REMUNOM
ISSN 2178-6925

Received: 13/01/2026 - **Accepted:** 02/03/2026
Vol: 13.02
DOI: 10.66104/hegxr115
Pages: 1-28

MODELAGEM PARAMÉTRICA LPC E PRONY INTEGRADA AO FILTRO DE KALMAN PARA REDUÇÃO DE RUÍDO EM SINAIS DE FALA: ANÁLISE COMPARATIVA E AVALIAÇÃO ESTATÍSTICA

LPC AND PRONY PARAMETRIC MODELING INTEGRATED INTO THE KALMAN FILTER FOR SPEECH NOISE REDUCTION: COMPARATIVE ANALYSIS AND STATISTICAL EVALUATION

MODELADO PARAMÉTRICO LPC Y PRONY INTEGRADO AL FILTRO DE KALMAN PARA LA REDUCCIÓN DE RUIDO EN SEÑALES DE VOZ: ANÁLISIS COMPARATIVO Y EVALUACIÓN ESTADÍSTICA

Leandro Aureliano da Silva
Doutor em Engenharia Elétrica, Universidade de Uberaba - Uniube, Brasil
E-mail: leandro.silva@uniube.br

Eduardo Silva Vasconcelos
Doutor em Engenharia Elétrica, Instituto Federal Goiano - IF-GOIANO, Brasil
E-mail: educelos1@gmail.com

Luiz Fernando Ribeiro de Paiva
Doutor em Educação, Universidade de Uberaba - Uniube, Brasil
E-mail: luiz.paiva@uniube.br

Adriano Dawison de Lima
Doutor em Agronomia, Universidade de Uberaba - Uniube, Brasil
E-mail: adriano.lima @uniube.br

Welington Mrad Joaquim
Mestrado em Ciências e Matemática, Universidade de Uberaba - Uniube, Brasil
E-mail: welington.joaquim@uniube.br

Edilberto Pereira Teixeira
Doutor em Engenharia Elétrica, Universidade de Uberaba - Uniube, Brasil
E-mail: edilberto.teixeira@uniube.br

Resumo

Este trabalho apresenta um estudo experimental comparativo entre duas estratégias de modelagem paramétrica incorporadas ao filtro de Kalman para redução de ruído em sinais de voz: modelos autorregressivos (LPC-AR) e modelos polo-zero estimados pelo método de Prony (ARMA). O protocolo experimental considerou controle da relação sinal-ruído segmentada de entrada (≈3 dB), múltiplas realizações estocásticas e dois cenários de ruído (branco e colorido AR(1)). O desempenho foi avaliado por meio de métricas complementares: relação sinal-ruído segmentada de saída

REMUNOM
ISSN 2178-6925

Received: 13/01/2026 - **Accepted:** 02/03/2026
Vol: 13.02
DOI: 10.66104/hegxr115
Pages: 1-28

(segSNR_out), divergência de Itakura–Saito (IS), intelligibilidade objetiva de curto prazo (STOI) e tempo de processamento. Os resultados, reportados como média ± desvio padrão e analisados por teste de Friedman com pós-hoc, indicam que a abordagem Kalman+LPC apresentou desempenho global superior e maior consistência estatística, especialmente sob ruído branco, com menor custo computacional. A formulação Kalman+Prony mostrou estabilidade numérica adequada, com polos majoritariamente dentro do círculo unitário, mas não apresentou ganho sistemático em relação ao modelo all-pole. A análise de Pareto evidenciou um compromisso entre distorção espectral e energia residual, sem vantagem clara do modelo ARMA nas condições avaliadas. Conclui-se que, para os cenários testados, o modelo LPC integrado ao filtro de Kalman constitui a alternativa mais robusta em termos de desempenho médio, estabilidade e eficiência computacional.
Palavras-chave: Filtro de Kalman; Supressão de ruído; Enriquecimento de voz; Prony; LPC.

Abstract

This work presents a comparative experimental study between two parametric modeling strategies integrated into the Kalman filter for speech denoising: autoregressive models (LPC-AR) and pole–zero models estimated via the Prony method (ARMA). The experimental protocol included controlled input segmental signal-to-noise ratio (segSNR_in ≈ 3 dB), multiple stochastic noise realizations, and two noise scenarios (white and colored AR(1)). Performance was evaluated using complementary metrics: output segmental signal-to-noise ratio (segSNR_out), Itakura–Saito divergence (IS), short-time objective intelligibility (STOI), and computational time. Results were reported as mean ± standard deviation and analyzed using the Friedman test with post-hoc comparisons. The findings indicate that the Kalman+LPC approach achieved superior overall performance and greater statistical consistency, particularly under white noise conditions, while also exhibiting lower computational cost. The Kalman+Prony formulation demonstrated adequate numerical stability, with poles predominantly located inside the unit circle; however, it did not provide systematic performance gains over the all-pole model. Pareto analysis revealed a trade-off between spectral distortion and residual energy, with no clear advantage of the ARMA model under the evaluated conditions. It is concluded that, for the tested scenarios, the LPC model integrated into the Kalman filter constitutes the most robust alternative in terms of average performance, numerical stability, and computational efficiency.

Keywords: Kalman filter; Noise suppression; Speech enhancement; Prony; LPC.

Resumen

Este trabajo presenta un estudio experimental comparativo entre dos estrategias de modelado paramétrico integradas en el filtro de Kalman para la reducción de ruido en señales de voz: modelos autorregresivos (LPC-AR) y modelos polo-cero estimados mediante el método de Prony (ARMA). El protocolo experimental consideró el control de la relación señal-ruido segmentada de entrada

REMUNOM
ISSN 2178-6925

Received: 13/01/2026 - **Accepted:** 02/03/2026
Vol: 13.02
DOI: 10.66104/hegxr115
Pages: 1-28

(segSNR_in ≈ 3 dB), múltiples realizaciones estocásticas del ruido y dos escenarios de ruido (blanco y coloreado AR(1)). El desempeño se evaluó mediante métricas complementarias: relación señal-ruido segmentada de salida (segSNR_out), divergencia de Itakura–Saito (IS), inteligibilidad objetiva de corto plazo (STOI) y tiempo de procesamiento. Los resultados se reportaron como media ± desviación estándar y se analizaron utilizando la prueba de Friedman con comparaciones post-hoc. Los resultados indican que el enfoque Kalman+LPC presentó un desempeño global superior y mayor consistencia estadística, especialmente bajo ruido blanco, además de menor costo computacional. La formulación Kalman+Prony mostró estabilidad numérica adecuada, con polos predominantemente dentro del círculo unitario; sin embargo, no evidenció una mejora sistemática respecto al modelo all-pole. El análisis de Pareto evidenció un compromiso entre distorsión espectral y energía residual, sin una ventaja clara del modelo ARMA en las condiciones evaluadas. Se concluye que, para los escenarios analizados, el modelo LPC integrado al filtro de Kalman constituye la alternativa más robusta en términos de desempeño promedio, estabilidad numérica y eficiencia computacional.
Palabras clave: Filtro de Kalman; Supresión de ruido; Mejora de voz; Prony; LPC.

1. Introdução

Nos últimos anos, o uso do filtro de Kalman para aprimoramento de fala tem sido revisitado sob diferentes perspectivas, incluindo cenários com ruído colorido, adaptação paramétrica e integração com técnicas híbridas. Estudos recentes demonstram a aplicação do filtro de Kalman combinado a estratégias espectrais para supressão de ruído não branco (SILVA, 2022), bem como versões adaptativas voltadas à melhoria da robustez estatística em ambientes variáveis (AMIT; VINAY, 2024). Além disso, abordagens contemporâneas têm explorado arquiteturas híbridas que combinam estimadores do tipo Kalman com modelos baseados em aprendizado profundo ou integração audiovisual, ampliando a capacidade de rastreamento do sinal de fala em condições adversas (GOLMAKANI et al., 2022).

Revisões sistemáticas recentes indicam que, apesar da ampla variedade de algoritmos de aprimoramento de fala atualmente disponíveis, incluindo métodos espectrais, estatísticos e baseados em aprendizado profundo, o filtro de Kalman permanece relevante como estrutura interpretável e *model-based*, especialmente quando associado a representações paramétricas do trato vocal (YOUSIF, 2025).

Contudo, observa-se que a formulação predominante na literatura utiliza

REMUNOM
ISSN 2178-6925

Received: 13/01/2026 - **Accepted:** 02/03/2026
Vol: 13.02
DOI: 10.66104/hegxr115
Pages: 1-28

modelos autorregressivos do tipo *all-pole* (LPC) como representação do sinal de fala dentro da estrutura do filtro de Kalman. A incorporação explícita de modelos estimados por métodos clássicos como Prony, ainda é pouco explorada em estudos comparativos sistemáticos no contexto de redução de ruído em fala. Tal lacuna metodológica fundamenta a proposta do presente trabalho, que investiga comparativamente o desempenho das abordagens Kalman+LPC e Kalman+Prony sob diferentes ordens de modelagem e cenários de ruído. Como contribuições, o artigo: (i) propõe e descreve um protocolo reprodutível para estudar o impacto da ordem de Prony na parametrização do filtro de Kalman sob ruído branco e colorido, com repetição estocástica e reporte estatístico; (ii) apresenta comparação experimental direta entre Kalman+Prony e Kalman+LPC sob o mesmo protocolo; e (iii) discute, de forma integrada, desempenho espectral, ganho de relação sinal ruído SNR, estabilidade dos polos e custo computacional em função da ordem do modelo.

2. Revisão da Literatura

2.1 Modelo em espaço de estados do Filtro de Kalman

Rudolf Emil Kalman publicou, em 1960, o artigo "*A New Approach to Linear Filtering and Prediction Problems*", no qual apresentou uma formulação recursiva para a filtragem e a predição linear de sistemas discretos, permitindo atualizar a estimativa do estado e sua incerteza à medida que novas observações são incorporadas (KALMAN, 1960). Desde então, impulsionado pelo desenvolvimento da computação digital e pela ampla disponibilidade de dados, o filtro de Kalman consolidou-se como uma ferramenta fundamental em aplicações que exigem estimação sob incerteza, incluindo navegação e rastreamento, monitoramento e controle de processos, modelagem econômica e reconstrução/atenuação de ruído em sinais.

Neste artigo, o desenvolvimento do filtro de Kalman segue a heurística descrita por Vaseghi (2000). Desta forma o processo de modelagem do sinal de voz, é iniciado por um processo auto-regressivo (*AR*) de ordem *M* (*AR*(*M*)), conforme a Equação (1).

REMUNOM
ISSN 2178-6925

Received: 13/01/2026 - **Accepted:** 02/03/2026
Vol: 13.02
DOI: 10.66104/hegxr115
Pages: 1-28

$$s(n) = \sum_{k=1}^{M} a_k \; s(n-k) + w(n) \tag{1}$$

em que a_k são os coeficientes de predição de ordem *M* e $w(n)$ representa a inovação (erro de predição), associada à excitação no modelo fonte–filtro de produção da fala.

Pode-se observar que no processo de aquisição dos sinais de áudio e voz, a maior parte dos sinais é capturada na presença de algum tipo de ruído que é adicionado ao sinal de voz. Consequentemente, pode-se modelar esse processo de acordo com a Equação (2).

$$y(n) = s(n) + v(n) \tag{2}$$

em que $y(n)$ é o sinal observado e $v(n)$ é um ruído aditivo com variância R.

As Equações (1) e (2), segundo Vaseghi (2000), podem ser reescritas em espaço de estados, definindo-se o vetor de estado:

$$\mathbf{x}(n) = [s(n) \; s(n-1) \; \cdots \; s(n-M+1)]^T \in \mathbb{R}^{M\times 1} \tag{3}$$

Dessa forma, obtém-se:

$$\mathbf{x}(n) = \mathbf{A}\,\mathbf{x}(n-1) + \mathbf{g}\,w(n) \tag{4}$$

$$y(n) = \mathbf{H}\,\mathbf{x}(n) + v(n) \tag{5}$$

REMUNOM
ISSN 2178-6925
Received: 13/01/2026 - **Accepted:** 02/03/2026
Vol: 13.02
DOI: 10.66104/hegxr115
Pages: 1-28

em que: $\mathbf{A} \in \mathbb{R}^{M\times M}$é a matriz de transição de estados, $\mathbf{g} = [1\ 0 \cdots 0]^T \in \mathbb{R}^{M\times 1}$ injeta a inovação no primeiro estado e $\mathbf{H} = [1\ 0 \cdots 0] \in \mathbb{R}^{1\times M}$seleciona o componente correspondente à amostra atual da fala (VASEGHI, 2000).

Assume-se que $w(n)$ e $v(n)$ são ruídos brancos, não correlacionados entre si, com:

$$\mathbb{E}\{v(n)v(k)\} = R\,\delta[n-k] \tag{6}$$

$$\mathbb{E}\{w(n)w(k)\} = Q\,\delta[n-k] \tag{7}$$

onde R e Q são, respectivamente, as variâncias (ou covariâncias) do ruído de medida e da inovação do processo.

Conforme Brown e Hwang (1997) e Vaseghi (2000), o filtro de Kalman é descrito por um conjunto de equações de predição e correção. A predição do estado é dada por:

$$\hat{\mathbf{x}}(n \mid n-1) = \mathbf{A}\,\hat{\mathbf{x}}(n-1 \mid n-1) \tag{8}$$

e a predição da covariância do erro por:

$$\mathbf{P}(n \mid n-1) = \mathbf{A}\,\mathbf{P}(n-1 \mid n-1)\mathbf{A}^T + \mathbf{Q} \tag{9}$$

O ganho de Kalman é calculado por:

$$\mathbf{K}(n) = \mathbf{P}(n \mid n-1)\mathbf{H}^T(\mathbf{H}\mathbf{P}(n \mid n-1)\mathbf{H}^T + R)^{-1} \tag{10}$$

seguido da atualização do estado:

REMUNOM
ISSN 2178-6925
Received: 13/01/2026 - **Accepted:** 02/03/2026
Vol: 13.02
DOI: 10.66104/hegxr115
Pages: 1-28

$$\hat{\mathbf{x}}(n \mid n) = \hat{\mathbf{x}}(n \mid n-1) + \mathbf{K}(n)(y(n) - \mathbf{H}\hat{\mathbf{x}}(n \mid n-1)) \tag{11}$$

e da covariância:

$$\mathbf{P}(n \mid n) = (\mathbf{I} - \mathbf{K}(n)\mathbf{H})\mathbf{P}(n \mid n-1) \tag{12}$$

Nessas expressões, $\mathbf{P}(n \mid n)$ é a matriz de covariância do erro de estimação no instante n, e $\boldsymbol{K}(n)$ regula o compromisso entre a confiança no modelo e a confiança na medida, influenciando diretamente a redução de ruído e a distorção introduzida no sinal estimado.

2.2 Prony

O método de Prony é uma técnica de identificação paramétrica que ajusta um modelo racional IIR a uma sequência discreta, estimando simultaneamente os coeficientes do numerador e do denominador de uma função de transferência (KIM et al., 2025). Diferentemente do modelo autorregressivo clássico (*all-pole*), o modelo estimado por Prony possui estrutura do tipo ARMA, podendo ser expresso como mostrado na Equação 13:

$$H(z) = \frac{B(z)}{A(z)} = \frac{1 + \sum_{k=1}^{N_b} b_k \, z^{-k}}{1 + \sum_{k=1}^{M} a_k \, z^{-k}}, \tag{13}$$

em que A(z)representa o polinômio associado aos polos e B(z)o polinômio associado aos zeros do sistema. A ordem *M* define o grau do denominador e, consequentemente, a dimensionalidade dinâmica do modelo.

No presente trabalho, a modelagem paramétrica não constitui apenas uma etapa de representação espectral, mas influencia diretamente a estrutura do filtro de Kalman. A partir do polinômio estimado A(z), o modelo autorregressivo equivalente pode ser escrito como (Equação 14).

REMUNOM
ISSN 2178-6925

Received: 13/01/2026 - **Accepted:** 02/03/2026
Vol: 13.02
DOI: 10.66104/hegxr115
Pages: 1-28

$$x[n] = -\sum_{k=1}^{M} a_k \, x[n-k] + w[n], \tag{14}$$

onde w[n] representa o termo de excitação (ou inovação) do sistema (Equação 15). A sequência proxy de inovação é obtida aplicando o polinômio estimado ao sinal observado y[n], isto é,

$$w[n] = A(z)\, y[n] = y[n] + \sum_{k=1}^{M} a_k \, y[n-k] \tag{15}$$

o que, computacionalmente, é implementado como $w[n] = \text{filter}(A, 1, y[n])$. A estatística dessa sequência é utilizada para estimar a matriz de covariância de processo (Equação 16),

$$\mathbf{Q}(M) = \text{Cov}(\mathbf{w}_M[n]) \tag{16}$$

de dimensão M×M, podendo incluir regularização diagonal para garantir positividade definida.

Os coeficientes a_k também determinam diretamente a matriz de transição do modelo em espaço de estados. Na forma canônica controlável, a matriz dinâmica é dada pela Equação (17).

$$\mathbf{A}(M) = \begin{bmatrix} -a_1 & -a_2 & \dots & -a_M \\ 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix} \tag{17}$$

Dessa forma, a ordem M influencia simultaneamente (i) a dimensionalidade do vetor de estado, (ii) a estrutura dinâmica do sistema por meio de $\mathbf{A}(M)$, (iii) a estatística da inovação $w[n]$e (iv) a matriz de covariância de processo $\mathbf{Q}(M)$. Como o desempenho do filtro de Kalman é altamente sensível tanto à modelagem dinâmica

REMUNOM
ISSN 2178-6925
Received: 13/01/2026 - **Accepted:** 02/03/2026
Vol: 13.02
DOI: 10.66104/hegxr115
Pages: 1-28

quanto às covariâncias envolvidas, a adoção de um estimador estruturalmente mais flexível capaz de representar polos e zeros, é coerente com o objetivo central do estudo: analisar como a variação da ordem do modelo altera simultaneamente a dinâmica do sistema, a covariância estimada e o desempenho final da filtragem. (BAI et al., 2023).

A escolha do método de Prony em detrimento do LPC para experimentar é fundamentada, primeiramente, pelo fato de o LPC assumir um modelo estritamente *all-pole*, adequado para modelar formantes, mas limitado na representação de antirressonâncias. Revisões recentes sobre modelagem de fala destacam as limitações do paradigma exclusivamente autorregressivo (O'SHAUGHNESSY, 2023). Em segundo lugar, no contexto deste estudo, a ordem do modelo altera diretamente ***A***(M)e ***Q***(M), impactando o comportamento global do filtro de Kalman. Por fim, estudos teóricos recentes demonstram a robustez do método de Prony sob condições ruidosas e discutem sua sensibilidade à ordem em regimes de submodelagem e supermodelagem (BENDORY; DE CASTRO; ELDAR, 2024).

2.3 Codificação por Predição Linear (LPC)

A Codificação por Predição Linear (LPC) é uma técnica clássica de modelagem paramétrica em fala baseada na hipótese de que, em janelas curtas, o sinal pode ser aproximado por uma combinação linear de amostras passadas. Assim, uma amostra *x*[n] é estimada por um preditor de ordem p: $\hat{x}[n] = \sum_{k=1}^{p} a_k\, x[n-k]$, e o resíduo $e[n] = x[n] - \hat{x}[n]$ representa a componente não prevista, frequentemente associada à excitação do sistema produtor de fala (Makhoul, 1975). Ao minimizar o erro quadrático médio, obtêm-se os coeficientes a_k, tipicamente via equações de autocorrelação (Yule–Walker) e solução eficiente por Levinson–Durbin, explorando a estrutura Toeplitz da matriz de autocorrelação (MARKEL; GRAY JR., 1976).

No domínio *Z*, o LPC descreve o trato vocal como um filtro *all-pole* (Equação 18):

REMUNOM
ISSN 2178-6925
Received: 13/01/2026 - **Accepted:** 02/03/2026
Vol: 13.02
DOI: 10.66104/hegxr115
Pages: 1-28

$$X(z) \approx \frac{G}{A(z)} E(z), A(z) = 1 - \sum_{k=1}^{p} a_k z^{-k}, \quad (18)$$

em que os polos de $1/A(z)$ capturam o envelope espectral e as ressonâncias (formantes) da fala, justificando sua ampla adoção em análise e síntese (ATAL; HANAUER, 1971). Entretanto, por ser estruturalmente *all-pole*, o LPC não modela explicitamente zeros espectrais, podendo ter limitações em fonemas com anti-ressonâncias (p.ex., nasais) e em cenários adversos, onde modelos ARMA (polos e zeros) podem oferecer maior flexibilidade (Makhoul, 1975). Neste trabalho, o LPC é empregado como *baseline* (Kalman+LPC) por sua consolidação na literatura e custo computacional relativamente baixo, permitindo comparação direta com a abordagem baseada em Prony/ARMA sob o mesmo protocolo experimental.

3. Materiais e Métodos

3.1 Base de dados e pré-processamento

Foram utilizados 10 sinais de fala isolados (palavras: bala, boi, bola, casa, dado, dia, elefante, elétrica, fala e igreja), amostrados a 22,05 kHz. Cada sinal foi convertido para mono, removida a média e normalizado por seu valor absoluto máximo. Para reduzir efeitos de borda na análise por blocos, o sinal foi preenchido com zeros (*padding*) antes e depois, e a região válida foi definida de forma a manter a extensão temporal do sinal original.

3.2 Geração e calibração de ruído

Foram considerados dois cenários: (i) ruído branco Gaussiano (*white*) e (ii) ruído colorido AR(1) (ar1) com parâmetro α = 0,9. Para cada par (sinal, tipo de ruído) foram geradas K = 20 realizações independentes, totalizando 200 “sujeitos” por condição (10 sinais × 20 realizações). Em cada realização, o ruído foi escalado para manter segSNR_in ≈ 3 dB, via busca binária no ganho do ruído usando a definição segmental com janelamento de 512 amostras e sobreposição de 50%. A mesma

REMUNOM
ISSN 2178-6925
Received: 13/01/2026 - **Accepted:** 02/03/2026
Vol: 13.02
DOI: 10.66104/hegxr115
Pages: 1-28

realização de ruído (mesma semente) foi reutilizada para todas as ordens *M* e para ambos os métodos (Prony e LPC), assegurando comparação pareada.

3.3 Filtro de Kalman e modelos paramétricos (Prony e LPC)

A filtragem foi realizada em regime quase estacionário, utilizando janelas de comprimento N_jan=512 amostras com sobreposição de 50%, correspondente a um deslocamento (*hop size*) de 256 amostras entre blocos consecutivos. Essa segmentação permite assumir estacionariedade local do sinal de fala, condição necessária para a estimação consistente dos modelos paramétricos em cada quadro temporal.

O vetor de estado ***x***[n] possui dimensão *M*, correspondente à ordem do modelo adotado. A equação de observação foi definida na forma da Equação 19.

$$y[n] = \mathbf{C}\mathbf{x}[n] + v[n] \tag{19}$$

em que $\mathbf{C} = [1\ 0\ \ldots\ 0]$seleciona o primeiro componente do vetor de estado como estimativa da amostra observada, e $v[n]$representa o ruído de medição. Essa estrutura corresponde à forma canônica controlável de um sistema autorregressivo.

Para cada bloco, os coeficientes do modelo foram estimados independentemente. No caso do LPC, os coeficientes a_k foram obtidos por meio da função lpc(·, M), resultando em um modelo puramente *all-pole*, cuja dinâmica é descrita pelo polinômio da Equação (20).

$$A(z) = 1 - \sum_{k=1}^{M} a_k\, z^{-k} \tag{20}$$

Para o método de Prony, foram estimados simultaneamente os coeficientes do numerador e do denominador por meio da função prony(·, nb, M), adotando $nb = 2$, caracterizando um modelo ARMA de ordem (nb, M), capaz de representar polos e zeros. Neste caso, os coeficientes do denominador $a_k$foram igualmente utilizados para compor a matriz de transição $\boldsymbol{A}(M)$, enquanto os coeficientes do numerador

REMUNOM
ISSN 2178-6925

Received: 13/01/2026 - **Accepted:** 02/03/2026
Vol: 13.02
DOI: 10.66104/hegxr115
Pages: 1-28

b_k, influenciam indiretamente a estatística da sequência de inovação utilizada na estimação da matriz de covariância de processo $\mathbf{Q}(M)$.

As ordens do modelo foram varridas no conjunto $M \in \{6,8,10,12,14,16,18,20\}$, permitindo analisar o efeito da complexidade do modelo tanto na estrutura dinâmica do sistema quanto na estimação das covariâncias e no desempenho final da filtragem.

3.4 Estimação da covariância de processo e estabilidade de polos

A covariância do ruído de medição foi fixada como Qv = var(n), onde n é o ruído aditivo na região válida. A covariância de processo Qw, foi estimada a partir de uma sequência proxy de inovação w obtida por w = filter(a, 1, y) em cada bloco, seguida da construção de uma matriz de covariância M×M por imersão de estados (vetores atrasados) e cálculo amostral completo, conforme apresentado na seção 2.2. Para Prony, os polos do denominador a(z) foram computados em cada bloco, armazenando-se a distribuição completa por (sinal, ruído, *M*, realização). A estabilidade foi quantificada por max|p| e pelo percentual de polos instáveis (|p| ≥ 1). Para a filtragem (apenas), polos instáveis foram estabilizados por reflexão p ← 1/conj(p), antes da reconstrução de a(z).

3.5 Métricas objetivas e custo computacional

O desempenho foi avaliado por métricas complementares que capturam aspectos energéticos, espectrais, perceptuais e computacionais da filtragem. Todas as medidas foram calculadas na região válida do sinal, desconsiderando efeitos de borda.

(i) segSNR_out

A relação sinal-ruído segmentada de saída (segSNR_out) calcula a SNR em quadros curtos e posteriormente realiza a média aritmética, reduzindo a influência de variações globais de energia (HANSEN; PELL, 1998). Foi definida como a Equação (21).

REMUNOM
ISSN 2178-6925
Received: 13/01/2026 - **Accepted:** 02/03/2026
Vol: 13.02
DOI: 10.66104/hegxr115
Pages: 1-28

$$\text{segSNR}_{out} = \frac{1}{L}\sum_{l=1}^{L} 10\log_{10}\frac{\sum x^2[n]}{\sum(x[n]-\hat{x}[n])^2}. \quad (21)$$

onde L é o número de quadros. Essa métrica é amplamente usada como indicador de atenuação do ruído residual em aprimoramento de fala, sendo frequentemente reportada junto a métricas espectrais e de inteligibilidade.

(ii) Divergência de Itakura–Saito (IS) via PSD (Welch)

A divergência de Itakura–Saito mede discrepância espectral de forma assimétrica, com alta sensibilidade ao envelope espectral (relevante em modelagem paramétrica). Estimando-se as densidades espectrais de potência (PSDs) $S_x(k)$e $S_{\hat{x}}(k)$ por Welch, define-se pela Equação (22):

$$D_{IS}(S_x \parallel S_{\hat{x}}) = \frac{1}{K}\sum_{k=1}^{K}\left(\frac{S_x(k)}{S_{\hat{x}}(k)} - \log\frac{S_x(k)}{S_{\hat{x}}(k)} - 1\right) \quad (22)$$

onde k indexa as bandas/frequências discretas e Ké o número total de bins. Valores menores indicam maior proximidade espectral. Welch é utilizado por reduzir variância da estimativa via média de periodogramas segmentados. (WANG et al., 2024; SELICATO et al., 2025).

(iii) LSD (*Log-Spectral Distance*)

A LSD mede distorção espectral em escala logarítmica (mais alinhada à percepção), sendo comum em estudos recentes de aprimoramento/*denoising*. Uma forma típica é dada pela Equação 23 (GRAY; MARKEL, 1976):

$$\text{LSD} = \sqrt{\frac{1}{K}\sum_{k=1}^{K}(10\log_{10}S_x(k) - 10\log_{10}S_{\hat{x}}(k))^2} \quad (23)$$

REMUNOM
ISSN 2178-6925

Received: 13/01/2026 - **Accepted:** 02/03/2026
Vol: 13.02
DOI: 10.66104/hegxr115
Pages: 1-28

onde $S_x(k)$e $S_{\hat{x}}(k)$ são PSDs (espectros de magnitude/potência, conforme implementado) e K é o número de bins.

(iv) STOI

O intelligibilidade objetiva de curto prazo - STOI (*Short-Time Objective Intelligibility*) estima inteligibilidade com base na correlação entre envelopes temporais em bandas de frequência (TAAL et al., 2011). Assume valores entre 0 e 1, sendo mais correlacionado com testes subjetivos do que métricas puramente energéticas.

(v) Tempo de Execução

O custo computacional foi representado pelo tempo médio (±DP) por configuração. Como o filtro de Kalman envolve operações matriciais cuja complexidade cresce aproximadamente com $O(M^2)$, essa medida permite analisar o impacto da ordem do modelo na eficiência do algoritmo.

3.6 Análise estatística

Como cada par sinal-realização foi avaliado em múltiplas ordens M (medidas repetidas), empregou-se o teste não paramétrico de Friedman para detectar diferenças entre ordens, separadamente por método e tipo de ruído. Quando significativo ($\alpha = 0{,}05$), aplicou-se pós-teste Wilcoxon pareado com correção de Holm. A comparação direta Prony vs. LPC em cada ordem M foi realizada por Wilcoxon signed-rank, também com correção de Holm ao longo das oito ordens. Além dos p-valores, reportou-se o tamanho de efeito $r = \frac{z}{\sqrt{N}}$ e o sinal da diferença mediana (Prony − LPC).

4. Resultados e Discussão

4.1 Visão global do desempenho (média ± DP)

REMUNOM
ISSN 2178-6925

Received: 13/01/2026 - **Accepted:** 02/03/2026
Vol: 13.02
DOI: 10.66104/hegxr115
Pages: 1-28

Os Quadros 1 e 2 apresentam os resultados consolidados (média ± DP) agregando todos os sinais e realizações (*N*=200 por *M*, método (LPC ou Prony) e tipo de ruído (Branco (*white*) ou colorido (ar1)). Assim, o Quadro 1 apresenta o resultado para ruído branco (*white*) e o Quadro 2 para ruído colorido AR(1) (ar1).

Quadro 1: Desempenho consolidado (média ± DP) para ruído branco (*white*), agregando N=200 sujeitos por ordem M.

Método	M	segSNR_out (dB)	IS	STOI	Tempo (s)
LPC	6	7.464 ± 0.396	2.103 ± 0.758	0.9582 ± 0.0173	0.1323 ± 0.0410
LPC	8	7.406 ± 0.432	2.134 ± 0.768	0.9602 ± 0.0163	0.1405 ± 0.0420
LPC	10	7.329 ± 0.456	2.160 ± 0.753	0.9615 ± 0.0162	0.1537 ± 0.0439
LPC	12	7.221 ± 0.465	2.193 ± 0.750	0.9629 ± 0.0157	0.1585 ± 0.0472
LPC	14	7.089 ± 0.463	2.224 ± 0.759	0.9640 ± 0.0158	0.1777 ± 0.0528
LPC	16	6.941 ± 0.453	2.254 ± 0.770	0.9649 ± 0.0159	0.1752 ± 0.0504
LPC	18	6.769 ± 0.443	2.296 ± 0.781	0.9657 ± 0.0161	0.2015 ± 0.0618
LPC	20	6.598 ± 0.428	2.331 ± 0.790	0.9664 ± 0.0162	0.2196 ± 0.0682
Prony	6	7.182 ± 0.405	2.112 ± 0.758	0.9547 ± 0.0176	0.3378 ± 0.0990
Prony	8	7.078 ± 0.456	2.145 ± 0.770	0.9567 ± 0.0162	0.3410 ± 0.0993
Prony	10	6.995 ± 0.474	2.173 ± 0.756	0.9582 ± 0.0158	0.3639 ± 0.1030
Prony	12	6.862 ± 0.480	2.210 ± 0.752	0.9593 ± 0.0159	0.3647 ± 0.1024
Prony	14	6.710 ± 0.524	2.244 ± 0.761	0.9606 ± 0.0157	0.3875 ± 0.1111
Prony	16	6.602 ± 0.496	2.276 ± 0.772	0.9623 ± 0.0159	0.3883 ± 0.1069
Prony	18	6.428 ± 0.479	2.319 ± 0.783	0.9631 ± 0.0162	0.4176 ± 0.1169
Prony	20	6.289 ± 0.467	2.356 ± 0.793	0.9643 ± 0.0161	0.4427 ± 0.1423

Fonte: Elaborado pelos autores (2026).

Analisando o Quadro 1, verifica-se que o aumento da ordem do modelo não resulta em melhoria global de desempenho para nenhum dos métodos. Observa-se redução progressiva da segSNR_out e aumento da divergência de Itakura–Saito (IS) à medida que *M* cresce, indicando que ordens mais elevadas tendem a introduzir maior distorção espectral e menor ganho energético. Embora o índice STOI apresente leve incremento com o aumento de *M*, essa melhoria é marginal e não

REMUNOM
ISSN 2178-6925

Received: 13/01/2026 - **Accepted:** 02/03/2026
Vol: 13.02
DOI: 10.66104/hegxr115
Pages: 1-28

compensa a degradação observada nas demais métricas. Em todas as ordens analisadas, o método baseado em LPC apresenta desempenho superior ao Prony em segSNR_out, menor IS e maior STOI, além de custo computacional substancialmente inferior (aproximadamente duas vezes mais rápido). De forma geral, os resultados sugerem que, sob ruído branco, ordens mais baixas (como M = 6 ou M = 8), oferecem melhor compromisso entre desempenho e complexidade, e que o modelo *all-pole* (LPC) é mais eficiente e robusto nesse cenário.

Quadro 2: Desempenho consolidado (média ± DP) para ruído colorido AR(1) (ar1), agregando N=200 sujeitos por ordem M.

Método	M	segSNR_out (dB)	IS	STOI	Tempo (s)
LPC	6	4.971 ± 0.514	0.799 ± 0.228	0.9028 ± 0.0310	0.1343 ± 0.0445
LPC	8	4.946 ± 0.470	0.697 ± 0.200	0.9106 ± 0.0294	0.1407 ± 0.0430
LPC	10	4.916 ± 0.447	0.681 ± 0.201	0.9146 ± 0.0293	0.1526 ± 0.0434
LPC	12	4.883 ± 0.430	0.681 ± 0.216	0.9177 ± 0.0284	0.1591 ± 0.0487
LPC	14	4.777 ± 0.461	0.920 ± 2.093	0.9181 ± 0.0313	0.1793 ± 0.0500
LPC	16	4.749 ± 0.411	0.663 ± 0.199	0.9226 ± 0.0274	0.1745 ± 0.0520
LPC	18	4.699 ± 0.397	0.656 ± 0.203	0.9251 ± 0.0268	0.2016 ± 0.0585
LPC	20	4.630 ± 0.383	0.639 ± 0.191	0.9266 ± 0.0269	0.2167 ± 0.0638
Prony	6	4.230 ± 0.808	2.364 ± 7.270	0.8884 ± 0.0511	0.3390 ± 0.1071
Prony	8	4.199 ± 0.762	2.198 ± 9.209	0.8974 ± 0.0497	0.3497 ± 0.1184
Prony	10	4.180 ± 0.661	2.124 ± 10.974	0.9011 ± 0.0497	0.3654 ± 0.0992
Prony	12	4.261 ± 0.573	0.925 ± 2.107	0.9106 ± 0.0370	0.3637 ± 0.1034
Prony	14	4.185 ± 0.527	2.009 ± 8.310	0.9090 ± 0.0572	0.3902 ± 0.1077
Prony	16	4.136 ± 0.543	0.973 ± 4.183	0.9179 ± 0.0320	0.3881 ± 0.1103
Prony	18	4.132 ± 0.499	0.649 ± 0.201	0.9220 ± 0.0268	0.4205 ± 0.1132
Prony	20	4.066 ± 0.450	0.652 ± 0.210	0.9236 ± 0.0269	0.4355 ± 0.1206

Fonte: Elaborado pelos autores (2026).

REMUNOM
ISSN 2178-6925

Received: 13/01/2026 - **Accepted:** 02/03/2026
Vol: 13.02
DOI: 10.66104/hegxr115
Pages: 1-28

Verificando o Quadro 2, percebe-se um comportamento mais “rico” do que no ruído branco, porque aparece claramente o *trade-off* entre métricas energéticas (segSNR_out) e métricas espectrais (IS), além de maior variabilidade no Prony.

Já no LPC, a segSNR_out decresce de forma quase monotônica com M (4,971 dB em $M = 6 \rightarrow$ 4,630 dB em $M = 20$), enquanto o IS melhora com o aumento de M (cai de 0,799 → 0,639), sugerindo que ordens maiores favorecem a aproximação espectral do sinal (menor divergência), porém sem traduzir isso em ganho energético. O STOI aumenta gradualmente (0,9028 → 0,9266), indicando melhora consistente na medida objetiva associada à inteligibilidade; como esperado, o tempo também cresce com M (≈ 0,134 s → 0,217 s).

No Prony, a segSNR_out é sistematicamente inferior à do LPC em todas as ordens (por exemplo, $M = 6$: 4,230 vs 4,971 dB), e apresenta maior dispersão (DP elevado), indicando maior sensibilidade estocástica do ARMA estimado em blocos curtos sob ruído colorido. O aspecto mais relevante é o IS: em ordens baixas ($M = 6$ –10 e $M = 14$), o IS médio é alto e o desvio padrão é considerado (ex.: 2,364 ± 7,270), sugerindo ocorrência de *outliers*/instabilidades numéricas em parte das realizações; já em ordens mais altas, especialmente $M = 18 – 20$, o IS torna-se baixo e estável (0,649 ± 0,201 e 0,652 ± 0,210), aproximando-se do desempenho do LPC e evidenciando que o Prony pode melhorar o ajuste espectral em ruído colorido quando M é suficientemente grande. O STOI também cresce com M (atingindo ≈0,922–0,924 em $M = 18 – 20$), mas o tempo permanece cerca de 2 a 3 vezes maior que no LPC em toda a faixa (≈ 0,34 – 0,44 s).

Em suma, o Quadro 2 indica que, sob ruído colorido AR(1), o LPC é mais vantajoso quando o objetivo é preservação energética (segSNR_out) e custo computacional, enquanto o Prony tende a se tornar competitivo em termos de distorção espectral (IS) apenas em ordens mais altas (principalmente $M \geq 18$), ao custo de menor segSNR_out e maior tempo. Além disso, os desvios padrão muito elevados do IS em algumas ordens do Prony reforçam a necessidade de discutir estabilidade/robustez para caracterizar corretamente o comportamento em ruído colorido.

A escolha da ordem do modelo M pode ser interpretada como um problema de otimização multiobjetivo, no qual critérios potencialmente conflitantes devem ser

REMUNOM
ISSN 2178-6925

Received: 13/01/2026 - **Accepted:** 02/03/2026
Vol: 13.02
DOI: 10.66104/hegxr115
Pages: 1-28

considerados simultaneamente. No presente estudo, dois objetivos principais foram analisados: (i) a minimização da divergência de Itakura–Saito (IS), associada à preservação do envelope espectral, e (ii) a maximização da segSNR_out, relacionada ao ganho energético após a filtragem. Como não existe, a priori, um único valor de M que maximize ambos os critérios de forma independente, emprega-se a análise por fronteira de Pareto, na qual cada ponto no plano (IS, segSNR_out) representa uma ordem específica do modelo. Soluções são consideradas Pareto-ótimas quando não é possível melhorar um objetivo sem deteriorar o outro. A Figura 1 apresenta essa análise para o cenário de ruído branco.

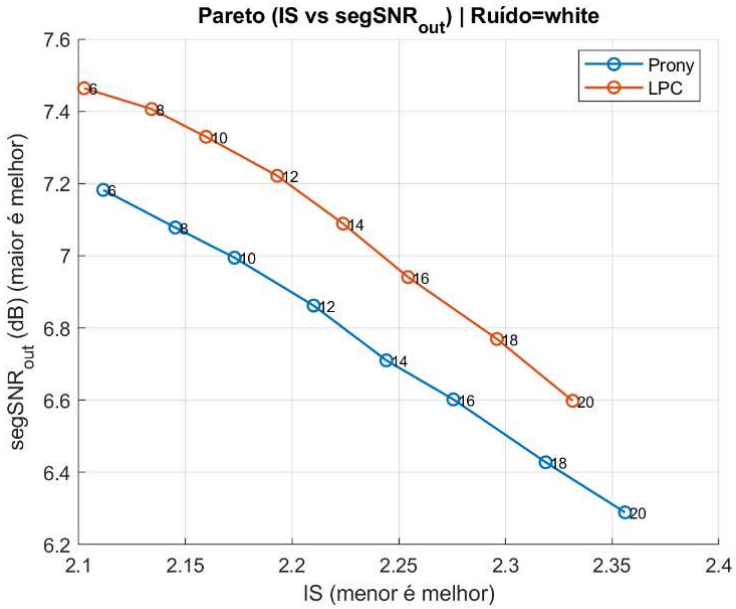
No cenário de ruído branco, observa-se na Figura 1 o comportamento aproximadamente monotônico para ambos os métodos. À medida que a ordem M aumenta, verifica-se deslocamento consistente dos pontos para a direita (aumento de IS) e para baixo (redução de segSNR_out), indicando deterioração simultânea das duas métricas. Esse padrão sugere que, sob ruído branco e processamento em blocos curtos, o aumento da complexidade do modelo não produz ganho espectral compensatório, mas sim efeitos associados à superparametrização.

Comparativamente, a curva correspondente ao método LPC permanece sistematicamente acima da curva do Prony ao longo de toda a faixa de ordens analisadas. Isso implica que, para um mesmo nível de IS, o LPC apresenta maior segSNR_out; alternativamente, para um mesmo nível de segSNR_out, o LPC mantém menor distorção espectral. Do ponto de vista multiobjetivo, tal comportamento caracteriza uma dominância prática do baseline *all-pole* no plano energia – distorção.

REMUNOM
ISSN 2178-6925

Received: 13/01/2026 - **Accepted:** 02/03/2026
Vol: 13.02
DOI: 10.66104/hegxr115
Pages: 1-28

Figura 1 – Curva de Pareto entre divergência de Itakura–Saito (IS) e segSNR_out sob ruído branco.



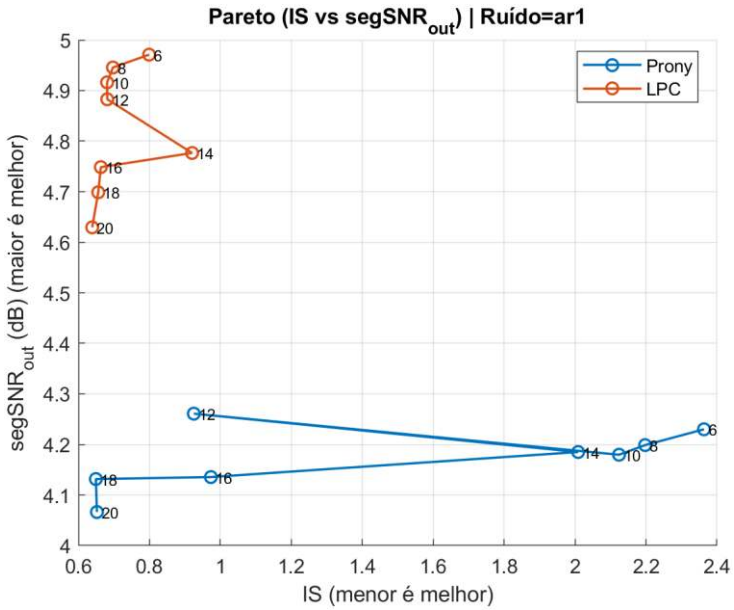
Fonte: Elaborado pelos autores (2026).

Observa-se (Figura 1) ainda que as melhores configurações concentram-se em ordens baixas (particularmente $M = 6$), localizadas na região superior esquerda da figura, onde coexistem menor IS e maior segSNR_out. Ordens mais elevadas deslocam-se progressivamente para regiões menos favoráveis do plano de Pareto. Assim, para ruído branco, a evidência experimental indica que modelos de baixa ordem oferecem o melhor compromisso entre preservação espectral e ganho energético, sendo o método LPC–Kalman o que apresenta desempenho globalmente superior nesse cenário.

A análise de Pareto também foi realizada para o cenário de ruído colorido AR(1), no qual a estrutura espectral do ruído é predominantemente autoregressiva. Nesse caso, a relação entre divergência espectral (IS) e ganho energético (segSNR_out) pode apresentar comportamento distinto daquele observado no ruído branco. A Figura 2 apresenta o plano (IS, segSNR_out) para as ordens $M \in \{6, 8, \ldots, 20\}$, permitindo avaliar o compromisso entre preservação do envelope espectral e desempenho energético em cada método.

REMUNOM
ISSN 2178-6925
Received: 13/01/2026 - **Accepted:** 02/03/2026
Vol: 13.02
DOI: 10.66104/hegxr115
Pages: 1-28

Figura 2 – Curva de Pareto entre divergência de Itakura–Saito (IS) e segSNR_out sob ruído colorido (ar1).



Fonte: Elaborado pelos autores (2026).

Diferentemente do comportamento monotônico observado no ruído branco, o cenário AR(1) (Figura2), revela um *trade-off* mais pronunciado entre IS e segSNR_out, especialmente no método baseado em Prony. Para o LPC, observa-se tendência relativamente organizada: à medida que M aumenta, a IS tende a reduzir (melhor ajuste espectral), enquanto a segSNR_out apresenta leve decréscimo. Esse comportamento caracteriza um compromisso gradual entre fidelidade espectral e ganho energético, com desempenho energético sistematicamente superior ao Prony ao longo de toda a faixa de ordens.

No caso do Prony, entretanto, a distribuição dos pontos no plano de Pareto é menos regular. Em ordens baixas (por exemplo, $M = 6, 8, 10$), observa-se maior IS e segSNR_out relativamente moderada. Já em ordens mais altas ($M = 18$ e $M = 20$), ocorre redução significativa da IS, aproximando-se dos valores obtidos pelo LPC, porém com segSNR_out inferior. Esse comportamento indica que o aumento da ordem no modelo ARMA favorece o ajuste espectral em ruído colorido, mas não se traduz em ganho energético correspondente.

REMUNOM
ISSN 2178-6925

Received: 13/01/2026 - **Accepted:** 02/03/2026
Vol: 13.02
DOI: 10.66104/hegxr115
Pages: 1-28

Em termos de dominância multiobjetivo, o LPC mantém-se globalmente em região mais favorável do plano (IS, segSNR_out), combinando menor IS e maior segSNR_out na maior parte das ordens. Ainda assim, o Prony evidencia capacidade de reduzir a divergência espectral em ordens elevadas, sugerindo que, sob ruído estruturalmente auto-regressivo, modelos ARMA de maior complexidade podem capturar melhor certas componentes espectrais, embora com penalização energética.

Assim, para o ruído colorido (ar1) a escolha da ordem M envolve um compromisso real entre energia e fidelidade espectral, sendo que o baseline LPC–Kalman apresenta melhor equilíbrio global, enquanto o Prony–Kalman pode tornar-se competitivo em termos espectrais apenas em ordens mais altas e sob maior custo computacional.

4.2 Resultados por palavra e variabilidade

O Quadro 3 reporta, para cada palavra, o valor de M* (ótimo) que minimiza IS (com desempate por maior segSNR_out e menor tempo), bem como as métricas correspondentes (média ± DP). Observa-se que o M* pode variar substancialmente entre palavras, sobretudo em ruído colorido (ar1), confirmando que conclusões baseadas em poucas amostras podem ser enviesadas pelo conteúdo fonético.

Quadro 3: Resultados por palavra (média ± DP) no ponto M* que minimiza IS (desempate por maior segSNR_out e menor tempo), para cada ruído e método.

Sinal	Ruído	Método	M*	IS (μ±DP)	segSNR_out (dB) (μ±DP)	STOI (μ±DP)	Tempo (s) (μ±DP)
bala	ar1	LPC	8	0.819 ± 0.032	5.436 ± 0.237	0.9054 ± 0.0057	0.1507 ± 0.0238
bala	ar1	Prony	8	0.909 ± 0.026	5.063 ± 0.263	0.9048 ± 0.0055	0.3488 ± 0.0307
bala	white	LPC	6	2.997 ± 0.015	7.647 ± 0.083	0.9675 ± 0.0026	0.1447 ± 0.0270
bala	white	Prony	6	2.999 ± 0.014	7.557 ± 0.097	0.9645 ± 0.0028	0.3537 ± 0.0596
boi	ar1	LPC	20	0.774 ± 0.039	5.071 ± 0.170	0.9445 ± 0.0063	0.1510 ± 0.0373

REMUNOM
ISSN 2178-6925

Received: 13/01/2026 - **Accepted:** 02/03/2026
Vol: 13.02
DOI: 10.66104/hegxr115
Pages: 1-28

boi	ar1	Prony	18	0.761 ± 0.039	4.041 ± 0.485	0.9321 ± 0.0097	0.3006 ± 0.0635
boi	white	LPC	6	2.800 ± 0.025	8.022 ± 0.066	0.9774 ± 0.0022	0.0833 ± 0.0091
boi	white	Prony	6	2.811 ± 0.023	7.019 ± 0.330	0.9678 ± 0.0037	0.2100 ± 0.0218
bola	ar1	LPC	18	0.804 ± 0.035	4.680 ± 0.161	0.9059 ± 0.0048	0.1899 ± 0.0154
bola	ar1	Prony	14	0.798 ± 0.034	4.158 ± 0.189	0.9003 ± 0.0047	0.3750 ± 0.0361
bola	white	LPC	6	3.390 ± 0.018	7.761 ± 0.069	0.9683 ± 0.0020	0.1235 ± 0.0218
bola	white	Prony	6	3.404 ± 0.020	7.559 ± 0.088	0.9647 ± 0.0027	0.3261 ± 0.0374
casa	ar1	LPC	20	0.356 ± 0.020	4.761 ± 0.099	0.9282 ± 0.0079	0.1538 ± 0.0155
casa	ar1	Prony	20	0.358 ± 0.018	4.538 ± 0.102	0.9320 ± 0.0077	0.3187 ± 0.0296
casa	white	LPC	6	1.043 ± 0.013	7.372 ± 0.045	0.9694 ± 0.0036	0.0939 ± 0.0190
casa	white	Prony	6	1.047 ± 0.013	7.347 ± 0.037	0.9694 ± 0.0035	0.2464 ± 0.0388
dado	ar1	LPC	10	0.657 ± 0.015	5.528 ± 0.085	0.8764 ± 0.0064	0.1612 ± 0.0126
dado	ar1	Prony	18	0.712 ± 0.016	4.829 ± 0.171	0.8764 ± 0.0057	0.4295 ± 0.0420
dado	white	LPC	6	1.992 ± 0.007	7.294 ± 0.044	0.9178 ± 0.0053	0.1393 ± 0.0151
dado	white	Prony	6	2.001 ± 0.007	7.076 ± 0.061	0.9114 ± 0.0043	0.3421 ± 0.0287
dia	ar1	LPC	6	0.343 ± 0.018	4.478 ± 0.101	0.9207 ± 0.0053	0.1161 ± 0.0110
dia	ar1	Prony	6	0.352 ± 0.017	3.840 ± 0.244	0.9119 ± 0.0050	0.2805 ± 0.0146
dia	white	LPC	8	1.454 ± 0.013	6.796 ± 0.078	0.9512 ± 0.0028	0.1226 ± 0.0179
dia	white	Prony	8	1.461 ± 0.013	6.647 ± 0.104	0.9457 ± 0.0040	0.3039 ± 0.0266

REMUNOM
ISSN 2178-6925

Received: 13/01/2026 - **Accepted:** 02/03/2026
Vol: 13.02
DOI: 10.66104/hegxr115
Pages: 1-28

elefante	ar1	LPC	20	0.725 ± 0.066	4.172 ± 0.094	0.9607 ± 0.0019	0.3163 ± 0.0363
elefante	ar1	Prony	20	0.632 ± 0.051	3.693 ± 0.123	0.9602 ± 0.0021	0.6406 ± 0.0545
elefante	white	LPC	6	1.350 ± 0.010	7.965 ± 0.134	0.9422 ± 0.0063	0.1861 ± 0.0188
elefante	white	Prony	6	1.356 ± 0.011	7.827 ± 0.138	0.9413 ± 0.0063	0.4804 ± 0.0356
elétrica	ar1	LPC	18	0.498 ± 0.013	4.125 ± 0.112	0.9593 ± 0.0026	0.2769 ± 0.0360
elétrica	ar1	Prony	12	0.516 ± 0.013	3.760 ± 0.184	0.9566 ± 0.0030	0.5106 ± 0.0332
elétrica	white	LPC	6	2.063 ± 0.011	6.922 ± 0.153	0.9652 ± 0.0055	0.1863 ± 0.0215
elétrica	white	Prony	6	2.073 ± 0.012	6.702 ± 0.168	0.9643 ± 0.0052	0.4740 ± 0.0433
fala	ar1	LPC	10	0.580 ± 0.027	5.348 ± 0.136	0.8746 ± 0.0102	0.0992 ± 0.0075
fala	ar1	Prony	8	0.650 ± 0.022	5.133 ± 0.176	0.8723 ± 0.0077	0.2329 ± 0.0294
fala	white	LPC	6	2.522 ± 0.020	7.356 ± 0.065	0.9521 ± 0.0046	0.0851 ± 0.0097
fala	white	Prony	6	2.529 ± 0.020	7.271 ± 0.078	0.9512 ± 0.0046	0.2290 ± 0.0255
igreja	ar1	LPC	20	0.466 ± 0.040	4.857 ± 0.068	0.9533 ± 0.0030	0.2668 ± 0.0390
igreja	ar1	Prony	20	0.450 ± 0.034	4.055 ± 0.266	0.9485 ± 0.0030	0.5161 ± 0.0501
igreja	white	LPC	6	1.404 ± 0.013	7.513 ± 0.039	0.9704 ± 0.0016	0.1589 ± 0.0266
igreja	white	Prony	6	1.421 ± 0.014	6.828 ± 0.177	0.9658 ± 0.0024	0.4030 ± 0.0640

Fonte: Elaborado pelos autores (2026).

A síntese dos “melhores casos” por palavra (Quadro 3) indica que Kalman+LPC é no conjunto avaliado, a técnica mais consistente e eficiente: em ruído branco, o LPC mantém IS muito próxima (frequentemente ligeiramente menor) e entrega segSNR_out maior que o Prony em praticamente todas as palavras, além de exigir tempo de processamento significativamente menor (tipicamente ~ 2 – 3× mais

REMUNOM
ISSN 2178-6925

Received: 13/01/2026 - **Accepted:** 02/03/2026
Vol: 13.02
DOI: 10.66104/hegxr115
Pages: 1-28

rápido). Em ruído colorido, há poucos casos pontuais em que o Prony apresenta IS marginalmente menor (p.ex., *boi*, *elefante*, *igreja*), porém esse ganho espectral não se traduz em melhoria energética e/ou de inteligibilidade: em geral o Prony vem acompanhado de segSNR_out inferior e custo computacional maior, reforçando a superioridade global do baseline *all-pole* para as condições testadas.

Quanto à inteligibilidade, os valores de STOI obtidos após a filtragem são altos em ambos os métodos (aproximadamente 0,87 a 0,98 dependendo da palavra e do ruído), o que sugere que os sinais recuperados preservam bem os padrões de curto prazo associados à inteligibilidade, isto é, ficam muito próximos do sinal limpo (antes da contaminação) do ponto de vista perceptual objetivo. Além disso, as diferenças de STOI entre LPC e Prony nos melhores M^* são na maioria das palavras, pequenas (da ordem de milésimos a centésimos), indicando que ambas as abordagens recuperam fala com inteligibilidade comparável, mas o LPC se destaca por manter desempenho global mais favorável (melhor compromisso IS, segSNR_out) com menor tempo de execução.

4.3 Análise de estabilidade de Prony

A análise de estabilidade dos polos do modelo ARMA estimado por Prony (Quadros 4 e 5) indica comportamento predominantemente estável em todas as ordens avaliadas ($M \in \{6, 8, \ldots, 20\}$) e para ambos os ruídos. Em termos agregados (todos os sinais e realizações), o maior módulo de polo permaneceu abaixo do círculo unitário, com max$|\ p\ |$ médio variando aproximadamente de 0,960 a 0,988, e desvios-padrão reduzidos. Para ruído colorido, não foram observados polos instáveis (0% em todas as ordens e nenhuma execução com $|\ p\ | \geq 1$). Para ruído branco, a instabilidade também foi praticamente ausente, ocorrendo apenas de forma esparsa nas maiores ordens (em $M = 18$ e $M = 20$), com 0,5% e 1,0% das execuções apresentando ao menos um polo instável e percentuais médios de instabilidade próximos de zero.

REMUNOM
ISSN 2178-6925

Received: 13/01/2026 - **Accepted:** 02/03/2026
Vol: 13.02
DOI: 10.66104/hegxr115
Pages: 1-28

Quadro 4: Estabilidade dos polos (Prony) – Ruído Colorido AR(1).

M	N	max|p| (média ± DP)	% polos instáveis (média ± DP)	Proporção execuções com |p|≥1
6	200	0,965 ± 0,017	0,000 ± 0,000	0,000
8	200	0,975 ± 0,007	0,000 ± 0,000	0,000
10	200	0,976 ± 0,010	0,000 ± 0,000	0,000
12	200	0,981 ± 0,010	0,000 ± 0,000	0,000
14	200	0,983 ± 0,008	0,000 ± 0,000	0,000
16	200	0,984 ± 0,010	0,000 ± 0,000	0,000
18	200	0,986 ± 0,009	0,000 ± 0,000	0,000
20	200	0,988 ± 0,008	0,000 ± 0,000	0,000

Fonte: Elaborado pelos autores (2026).

Quadro 5: Estabilidade dos polos (Prony) – Ruído Branco.

M	N	max|p| (média ± DP)	% polos instáveis (média ± DP)	Proporção execuções com |p|≥1
6	200	0,960 ± 0,020	0,000 ± 0,000	0,000
8	200	0,970 ± 0,014	0,000 ± 0,000	0,000
10	200	0,976 ± 0,013	0,000 ± 0,000	0,000
12	200	0,979 ± 0,010	0,000 ± 0,000	0,000
14	200	0,984 ± 0,009	0,000 ± 0,000	0,000
16	200	0,985 ± 0,009	0,000 ± 0,000	0,000
18	200	0,986 ± 0,010	0,001 ± 0,016	0,005
20	200	0,987 ± 0,010	0,002 ± 0,021	0,010

Fonte: Elaborado pelos autores (2026).

4.5 Significância estatística

O Quadro 6 resume os testes estatísticos. Para ambos os tipos de ruído, o teste de Friedman indica diferenças estatisticamente significativas entre as ordens *M* dentro de cada método ($p<0,05$ para todas as métricas).

Quadro 6: Síntese dos testes estatísticos: Friedman (diferenças entre ordens *M* dentro de cada método) e Wilcoxon pareado Prony vs LPC por *M* (Holm).

Ruído	Métrica	M_sig(Prony≠LPC)	Mediana(Prony-LPC)	Friedman p (LPC)	Friedman p (Prony)
ar1	segSNR_out	8/8	-0.5196	1.12122047085914 08e-158	1.95349630082663 94e-33
ar1	IS_out	2/8	0.0064	2.63028774987055 2e-43	2.90013406877246 66e-26
ar1	LSD_out	7/8	0.1713	4.00647670735396 1e-134	2.67554487971645 1e-117

REMUNOM
ISSN 2178-6925

Received: 13/01/2026 - **Accepted:** 02/03/2026
Vol: 13.02
DOI: 10.66104/hegxr115
Pages: 1-28

ar1	STOI	8/8	-0.0024	9.56068063818664 8e-201	3.86543465571043 96e-165
ar1	Tempo_s	8/8	0.2027	1.20082504186489 74e-206	2.42112363694995 95e-165
white	segSNR_out	8/8	-0.1960	5.56083143626987 7e-289	3.40649083581232 9e-226
white	IS_out	8/8	0.0180	4.68685511660261 84e-294	2.26648389440652 7e-294
white	LSD_out	8/8	0.1035	6.32634114438543 6e-298	7.43829790026042 1e-298
white	STOI	8/8	-0.0023	8.97068859415494 7e-249	2.27285744043749 68e-185
white	Tempo_s	8/8	0.2050	1.73982130398161 33e-199	7.95183697794535 e-164

Fonte: Elaborado pelos autores (2026).

Analisando o Quadro 6, na comparação direta Prony vs LPC por ordem *M* (Wilcoxon pareado com correção de Holm), observou-se diferença significativa em 8/8 ordens para segSNR_out, STOI e tempo em ambos os ruídos; para IS em ar1, apenas 2/8 ordens apresentaram diferença significativa, enquanto em *white*, a diferença foi significativa em todas as ordens. O sinal da mediana (Prony − LPC) indica, de forma consistente, maior tempo computacional para Prony e desempenho inferior em segSNR_out e STOI na maioria dos cenários, ao passo que as diferenças em IS dependem do ruído e da ordem. O teste de Friedman indicou efeito significativo da ordem M sobre todas as métricas ($p \ll 0{,}05$) em ambos os métodos, confirmando que a escolha de M impacta de forma sistemática o desempenho.

5 Conclusão

O presente estudo investigou, de forma estatisticamente controlada, o impacto da modelagem paramétrica AR (LPC) e ARMA (Prony) quando integradas ao filtro de Kalman para redução de ruído em sinais de fala. Diferentemente de abordagens predominantemente descritivas encontradas na literatura, foi adotado um protocolo experimental reprodutível, com controle rigoroso da segSNR de entrada, múltiplas realizações estocásticas e análise estatística não paramétrica

REMUNOM
ISSN 2178-6925

Received: 13/01/2026 - **Accepted:** 02/03/2026
Vol: 13.02
DOI: 10.66104/hegxr115
Pages: 1-28

para medidas repetidas.

Os resultados evidenciaram que a ampliação estrutural proporcionada pelo modelo ARMA não implica, necessariamente, melhoria sistemática no desempenho do filtro de Kalman. Embora o método de Prony tenha apresentado estabilidade numérica adequada e maior flexibilidade estrutural (representação de polos e zeros), o modelo LPC demonstrou desempenho energético e espectral globalmente superior, especialmente sob ruído branco, além de menor custo computacional. Em ruído colorido AR(1), o Prony mostrou-se competitivo apenas em ordens mais elevadas e predominantemente na métrica espectral (IS), sem ganhos consistentes em segSNR_out ou STOI.

A análise multiobjetivo via fronteira de Pareto confirmou que o aumento da ordem do modelo tende a deslocar o sistema para regiões menos favoráveis do plano energia/distorção, caracterizando possível superparametrização em blocos curtos. Os testes de Friedman e Wilcoxon corroboraram que tanto a escolha da ordem quanto o tipo de modelagem influenciam significativamente o desempenho, mas sem evidência de dominância prática do modelo ARMA nas condições analisadas.

Conclui-se que, para o regime de processamento em blocos curtos e sob controle estatístico rigoroso, o modelo *all-pole* (LPC) integrado ao filtro de Kalman constitui a alternativa mais robusta e eficiente. O estudo contribui ao esclarecer, com base experimental e estatística, os limites e potencialidades da inclusão explícita de zeros (ARMA) na estrutura de estimação de estado para aprimoramento de fala.

Referências

ATAL, B. S.; HANAUER, S. L. Speech analysis and synthesis by linear prediction of the speech wave. ***The Journal of the Acoustical Society of America***, v. 50, n. 2B, p. 637–655, 1971. DOI: https://doi.org/10.1121/1.1912679

AMIT, K. S.; VINAY, K. S. Noise free speech enhancement based on fast adaptive Kalman filtering algorithm. ***International Journal of Computer Applications***, v. 185, n. 9, p. 1–7, 2024. Disponível em: https://research.ijcaonline.org/volume93/number9/pxc3895805.pdf. Acesso em: 21 fev. 2026.

REMUNOM
ISSN 2178-6925

Received: 13/01/2026 - **Accepted:** 02/03/2026
Vol: 13.02
DOI: 10.66104/hegxr115
Pages: 1-28

BAI, Y. et al. State of the art on state estimation: Kalman filter driven by machine learning. ***Annual Reviews in Control***, v. 56, p. 100909, 2023. DOI: https://doi.org/10.1016/j.arcontrol.2023.100909

BENDORY, T.; DE CASTRO, Y.; ELDAR, Y. C. On the accuracy of Prony's method for recovery of sparse measures from noisy frequency samples. *arXiv*, 2024. Disponível em: https://arxiv.org/abs/2302.05883. Acesso em: 21 fev. 2026.

BROWN, R. G.; HWANG, P. Y. C. *Introduction to random signals and applied Kalman filtering*. 3. ed. New York: John Wiley & Sons, 1997.

GRAY, A. H.; MARKEL, J. D. Distance measures for speech processing. *IEEE Transactions on Acoustics, **Speech, and Signal Processing***, v. 24, n. 5, p. 380–391, 1976. DOI: https://doi.org/10.1109/TASSP.1976.1162849

KALMAN, R. E. A new approach to linear filtering and prediction problems. ***Transactions of the ASME – Journal of Basic Engineering***, v. 82, p. 35–45, 1960. DOI: https://doi.org/10.1115/1.3662552

KIM, K. et al. Neural network regression for sound source localization using classical spectral estimation techniques (Yule–Walker, Prony, Steiglitz–McBride). ***Applied Sciences***, v. 15, n. 17, 2025. Disponível em: https://www.mdpi.com/2076-3417/15/17/9272. Acesso em: 21 fev. 2026.

MAKHOUL, J. Linear prediction: a tutorial review. ***Proceedings of the IEEE***, v. 63, n. 4, p. 561–580, 1975. DOI: https://doi.org/10.1109/PROC.1975.9792

MARKEL, J. D.; GRAY JR., A. H. *Linear prediction of speech*. New York: Springer, 1976.

O'SHAUGHNESSY, D. Review of methods for coding of speech signals. ***EURASIP Journal on Audio, Speech, and Music Processing***, 2023. DOI: https://doi.org/10.1186/s13636-023-00274-x

SILVA, L. A. da. Utilização dos filtros de Kalman de tempo discreto e frequência combinados com subtração espectral de potência na redução de ruído aditivo em sinais de voz contaminados por ruído colorido. ***Research, Society and Development***, v. 11, n. 7, 2022. DOI: https://doi.org/10.33448/rsd-v11i7.30369

TAAL, C. H. et al. An algorithm for intelligibility prediction of time–frequency weighted noisy speech. ***IEEE Transactions on Audio, Speech, and Language Processing***, v. 19, n. 7, p. 2125–2136, 2011. DOI: https://doi.org/10.1109/TASL.2011.2114881

VASEGHI, S. V. *Advanced digital signal processing and noise reduction*. 2. ed. New York: John Wiley & Sons, 2000.

WANG, J. et al. Independent low-rank matrix analysis for determined blind source separation using Itakura–Saito divergence. *arXiv*, 2024. Disponível em: https://arxiv.org/pdf/2401.01762. Acesso em: 21 fev. 2026.

YOUSIF, S. T. Speech enhancement algorithms: a systematic literature review. *Algorithms*, v. 18, n. 5, 2025. DOI: https://doi.org/10.3390/a18050272