

# Nouvelle Entropie de permutation basée sur le déterminant de matrices circulantes

Meryem JABLOUN

Laboratoire PRISME - équipe signal, 12 rue de Blois, 45000 Orléans, France  
meryem.jabloun@univ-orleans.fr

**Résumé** – Au cours des deux dernières décennies, un nombre croissant de variantes de l’entropie de permutation (PE) ont été proposées pour mesurer la complexité des signaux. Leur principe fondamental repose sur l’identification des motifs d’ordre, mais peu d’entre elles ont pris en compte l’information relative à la variation de l’amplitude. Dans cet article, nous proposons une nouvelle variante de PE qui prend explicitement en compte à la fois l’amplitude et l’information spectrale. Cette nouvelle PE bénéficie d’une part des propriétés spectrales des déterminants des matrices circulantes, et d’autre part des propriétés filtrantes des motifs d’ordre que nous avons établies dans une étude précédente. Lorsqu’elle est appliquée à des signaux simulés, cette nouvelle PE surpasse les variantes de PE existantes qui intègrent l’information sur l’amplitude.

**Abstract** – Over the past two decades, an increasing number of permutation entropy (PE) variants have been proposed to measure signal complexity. Their fundamental principle is primarily based on the identification of order patterns, though fewer have incorporated amplitude information. In this paper, we introduce a new PE variant that explicitly accounts for both amplitude and spectral information by leveraging the well-established spectral properties of circulant determinants, as well as the spectral properties we previously established for  $d$ -dimensional order patterns. Simulated signals are used to illustrate the results, comparing our approach to existing amplitude-aware PE variants.

## 1 Introduction

Parmi les approches statistiques en traitement du signal, l’entropie de permutation (PE) et ses diverses extensions ont émergé comme des mesures de complexité prometteuses [1]. En classant chaque  $d$  échantillons successifs d’un signal donné, par ordre croissant, la PE transforme efficacement et sans seuillage des signaux faiblement non stationnaires en des séquences de codes, appelés motifs d’ordre (OPs) [2]. La distribution de probabilité de ces OPs est estimée à l’aide de leur fréquence d’apparition. L’application de l’entropie de Shannon à cette distribution des OPs définit donc la PE.

Étant donné que la PE repose uniquement sur les OPs et ignore l’information d’amplitude, quelques extensions de la PE incorporant l’amplitude, telles que la PE pondérée (ImpPE) [3, 4, 5, 6, 7], la PE sensible à l’amplitude (AAPE) [8, 9], et l’entropie de dispersion (DispEn) [10, 11] ont été proposées pour palier à ce problème.

La ImpPE attribue des poids à la fréquence d’apparition de chaque motif d’ordre (OP) en fonction de la variance des valeurs des échantillons qui le composent. En revanche, l’AAPE introduit un paramètre de réglage pour ajuster l’importance de la valeur moyenne de l’amplitude par rapport à sa variabilité. La DispEn commence par construire un signal artificiel en associant les échantillons du signal d’origine à un ensemble de classes prédéfinies

selon une distribution de probabilité présélectionnée. Ensuite, elle applique la PE à ce signal artificiel.

Dans le présent article, nous proposons une nouvelle variante de la PE qui capture efficacement à la fois les signatures spectrales des OPs [12] et la variation d’amplitude. Dans cette nouvelle PE, les poids attribués à la fréquence d’apparition des OPs sont calculés en se basant sur les déterminants des matrices circulantes. Ces matrices sont construites pour chaque segment du signal correspondant à un OP donné. La robustesse, la sensibilité aux variations des fréquences et de l’amplitude, ainsi que l’intérêt de cette nouvelle PE par rapport à celles existantes, sont démontrés au travers de l’analyse de plusieurs signaux académiques.

L’article est organisé comme suit. La section 2 passe en revue le concept de PE et les PE existantes sensibles à l’amplitude. La section 3 présente la PE basée sur les déterminants circulants (CDPE). Dans la section 4, une étude comparative de ces variantes est réalisée à l’aide de signaux académiques. Enfin, la section 5 conclut l’article et mentionne les directions futures de la recherche.

## 2 PE et PEs sensibles à l’amplitude

Cette section donne un aperçu rapide de la PE et ses variantes qui intègrent l’information relative à l’amplitude.

## 2.1 Entropie de permutation - PE

Une représentation symbolique sans seuil d'une série temporelle faiblement stationnaire et échantillonnée uniformément  $x_t$  pour  $t = 0, 1, \dots, N-1$  peut être obtenue en classant les valeurs au sein de chaque ensemble de  $d$  échantillons consécutifs,  $x_t, x_{t+1}, \dots, x_{t+d-1}$ . La permutation de l'ensemble  $\{0, 1, 2, \dots, d-1\}$  qui trie ces valeurs d'échantillons dans l'ordre croissant est appelée un motif ordinal (OP). Par exemple, il y a deux OPs possibles pour  $d=2$  :  $\Pi_1='01'$  si  $x_t < x_{t+1}$ , et  $\Pi_2='10'$  dans le cas contraire. La fréquence d'apparition de chaque OP, noté  $\Pi$ , peut être estimée comme suit :

$$p_{\Pi} = \frac{\#\{t \mid \{x_t, x_{t+1}, \dots, x_{t+d-1}\} \text{ de type } \Pi\}}{N-d+1}, \quad (1)$$

où  $\#$  désigne la cardinalité. La PE normalisée est alors définie par [2]

$$H = -(\log(d!))^{-1} \sum_i p_{\Pi_i} \log(p_{\Pi_i}), \quad (2)$$

où  $d!$  représente le nombre total des OPs possibles de longueur  $d$ , en supposant que les valeurs égales des échantillons sont rares. Notez que pour du bruit gaussien non corrélé,  $p_{\Pi} = \frac{1}{d!}$ , ce qui implique que la PE dans (2) est égale à 1. Rappelons aussi que la PE dans (2) présente deux limitations clés : l'ambiguïté dans la gestion des valeurs égales et l'omission de la variation d'amplitude.

## 2.2 PE pondérée - ImpPE

Contrairement à la PE, qui ne prend en compte que l'aspect ordinal des données, la PE pondérée (ImpPE) intègre des poids dans la probabilité des OPs (1), ce qui permet de distinguer les fluctuations mineures d'amplitude (potentiellement causées par le bruit) des variations significatives [3]. Par conséquent, la fréquence d'apparition d'un OP (1) est modifiée comme suit :

$$p_{\Pi} = \frac{\sum_t |\{x_t, x_{t+1}, \dots, x_{t+d-1}\} \text{ de type } \Pi| w_t}{\sum_{t=1}^{N-d+1} w_t}, \quad (3)$$

$$w_t = \frac{1}{d-1} \sum_{i=0}^{d-1} (x_{t+i} - \bar{x}_t)^2, \quad \text{et } \bar{x}_t = \frac{1}{d} \sum_{i=0}^{d-1} x_{t+i}. \quad (4)$$

Notez que pour de petites valeurs de  $d$ , (4) adopte l'estimateur sans biais de la variance, contrairement à la formulation originale [3].

## 2.3 PE Sensible à l'Amplitude - AAPE

L'AAPE diffère de la ImpPE en intégrant à la fois la moyenne absolue et les différences d'amplitude dans le calcul du poids pour définir la probabilité d'un OP. Par

conséquent, le poids dans (3) devient [9] :

$$w_t = A \frac{1}{d} \sum_{i=0}^{d-1} |x_{t+i}| + (1-A) \frac{1}{d-1} \sum_{i=1}^{d-1} |x_{t+i} - x_{t+i-1}|, \quad (5)$$

où  $A \in [0, 1]$  est un paramètre de réglage qui équilibre l'influence de la moyenne locale et des fluctuations.

## 2.4 Entropie de dispersion - DispEn

La DispEn est une PE qui commence par attribuer des labels à la série temporelle originale  $\{x_t\}_{t=0,1,\dots,N-1}$  à l'aide de  $c$  classes afin de former une série temporelle symbolique  $\{u_t\}_{t=0,1,\dots,N-1}$ , où  $u_t \in [1, c]$ . Cela peut être réalisé de manière linéaire, en triant toutes les valeurs des échantillons et en les divisant en classes de taille égale, ou de manière non linéaire, en utilisant des fonctions sigmoïdes ou la distribution cumulative normale. Chaque sous-séquence  $\{u_t, u_{t+1}, \dots, u_{t+d-1}\}$  définit un motif de dispersion. La fréquence d'apparition d'un motif donné est obtenue à partir de (1) en remplaçant  $x_t$  par la série temporelle symbolique  $u_t$ . Enfin, la DispEn normalisée est donnée par :

$$H = -\left(\log(c^d)\right)^{-1} \sum_i p_{\Pi_i} \log(p_{\Pi_i}), \quad (6)$$

où  $c^d$  représente le nombre total de motifs de dispersion possibles.

## 3 PE proposé - CDPE

Pour prendre en compte à la fois l'information d'amplitude et les caractéristiques spectrales de chaque sous-séquence  $\{x_t, x_{t+1}, \dots, x_{t+d-1}\}$ , nous proposons une nouvelle variante de la PE où les poids dans (3) sont redéfinis comme étant le déterminant d'une matrice circulante construite à partir de cette sous-séquence :

$$w_t = |\det(X_t)| \quad \text{où} \quad (7)$$

$$X_t = \begin{bmatrix} x_t & x_{t+1} & x_{t+2} & \cdots & x_{t+d-1} \\ x_{t+d-1} & x_t & x_{t+1} & \cdots & x_{t+d-2} \\ x_{t+d-2} & x_{t+d-1} & x_t & \cdots & x_{t+d-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{t+1} & x_{t+2} & x_{t+3} & \cdots & x_t \end{bmatrix}. \quad (8)$$

Le déterminant de cette matrice circulante (8) est directement lié à la transformée de Fourier discrète (TFD) de son vecteur générateur [13] et peut s'exprimer comme :

$$w_t = \prod_{k=0}^{d-1} |\lambda_k|, \quad \text{avec les valeurs propres } \lambda_k = \sum_{i=0}^{d-1} x_{t+i} e^{-j \frac{2\pi i k}{d}}, \quad (9)$$

et  $j^2 = -1$ . Ce déterminant circulant, et par conséquent le poids attribué à chaque OP en utilisant (7) en combinaison avec (3), capture les caractéristiques spectrales des

sous-séquences associées aux OPs. Notez qu'en présence de motifs structurés, une redondance ou une périodicité dans les données, ce déterminant circulant sera quasi-nul.

Pour éviter les cas triviaux où  $\lambda_0 = \sum_{i=0}^{d-1} x_{t+i} = 0$ , on peut calculer :

$$w_t = |\det(X_t - \alpha \mathbf{1})| = |\lambda_0 - \alpha d| \prod_{k=1}^{d-1} |\lambda_k|, \quad (10)$$

où  $\mathbf{1}$  est une matrice  $d \times d$  remplie de 1, et  $\alpha$  est une constante, par exemple,  $\alpha = \min(\{x_t\}_{t=0,1,\dots,N-1})$ . Dans la suite, les performances de la CDPE sont examinées au moyen de signaux académiques.

## 4 Résultats

Dans cette section, la PE, la ImpPE, l'AAPE, la DispEn et notre méthode CDPE sont comparées sur des signaux sinusoïdaux et du bruit gaussien.

### 4.1 CDPE pour les signaux sinusoïdaux

Les figures 1.a et c illustrent les performances des variantes considérées de la PE sur un signal sinusoïdal sans bruit  $x_t = \sin(2\pi\nu t)$ , en fonction de la fréquence normalisée  $\nu$ . Le nombre des échantillons est  $N = \lfloor 500/\nu \rfloor$ . Comme le montrent ces figures, la CDPE surpasse les autres PEs pour  $\nu \in [0.01, 0.2]$ , capturant précisément la PE théorique d'un sinus : 0.3869 pour  $d=3$  et 0.2181 pour  $d=4$ . Cette performance supérieure est confirmée dans les figures 1b–1d, où le signal sinusoïdal est noyé dans un bruit blanc gaussien (BBG) avec un rapport signal-sur-bruit (RSB) de 12 dB. Les résultats, moyennés sur 150 simulations Monte-Carlo, confirment les valeurs plus faibles de la CDPE et sa plus grande robustesse au bruit pour  $\nu \in [0.01, 0.2]$ .

Capter précisément la complexité du signal dans cette gamme de fréquences est essentiel pour des futures applications telles que l'analyse des signaux électromyographiques de surface (sEMG). En effet, ces signaux, qui présentent généralement des niveaux de RSB modérés, couvrent souvent une plage de fréquences de 5 à 500 Hz [14], ce qui correspond à une gamme de fréquences normalisées de  $[0.0024, 0.2441]$  pour une fréquence d'échantillonnage  $F_s = 2048$  Hz.

### 4.2 CDPE pour un bruit blanc gaussien

La CDPE théorique (3) (7) d'un BBG est égale à 1 (voir l'Annexe pour la démonstration avec  $d = 3$ ). Le Tableau 1 présente la moyenne  $\pm$  l'écart-type des PEs considérées, moyennées sur 150 simulations de Monte Carlo avec  $N = 5000$  échantillons. Pour  $d = 3$  à 6, la PE reste la plus proche de la valeur théorique. Bien que la CDPE présente l'écart-type le plus élevé pour  $d \geq 4$ , elle se comporte de

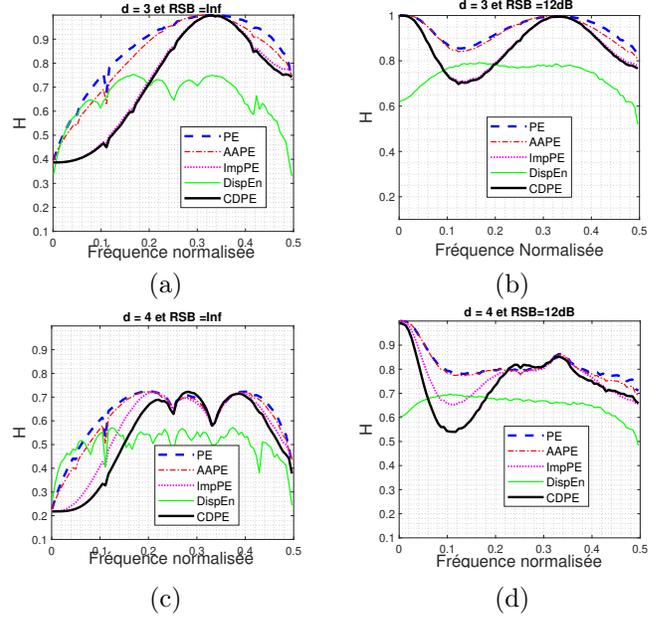


FIGURE 1 – Comparaison des PEs sensibles à l'amplitude appliquées à : (a) et (c) un signal sinusoïdal pur avec  $d=3$  et 4, respectivement. (b)-(d) un signal sinusoïdal noyé dans un BBG à un RSB=12 dB avec  $d=3$  et 4. Les résultats sont moyennés sur 150 réalisations Monte-Carlo. L'AAPE et DispEn sont calculées avec  $A=0.5$  et  $c = d$ , respectivement.

manière similaire à l'AAPE et à la ImpPE. En revanche, la DispEn se détériore pour  $c = d > 4$ .

## 5 Conclusions

Nous avons présenté une nouvelle variante de l'entropie de permutation basée sur le déterminant de matrices circulantes (CDPE). Cette approche innovante intègre les variations d'amplitude et de fréquence en exploitant les propriétés spectrales de ce type de déterminant. Appliquée à des signaux purement sinusoïdaux dans la gamme de fréquences réduite  $[0.01, 0.2]$ , la CDPE est la variante de PE qui se rapproche le mieux de la valeur théorique attendue pour ce type de signaux. Ce résultat met en évidence la pertinence de la CDPE pour des applications futures, telles que l'analyse des signaux électromyographiques, notamment ceux ayant un contenu spectral dans cette gamme de fréquences.

## Annexe - CDPE d'un BBG

Considérons un bruit blanc gaussien centré (BBGC)  $x_t$  avec une variance unitaire, et notons  $\{r_0, r_1, r_2\}$  la séquence triée de  $\{x_t, x_{t+1}, x_{t+2}\}$ . Pour évaluer l'espérance du poids dans (10) pour  $d = 3$ , le déterminant  $\det(X_t - \alpha \mathbf{1})$

TABLE 1 – Moyenne  $\pm$  Ecart type des mesures PEs pour du BBG :  $N = 5000$ , 150 MC réalisations.

ADPE	$d = 3$		$d = 4$		$d = 5$		$d = 6$	
	Moyenne	Écart-type	Moyenne	Écart-type	Moyenne	Écart-type	Moyenne	Écart-type
PE	0.9998	0.0002	0.9994	0.0002	0.9976	0.0004	0.9888	0.0007
CDPE	0.9997	0.0002	0.9893	0.0023	0.9786	0.0027	0.9366	0.0056
ImpPE	0.9997	0.0002	0.9989	0.0004	0.9964	0.0005	0.9846	0.0009
AAPE	0.9993	0.0003	0.9990	0.0003	0.9973	0.0005	0.9885	0.0007
DispEn (c=3)	0.9993	0.0002	0.9983	0.0004	0.9956	0.0006	0.9370	0.0052
DispEn (c=d)	0.9993	0.0002	0.9955	0.0005	0.9553	0.0014	0.7854	0.0005

s'exprime comme suit :

$$\det(X_t - \alpha \mathbf{1}) = r_0^3 + r_1^3 + r_2^3 - 3r_0r_1r_2 - 3\alpha(r_0^2 + r_1^2 + r_2^2 - r_0r_1 - r_0r_2 - r_1r_2). \quad (11)$$

Étant donné la densité de probabilité conjointe de  $\{r_i\}_{i=0,1,2}$  :

$$f_{r_0, r_1, r_2}(u, v, w) = 3!p(u)p(v)p(w)\mathbb{1}_{(u \leq v \leq w)}, \quad (12)$$

où  $p(u)$  est la densité de probabilité de la loi normale centrée réduite, les valeurs d'espérance suivantes peuvent être déduites [12] :

$$\begin{aligned} E[r_2] &= -E[r_0] = \frac{3}{2\sqrt{\pi}}, & E[r_1] &= 0, \\ E[r_1^2] &= 1 - \frac{\sqrt{3}}{\pi}, & E[r_0^2] &= E[r_2^2] = 1 + \frac{\sqrt{3}}{2\pi}, \\ E[r_0r_2] &= -\frac{\sqrt{3}}{\pi}, & E[r_1r_2] &= E[r_0r_1] = \frac{\sqrt{3}}{2\pi}, \\ E[r_1^3] &= 0, & E[r_2^3] &= -E[r_0^3] = \frac{15}{4\sqrt{\pi}}, \\ E[r_0r_1r_2] &= 0. \end{aligned} \quad (13)$$

Pour une valeur appropriée de  $\alpha$  garantissant  $\det(X_t - \alpha \mathbf{1}) > 0$  dans (11), l'espérance du poids associé à chaque motif ordinal est  $9|\alpha|$ . Ce résultat conduit à une distribution uniforme des motifs avec une fréquence d'occurrence  $p_i = \frac{1}{6}$  et une CDPE normalisée égale à 1.

## Références

- [1] M. Jabloun, O. Buttelli, and P. Ravier, "Legendre polynomial fitting-based permutation entropy offers new insights into the influence of fatigue on surface electromyography signal complexity," *Entropy*, vol. 26, no. 10, 2024.
- [2] C. Bandt and B. Pompe, "Permutation entropy : A natural complexity measure for time series," *Phys. Rev. Lett.*, vol. 88, p. 174102, Apr 2002.
- [3] B. Fadlallah, B. Chen, A. Keil, and J. Principe, "Weighted-permutation entropy : A complexity measure for time series incorporating amplitude information," *Phys. Rev. E.*, vol. 87, no. 2, p. 022911, 2013.
- [4] J. Xia, P. Shang, J. Wang, and W. Shi, "Permutation and weighted-permutation entropy analysis for the complexity of nonlinear time series," *Commun. Nonlinear Sci. Numer. Simul.*, vol. 31, no. 1, pp. 60–68, 2016.
- [5] M. El Sayed Hussein Jomaa, P. Van Bogaert, N. Jrad, N. E. Kadish, N. Japaridze, M. Siniatchkin, M. A. Colominas, and A. Humeau-Heurtier, "Multivariate improved weighted multiscale permutation entropy and its application on EEG data," *Biomed. Sig. Process. Control*, vol. 52, pp. 420–428, 2019.
- [6] D. Stosic, D. Stosic, T. Stosic, and B. Stosic, "Generalized weighted permutation entropy," *Chaos : An Interdisciplinary Journal of Nonlinear Science*, vol. 32, no. 10, p. 103105, 10 2022.
- [7] D. Bai, W. Yao, S. Wang, and J. Wang, "Multiscale weighted permutation entropy analysis of schizophrenia magnetoencephalograms," *Entropy*, vol. 24, no. 3, 2022.
- [8] A. S. Gaudêncio, M. Hilal, J. M. Cardoso, A. Humeau-Heurtier, and P. G. Vaz, "Texture analysis using two-dimensional permutation entropy and amplitude-aware permutation entropy," *Pattern Recognition Letters*, vol. 159, pp. 150–156, 2022.
- [9] H. Azami and J. Escudero, "Amplitude-aware permutation entropy : Illustration in spike detection and signal segmentation," *Computer Methods and Programs in Biomedicine*, vol. 128, pp. 40–51, 2016.
- [10] M. Rostaghi and H. Azami, "Dispersion entropy : A measure for time-series analysis," *IEEE Sig. Proc. Lett.*, vol. 23, no. 5, pp. 610–614, 2016.
- [11] F. Nieto-del Amor, R. Beskhani, Y. Ye-Lin, J. Garcia-Casado, A. Diaz-Martinez, R. Monfort-Ortiz, V. J. Diago-Almela, D. Hao, and G. Prats-Boluda, "Assessment of dispersion and bubble entropy measures for enhancing preterm birth prediction based on electrohysterographic signals," *Sensors*, vol. 21, no. 18, 2021.
- [12] M. Jabloun, "Éxploration théorique des motifs d'ordre pour une meilleure compréhension de leur utilité dans les méthodes ordinales de traitement du signal," in *GRETSI 2023*, Grenoble, France, Aug. 2023.
- [13] R. M. Gray, "Toeplitz and circulant matrices : A review," *Found. Trends Commun. Inf. Theory*, vol. 2, no. 3, pp. 155–239, 2006.
- [14] S. Muceli and R. Merletti, "Tutorial. frequency analysis of the surface EMG signal : Best practices," *J. Electromyogr. Kinesiol.*, vol. 79, p. 102937, 2024.