat{J}_e(\varepsilon_i(n)) - \beta \bar{s}_k(n)\right\} < 0. \quad (10)$$

Se o nó k é amostrado na iteração n , pode-se reescrever (10) como

$$\sum_{i \in N_k} c_{ik} \mathbb{E}\{\hat{J}_e(e_i(n))\} - \beta < 0. \quad (11)$$

Logo, para cada nó k existe um $\hat{\beta}_k$ dado por

$$\hat{\beta}_k \triangleq \sum_{i \in N_k} c_{ik} \mathbb{E}\{\hat{J}_e(e_i(n))\}, \quad (12)$$

tal que a adoção de $\beta > \hat{\beta}_k$ faz com que (9) seja satisfeita em regime permanente.

Seguindo uma abordagem comum na literatura, pode-se escrever $e_i(n)$ como $e_i(n) = e_{a_i}(n) + v_i(n)$, em que $e_{a_i}(n) \triangleq \mathbf{u}_i^T(n)[\mathbf{w}_i(n) - \mathbf{w}_o]$ é o erro *a priori* [9], [10]. Assim, pode-se tomar a expansão de $\hat{J}_e(e_i(n)) = 1 - \kappa_\sigma(e_i(n))$ em Série de Taylor no entorno de $v_i(n)$, semelhante ao que foi feito em [12]. Como se está interessado no número de nós amostrados em regime permanente, $n \rightarrow \infty$, o índice n será omitido daqui em diante. Assim, tem-se

$$\hat{J}_e(e_i) = \hat{J}_e(v_i) + \hat{J}'_e(v_i)e_{a_i} + \frac{1}{2}\hat{J}''_e(v_i)e_{a_i}^2 + o(e_{a_i}^2), \quad (13)$$

em que $o(e_{a_i}^2)$ denota termos de ordem elevada. Assumindo-se que e_{a_i} e v_i são independentes entre si e que $\mathbb{E}\{e_{a_i}\} = 0$ [9], [10], tomando-se a esperança de ambos os lados de (13) e desprezando-se $o(e_{a_i}^2)$, obtém-se

$$\mathbb{E}\{\hat{J}_e(e_i)\} \approx 1 - \left[\mathbb{E}\{\kappa_\sigma(v_i)\} + \frac{1}{2}\mathbb{E}\{\kappa''_\sigma(v_i)\}\zeta_i\right], \quad (14)$$

em que ζ_i denota o erro quadrático médio em excesso (EMSE — *excess MSE*) em regime permanente do nó i , dado por $\zeta_i \triangleq \lim_{n \rightarrow \infty} \mathbb{E}\{e_{a_i}^2(n)\}$ [9], [10]. Substituindo-se (14) em (12), conclui-se que, na média, o nó k deixa de ser amostrado em regime permanente se

$$\beta > \hat{\beta}_k = \sum_{i \in N_k} c_{ik} \left[1 - \left(\mathbb{E}\{\kappa_\sigma(v_i)\} + \frac{1}{2}\mathbb{E}\{\kappa''_\sigma(v_i)\}\zeta_i\right)\right]. \quad (15)$$

Usando (3) para calcular a esperança de $\kappa_\sigma(v_i)$, chega-se a

$$\begin{aligned} \mathbb{E}\{\kappa_\sigma(v_i)\} &= \int_{-\infty}^{\infty} \kappa_\sigma(v_i) f_{v_i}(v_i) dv_i \\ &= \int_{-\infty}^{\infty} \exp\left(\frac{-v_i^2}{2\sigma^2}\right) \frac{1-p_i}{\sqrt{2\pi\sigma_{b,i}^2}} \exp\left(\frac{-v_i^2}{2\sigma_{b,i}^2}\right) dv_i \\ &\quad + \int_{-\infty}^{\infty} \exp\left(\frac{-v_i^2}{2\sigma^2}\right) \frac{p_i}{\sqrt{2\pi\sigma_{t,i}^2}} \exp\left(\frac{-v_i^2}{2\sigma_{t,i}^2}\right) dv_i. \end{aligned} \quad (16)$$

Calculando-se ambas as integrais, obtém-se

$$\mathbb{E}\{\kappa_\sigma(v_i)\} = \sigma \left[\frac{1-p_i}{\sqrt{\sigma^2 + \sigma_{b,i}^2}} + \frac{p_i}{\sqrt{\sigma^2 + \sigma_{t,i}^2}} \right]. \quad (17)$$

Por sua vez, $\kappa''_\sigma(x)$ é dado por

$$\kappa''_\sigma(x) = \kappa_\sigma(x) \left[\frac{x^2}{\sigma^4} - \frac{1}{\sigma^2} \right]. \quad (18)$$

Substituindo $\kappa_\sigma(v_i)$ e $\kappa''_\sigma(v_i)$ em (16) e usando (18), $\mathbb{E}\{\kappa''_\sigma(v_i)\}$ pode ser calculada de maneira análoga, i.e.,

$$\mathbb{E}\{\kappa''_\sigma(v_i)\} = -\sigma \left[\frac{1-p_i}{(\sigma^2 + \sigma_{b,i}^2)^{3/2}} + \frac{p_i}{(\sigma^2 + \sigma_{t,i}^2)^{3/2}} \right]. \quad (19)$$

Finalmente, usando o resultado de [12], é possível mostrar que o EMSE alcançado por um filtro adaptativo independente executando o algoritmo MCC no nó i é dado por

$$\zeta_i = \frac{\mu_i \text{Tr}(\mathbf{R}_i)}{2} \frac{\frac{(1-p_i)\sigma_{b,i}^2}{(\sigma^2 + 2\sigma_{b,i}^2)^{3/2}} + \frac{p_i\sigma_{t,i}^2}{(\sigma^2 + 2\sigma_{t,i}^2)^{3/2}}}{\frac{1-p_i}{(\sigma^2 + \sigma_{b,i}^2)^{3/2}} + \frac{p_i}{(\sigma^2 + \sigma_{t,i}^2)^{3/2}}}, \quad (20)$$

em que $\text{Tr}(\mathbf{R}_i)$ denota o traço da matriz de autocorrelação do sinal de entrada, i.e., $\mathbf{R}_i = \mathbb{E}\{\mathbf{u}_i(n)\mathbf{u}_i^T(n)\}$.

Substituindo (17), (19) e (20) em (15), obtém-se finalmente a estimativa de $\hat{\beta}_k$ e a condição para para que o nó k deixe de ser amostrado na média, em regime permanente. Devido à cooperação entre os nós, (20) pode superestimar ligeiramente o EMSE real do nó i . Ainda assim, simulações sugerem que (20) representa uma boa aproximação para a seleção de β .

A escolha de μ_s também requer cuidado, embora simulações indiquem que esse parâmetro seja menos crítico para o desempenho do RAS-dMCC do que β . Uma análise detalhada de μ_s é deixada para trabalhos futuros.

B. Distribuição do erro de nós não amostrados

Para cada nó k , considera-se que em uma dada iteração n , tem-se $\alpha_k(n) > 0$, mas com $\alpha_k(n) \approx 0$. Por simplicidade, essa situação é denotada por $\alpha_k(n) = 0^+$. Nessas condições, $\alpha_k(n+1)$ é negativo se

$$g_k(\mathcal{N}_k) \triangleq \sum_{i \in \mathcal{N}_k} c_{ik} \hat{J}_e(\varepsilon_i(n)) - \beta < 0. \quad (21)$$

Sejam $\tilde{\varepsilon}_k$ o erro na iteração anterior ao nó k deixar de ser amostrado e o evento $A_k = \{g_k(\mathcal{N}_k) < 0\} \cap \{\alpha_k = 0^+\}$. Pelo Teorema de Bayes, a fdp de $\tilde{\varepsilon}_k$ é

$$f_{\tilde{\varepsilon}_k}(\tilde{\varepsilon}_k) = f_{\varepsilon_k|A_k}(\varepsilon_k|A_k) = \frac{f_{\varepsilon_k}(\varepsilon_k)P(A_k|\varepsilon_k)}{P(A_k)}. \quad (22)$$

Sejam os eventos $B_k = \{\sum_{i \in \mathcal{N}_k} b_i = 0\}$ e $\bar{B}_k$ o seu complementar. Assim, o numerador de (22) é dado por

$$\begin{aligned} f_{\varepsilon_k}(\varepsilon_k)P(A_k|\varepsilon_k) &= f_{\varepsilon_k|B_k}(\varepsilon_k|B_k)P(A_k|\varepsilon_k, B_k) \\ &\quad + f_{\varepsilon_k|\bar{B}_k}(\varepsilon_k|\bar{B}_k)P(A_k|\varepsilon_k, \bar{B}_k). \end{aligned} \quad (23)$$

Na ocorrência de um ruído impulsivo para algum nó $i \in \mathcal{N}_k$, tipicamente se obtém $g_k(\mathcal{N}_k) > 0$, dado que $\hat{J}_e(\varepsilon_i(n))$ tende a valores elevados. Logo, assume-se que $P(A_k|\varepsilon_k, \bar{B}_k) = 0$. Um raciocínio análogo pode ser aplicado à probabilidade $P(A_k)$ do denominador, o que permite reescrever (22) como

$$f_{\tilde{\varepsilon}_k}(\tilde{\varepsilon}_k) = \frac{f_{\varepsilon_k|B_k}(\varepsilon_k|B_k)P(A_k|\varepsilon_k, B_k)}{P(A_k|B_k)}. \quad (24)$$

Assumindo-se que $\varepsilon_i \approx v_{b,i}$, tem-se $f_{\varepsilon_i|B_k}(\varepsilon_i|B_k) \approx N(0; \sigma_{b,i}^2)$, $\forall i \in \mathcal{N}_k$. Falta então calcular as probabilidades em (24). Sob a hipótese de ocorrência de B_k , pode-se aproximar o *kernel* por sua série de Taylor de primeira ordem em torno de $\varepsilon \approx 0$, i.e., $\kappa_\sigma(\varepsilon) = \exp\left(\frac{-\varepsilon^2}{2\sigma^2}\right) \approx 1 - \frac{\varepsilon^2}{2\sigma^2}$. Nessas condições, pode-se reescrever (21) como

$$g_k(\mathcal{N}_k) \approx \sum_{i \in \mathcal{N}_k} c_{ik} \frac{\varepsilon_i^2}{2\sigma^2} - \beta < 0. \quad (25)$$

Define-se $T \triangleq 2\sigma^2\beta$ e $Q'_k \triangleq \sum_{i \in \mathcal{N}'_k} c_{ik} \varepsilon_i^2$ para facilitar a notação, em que $\mathcal{N}'_k \triangleq \mathcal{N}_k \setminus \{k\}$. Como a fdp condicional $f_{\varepsilon_i|B_k}(\varepsilon_i|B_k)$ possui distribuição gaussiana, Q'_k tem distribuição qui-quadrado sob a hipótese de ocorrência de B_k , ou seja, $Q'_k \sim \sum_{i \in \mathcal{N}'_k} c_{ik} \sigma_{b,i}^2 \chi_1^2$ com média e variância dadas por

$$\mathbb{E}\{Q'_k\} = \sum_{i \in \mathcal{N}'_k} c_{ik} \sigma_{b,i}^2 \approx \lambda'_k \nu'_k \quad (26a)$$

$$\text{Var}(Q'_k) = \sum_{i \in \mathcal{N}'_k} c_{ik} 2\sigma_{b,i}^4 \approx 2(\lambda'_k)^2 \nu'_k, \quad (26b)$$

sendo que $\lambda'_k = \frac{\sum_{i \in \mathcal{N}'_k} c_{ik}^2 \sigma_{b,i}^4}{\sum_{i \in \mathcal{N}'_k} c_{ik} \sigma_{b,i}^2}$ e $\nu'_k = \frac{\left(\sum_{i \in \mathcal{N}'_k} c_{ik} \sigma_{b,i}^2\right)^2}{\sum_{i \in \mathcal{N}'_k} c_{ik}^2 \sigma_{b,i}^4}$ representam a escala e número de graus de liberdade efetivo de Q'_k , respectivamente [11].

Dado ε_k , conclui-se a partir de (25) que, se B_k ocorre, $g_k(\mathcal{N}_k) < 0$ implica $Q'_k < (T - c_{kk}\varepsilon_k^2)/\lambda'_k$. Logo, a probabilidade condicional $P(A_k|\varepsilon_k, B_k)$ no numerador de (24) pode ser aproximada usando a fda, i.e.,

$$P\left(Q'_k < \frac{T - c_{kk}\varepsilon_k^2}{\lambda'_k}\right) \approx \mathcal{F}\chi_{\nu'_k}^2\left(\frac{T - c_{kk}\varepsilon_k^2}{\lambda'_k}\right). \quad (27)$$

Analogamente, definindo-se $Q_k \triangleq \sum_{i \in \mathcal{N}_k} c_{ik} \varepsilon_i^2 = Q'_k + c_{kk}\varepsilon_k^2$, nota-se que $\mathbb{E}\{Q_k\} = \lambda_k \nu_k$ e $\text{Var}(Q_k) = 2\lambda_k^2 \nu_k$, com

$$\lambda_k = \frac{\sum_{i \in \mathcal{N}_k} c_{ik}^2 \sigma_{b,i}^4}{\sum_{i \in \mathcal{N}_k} c_{ik} \sigma_{b,i}^2} \quad \text{e} \quad \nu_k = \frac{\left(\sum_{i \in \mathcal{N}_k} c_{ik} \sigma_{b,i}^2\right)^2}{\sum_{i \in \mathcal{N}_k} c_{ik}^2 \sigma_{b,i}^4}.$$

É possível constatar que a probabilidade no denominador de (24) é dada por $P(A_k|B_k) = P(Q_k < T/\lambda_k) \approx \mathcal{F}\chi_{\nu_k}^2\left(\frac{T}{\lambda_k}\right)$. Substituindo-se este resultado e (27) em (22), conclui-se que

$$f_{\tilde{\varepsilon}_k}(\tilde{\varepsilon}_k) = \frac{f_{\tilde{\varepsilon}_k}(\varepsilon_k) \mathcal{F}\chi_{\nu'_k}^2\left(\frac{T - c_{kk}\varepsilon_k^2}{\lambda'_k}\right)}{\mathcal{F}\chi_{\nu_k}^2\left(\frac{T}{\lambda_k}\right)}. \quad (28)$$

Esse resultado é usado em seguida para a estimação do número de nós amostrados em regime estacionário.

C. Número de nós amostrados em regime estacionário

Em regime permanente, caso (15) seja satisfeita, a amostragem dos nós é interrompida e retomada de maneira cíclica. O nó k deve permanecer amostrado por um certo número ξ_k de iterações até que sua amostragem cesse, durante as quais tem-se $\alpha_k(n) > 0$ e $\Delta\alpha_k(n) \triangleq \alpha_k(n+1) - \alpha_k(n) < 0$. Quando α_k se torna negativo, o nó k deixa de ser amostrado e seguem-se $\bar{\xi}_k$ iterações em que se tem $\alpha_k(n) < 0$ e $\Delta\alpha_k(n) > 0$. Eventualmente, $\alpha_k(n)$ e $\Delta\alpha_k(n)$ trocam de sinal e o ciclo se repete. Por exemplo, consideremos que $\alpha_k(n) = 0^-$, ou seja, $\alpha_k(n) < 0$ e $\alpha_k(n) \approx 0$. Neste cenário, $\mathbb{E}\{\alpha_k(n+1)\} = \mathbb{E}\{\Delta\alpha_k(n)\} > 0$ e a amostragem do nó k será retomada no instante $n+1$. A partir daí, na média, α_k decresce

a cada iteração de um valor de $\mathbb{E}\{\Delta\alpha_k(n)|\alpha_k(n) > 0\} < 0$. Calculando-se a razão entre essas grandezas, pode-se estimar $\theta_k \triangleq \mathbb{E}\{\xi_k\}$ como

$$\theta_k = \frac{\mathbb{E}\{\alpha_k(n+1)|\alpha_k(n) = 0^-\}}{\mathbb{E}\{\Delta\alpha_k(n)|\alpha_k(n) > 0\}}. \quad (29)$$

Analogamente, $\bar{\theta}_k \triangleq \mathbb{E}\{\bar{\xi}_k\}$ pode ser estimado por

$$\bar{\theta}_k = \frac{\mathbb{E}\{\alpha_k(n+1)|\alpha_k(n) = 0^+\}}{\mathbb{E}\{\Delta\alpha_k(n)|\alpha_k(n) < 0\}}. \quad (30)$$

Para prosseguir com as estimativas, é necessário tomar a esperança de ambos os lados de (7) e fazer algumas manipulações. Nota-se que essas esperanças dependem de $\hat{J}_e(\varepsilon_i)$ para $i \in \mathcal{N}_k$. A seguir, considera-se que $\mu_s \phi'_k(n)$ se mantém aproximadamente constante ao longo dessas iterações, de forma que o efeito desse fator pode ser desprezado, pois ele aparece no numerador e no denominador em (29) e (30).

Define-se $\bar{J}_{f,i} \triangleq \mathbb{E}\{\hat{J}_e(\varepsilon_i)\}$, i.e., o valor esperado de $\hat{J}_e(\varepsilon_i)$ considerando a distribuição f . As esperanças no numerador de (29) e no denominador de (30) devem ser calculadas usando (28) devido às suas condicionais. Por sua vez, a esperança no denominador de (29) deve considerar uma distribuição impulsiva para ε_k conforme (3), denotada por BG . Estima-se um valor máximo para θ_k considerando o caso em que todos os nós $i \in \mathcal{N}_k$ estão sendo amostrados, pois neste cenário $J_e(\varepsilon_i(n))$ tende a ser elevado devido a ε_i seguir a distribuição impulsiva. Isso faz com que o numerador em (29) assuma valores altos e o denominador, baixos. Assim, obtém-se

$$\theta_{k,max} = \frac{c_{kk} \bar{J}_{f_{\tilde{\varepsilon}_k}}(\tilde{\varepsilon}_k),k + \sum_{i \in \mathcal{N}'_k} c_{ik} \bar{J}_{BG,i}}{\beta - \sum_{i \in \mathcal{N}_k} c_{ik} \bar{J}_{BG,i}}. \quad (31)$$

Em contrapartida, um valor mínimo para θ_k pode ser estimado fazendo-se a hipótese contrária, ou seja, que nenhum nó $i \in \mathcal{N}'_k$ está sendo amostrado. Com isso, tem-se

$$\theta_{k,min} = \frac{\sum_{i \in \mathcal{N}_k} c_{ik} \bar{J}_{f_{\tilde{\varepsilon}_i}}(\tilde{\varepsilon}_i),i}{\beta - c_{kk} \bar{J}_{BG,k} - \sum_{i \in \mathcal{N}'_k} c_{ik} \bar{J}_{f_{\tilde{\varepsilon}_i}}(\tilde{\varepsilon}_i),i}. \quad (32)$$

Analogamente, obtém-se estimativas para os valores máximo e mínimo de $\bar{\theta}_k$. Pela hipótese da Seção III-B, a contribuição dos nós amostrados no numerador de (30) segue uma fdp gaussiana denotada por N . Isso permite obter

$$\bar{\theta}_{k,max} = \frac{\beta - \sum_{i \in \mathcal{N}'_k} c_{ik} \bar{J}_{f_{\tilde{\varepsilon}_i}}(\tilde{\varepsilon}_i),i - c_{kk} \bar{J}_{N,k}}{\sum_{i \in \mathcal{N}_k} c_{ik} \bar{J}_{f_{\tilde{\varepsilon}_i}}(\tilde{\varepsilon}_i),i} \quad (33)$$

e

$$\bar{\theta}_{k,min} = \frac{\beta - \sum_{i \in \mathcal{N}_k} c_{ik} \bar{J}_{N,i}}{c_{kk} \bar{J}_{f_{\tilde{\varepsilon}_k}}(\tilde{\varepsilon}_k),k + \sum_{i \in \mathcal{N}'_k} c_{ik} \bar{J}_{BG,i}}, \quad (34)$$

respectivamente. Pode-se estimar a probabilidade p_{s_k} do nó k ser amostrado em regime permanente como $p_{s_k} \approx \theta_k/(\theta_k + \bar{\theta}_k)$. Com base em (31)–(34), é possível estimar valores máximos e mínimos para p_{s_k} por meio de

$$p_{s_k,max} = \frac{\max\{1, \theta_{k,max}\}}{\max\{1, \theta_{k,max}\} + \max\{0, \bar{\theta}_{k,min}\}} \quad (35)$$

e

$$p_{s_k,min} = \frac{\max\{1, \theta_{k,min}\}}{\max\{1, \theta_{k,min}\} + \max\{0, \bar{\theta}_{k,max}\}}, \quad (36)$$

respectivamente. O uso do operador $\max$ em (35) e (36) se deve ao fato de que todo nó permanece amostrado por pelo menos uma iteração e não necessariamente deixa de ser amostrado. Caso (15) não seja satisfeita, as estimativas dadas

por (31)–(34) podem ser negativas e perdem a validade. Neste caso, pode-se fixar $p_{s_k, \max} = p_{s_k, \min} = 1$.

Por fim, limites inferiores e superiores para o número $\bar{V}_s$ esperado de nós amostrados por iteração em regime permanente podem ser obtidos por meio de

$$\sum_{k=1}^V p_{s_k, \min} \leq \bar{V}_s \leq \sum_{k=1}^V p_{s_k, \max}. \quad (37)$$

IV. RESULTADOS DE SIMULAÇÕES

Os resultados a seguir foram obtidos com uma média de 120 realizações considerando a rede com $V=20$ nós da Fig. 2(a). Na Fig. 2(b) é mostrado o valor de *signal-to-background-noise ratio* (SBNR) para cada nó. Considera-se $p_k = 0,01$ e $\sigma_{\text{imp},k}^2 = 10^6 \sigma_{b,k}^2$ para $k=1, \dots, V$. Os $M=10$ coeficientes do vetor $\mathbf{w}_o$ foram gerados aleatoriamente com distribuição normal com média nula e variância unitária e posteriormente normalizados de modo que $\|\mathbf{w}_o\| = 1$. O sinal $u_k(n)$ é gaussiano branco com média nula e variância $\sigma_{u_k}^2 = 1$. Por fim, adotou-se $\mu_k = 10^{-3}$ para todos os nós e $\sigma^2 = 0,5$ para o kernel gaussiano.

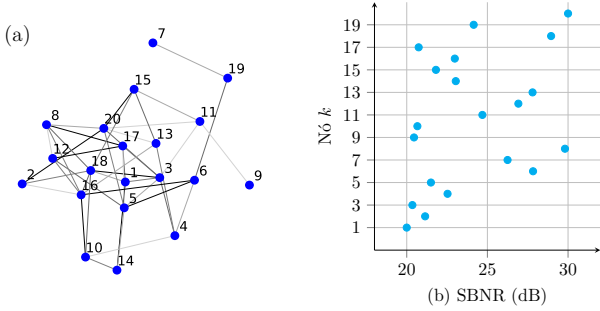


Fig. 2. (a) Rede adaptativa. (b) Distribuição dos SBNR entre os nós da rede.

Os resultados da Fig. 1 foram obtidos considerando-se $\beta = 0,02$ e $\mu_s = 1,25$. A fim de validar a análise da Seção III-A, em um cenário estacionário, variou-se o valor de β , mantendo-se sempre $\mu_s = 1/(40\beta)$. Para fins de ilustração, consideremos os nós 6, 8 e 10. Utilizando (15), obtiveram-se $\hat{\beta}_6 \approx 1,270 \cdot 10^{-2}$, $\hat{\beta}_8 \approx 1,767 \cdot 10^{-2}$ e $\hat{\beta}_{10} \approx 1,539 \cdot 10^{-2}$, respectivamente. A Fig. 3 mostra a probabilidade de amostragem p_{s_k} desses nós em função de β . Os valores teóricos são indicados pelas linhas verticais. Nota-se que, para cada um desses nós, a probabilidade de amostragem diminui quando β supera o correspondente $\hat{\beta}_k$, como previsto na análise teórica.

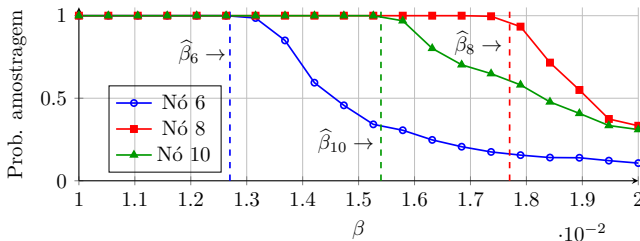


Fig. 3. Valores de p_{s_k} para $k \in \{6, 8, 10\}$ e diferentes valores de β .

Para validar a análise da Seção III-B, na Fig. 4 mostram-se o histograma de $\tilde{\varepsilon}_6$ e a estimativa de sua fdp dada por (28). Nota-se que a previsão teórica modela bem a distribuição dos dados. Por fim, na Fig. 5, apresenta-se o número médio de nós amostrados por iteração em regime permanente para diferentes valores de β . Também são apresentados os limites inferior e

superior para $\bar{V}_s$ dados por (37). Observa-se que os resultados de simulação permanecem entre ambos os limiares para todos os valores de β , o que corrobora a análise da Seção III-C.

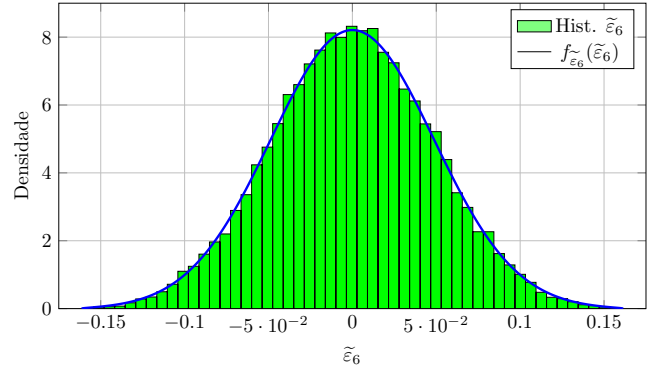


Fig. 4. Histograma dos $\tilde{\varepsilon}_6$ e a fdp estimada por (28).

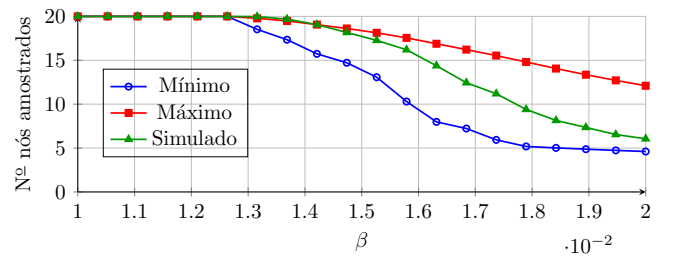


Fig. 5. Curva e estimativas do número esperado de nós amostrados em regime permanente para diferentes valores de β .

V. CONCLUSÕES

Neste trabalho, foi apresentada uma análise do algoritmo RAS-dMCC, voltado para estimação distribuída em ambientes com ruído impulsivo. Foram calculados limiares que auxiliam na escolha do seu parâmetro β . A fdp dos erros que levam os nós ao estado não amostrado foi modelada e esse resultado foi usado para a estimação de um intervalo para o número esperado de nós amostrados em regime estacionário. Os resultados de simulação corroboram a análise teórica apresentada. Em trabalhos futuros, pretende-se realizar uma análise acerca do parâmetro μ_s e testar o algoritmo em mais cenários.

REFERÊNCIAS

- [1] A. H. Sayed, *Adaptation, Learning, and Optimization over Networks*, vol. 7, Found. Trends Mach. Learn., now Publishers Inc., 2014.
- [2] S. Vlaski *et al.*, “Networked signal and inform. process.: Learn. multi-agent systems,” *IEEE Signal Process. Mag.*, vol. 40, pp. 92–105, 2023.
- [3] D. G. Tiglea, R. Candido, and M. T. M. Silva, “Adaptive diffusion networks: An overview,” *Signal Process.*, vol. 223, 2024.
- [4] D. G. Tiglea, R. Candido, and M. T. M. Silva, “A low-cost algorithm for adaptive sampling and censoring in diffusion networks,” *IEEE Trans. Signal Process.*, vol. 69, pp. 58–72, 2021.
- [5] V. H. Duque, D. G. Tiglea, and M. T. M. Silva, “Amostragem em redes de difusão adaptativas na presença de ruído impulsivo,” in *Anais do SBrt*, Natal, 2025.
- [6] J. Jeong and S. Kim, “Robust adaptive filter algorithms against impulsive noise,” *Circuits Syst. Signal Process.*, vol. 38, 12 2019.
- [7] A. Wilson, T. Panigrahi, and A. Dubey, “Robust distributed lorentzian adaptive filter with diffusion strategy in impulsive noise environment,” *Digit. Signal Process.*, vol. 96, 2019.
- [8] T. Shongwey, A.H.Vinck, H.C.Ferreira, “On impulse noise and its models,” in *IEEE Int. Symp. Power Line Commun. Appl.*, 2014, pp.12-17.
- [9] A. H. Sayed, *Adaptive Filters*, John Wiley & Sons, 2008.
- [10] S. Haykin, *Adaptive Filter Theory*, Pearson, 5th ed., 2014.
- [11] N. L. Johnson, S. Kotz, and N. Balakrishnan, *Continuous Univariate Distributions*, Wiley, 2nd ed., 1994.
- [12] B. Chen *et al.*, “Steady-state mean-square error analysis for adaptive filtering under the maximum correntropy criterion,” *IEEE Signal Process. Lett.*, vol. 21, pp. 880–884, 2014.