

Unakrsna entropija uzoraka transformisanih kardiovaskularnih podataka: pouzdanost procene

Dragana Bajić, Tamara Škorić

Univerzitet u Novom Sadu, Departman za energetiku,
elektroniku i telekomunikacije
Fakultet tehničkih nauka
Novi Sad, Srbija
tamara.ceranic@gmail.com, dragana.bajic@gmail.com

Nina Japundžić-Žigon

Medicinski fakultet,
Univerzitet u Beogradu,
Beograd, Srbija
nzigon@med.bg.ac.rs

Sažetak—Unakrsna entropija uzoraka (Cross-Sample Entropy, *XSampEn*) kvantifikuje međusobnu predvidljivost dva signala snimljenih iz srodnih izvora. Cilj rada je da redukuje nepouzdanost procene *XSampEn* signala koja nastaje usled nesaglasnih raspodela. Redukcija je ostvarena transformacijom posmatranih nizova putem funkcija raspodele (probability integral transform, PIT), čime se dobija uniformna raspodela oba signala. Predložena metoda je validirana na signalima sistolonog krvnog pritiska i pulsog interval snimljenih na normatenzivnim i granično hipertenzivnim laboratorijskim životinjama podvrgnutim različitim vrstama stresnih situacija.

Ključne reči – unakrsna entropija uzoraka; transformacija slučajne promenljive; pouzdanost; (key words)

I. UVOD

Merenje neizvesnosti vremenskih nizova oduvek je bio intrigantan zadatak. Entropija kao alat za procenu neizvesnost komunikacionih signala uvedena je u klasičnom Šenonovom radu [1], nakon kojeg su usledili i drugačiji koncepti, kao što su Kolmogorov-Sinai [2], Grassberger i Procaccia [3] i Eckmann i Ruelle [4]. Pincus [5] je predložio aproksimativnu entropiju (*ApEn*) kojom su izbegnuti stroge matematičke formulacije i omogućena primena na biomedicinskim signalima. *ApEn* nema teorijsku povezanost sa strogim matematičkim konceptima entropije (zato se i naziva aproksimativna), ali je njena implementacija jednostavna. Veoma je zastupljena u pretkliničkim istraživanjima, na primer: analiza varijabiliteta srčanog ritma pacijenata sa dva tipa dijabetesa [6], srčanom insuficijencijom [7], zdravih pacijenata [8] u toku vežbanja i odmora [9]. Međutim, *ApEn* je parametarska metoda pa se ubrzo pokazalo da je nekonzistentna jer eksperimentalni rezultati zavise od parametara [10]-[14]. Stoga su Richmann i Moorman predložili korekciju, entropiju uzorka (*Sample Entropy - SampEn*) [15]. To je najzastupljenija korekcija aproksimativne entropije, a trend rasta citiranosti *SampEn* prati rast citiranosti *ApEn* [14]. Jednostavne modifikacije *ApEn* kao što su zamene redosleda operacija sumiranje i logaritma uspešni su da ublaže nekonzistentnost u proceni [15] ali ne i da je u potpunosti uklone [13]. Unakrsna *SampEn* (*Cross Sample Entropy - XSampEn*) je predložena kao alternativa unakrsnoj aproksimativnoj entropiji *XApEn* (*Cross Approximate Entropy*) sa namerom da se pouzdano proceni međusobnu predvidljivost

dva signala [15]. *XSampEn* bi trebala da oslika uticaj jednog signala na drugi, bez obzira na statistička svojstva signala. Međutim, ovaj zahtev nije uvek lako ispuniti, posebno ako jedan od signala ima izrazito asimetričnu raspodelu.

Metoda koja uspešno otklanja statističke osobine vremenskih nizova je integralna transformacija verovatnoće (*Probability Integral Transform - PIT*) [16]. *PIT* transformiše slučajnu promenljivu proizvoljne raspodele u slučajnu promenljivu uniformne raspodele. Ova transformacija menja raspodelu, izjednačavajući verovatnoće amplituda uzoraka ali očuvavajući statističku zavisnost između odmeraka, kako u pojedinačnom nizu tako i u oba vremenska niza.

U ovom radu ćemo kombinovati *XSampEn* i *PIT* kako bismo izbegli distorziju *XSampEn* prouzrokovanu procenom na nepouzdanim i nesaglasnim raspodelama. Pokazano je da procena na *PIT* signalima (*PIT-XSampEn*) vodi više konzistentnim rezultatima u smislu ispunjenosti kriterijuma za pouzdanost. Predložena metoda je validirana na sistolnom krvnom pritisku i impulsnim intervalima snimljenim na normatenzivnim i graničnim hipertenzivnim laboratorijskim pacovima izloženim dvema vrstama stresa.

II. MATERIJALI I METODE

A. Eksperimentalni podaci

Za testiranje *PIT-XSampEn* procene koristili smo kardiovaskularne signale snimljene na mužjacima Vistar normatenzivnih pacova (NRM) i granično hipertenzivnih pacovima (BHR). Za snimanje talasnog oblika krvnog pritiska i prenos podataka koristi se telemetrijska sonda (TA11PA-C40, DSI, Transoma Medical) implantirana u abdominalnu aortu pacova 10 dana pre eksperimenta. Vremenski nizovi koji su izdvojeni iz talasnog oblika krvnog pritiska (BP) su sistolni krvni pritisak (SBP) – lokalni maksimumi BP i pulsni interval (PI) - interval između susjednih lokalnih maksimalnih rastućih gradijenata $(\Delta(BP)/\Delta T)_{max}$, tj. kao maksimalna promena $(\Delta)BP$ po intervalu odmeravanja $\Delta T=1ms$.

Druga eksperimentalna komponenta bila je izlaganje stresu: pacovi su bili postavljeni na platformu koja se tresu 200 puta/minuti (stres trešenja) ili postavljeni u pleksiglas cev

unutrašnjeg prečnika 5.5 cm sa porama u ležećem položaju (stres ograničavanja). Ovi signali su već korišćeni u drugim istraživanjima [13], [17]. Detaljniji opis protokola eksperimenta i prikupljenih signala dat je u [18].

Signali su detrendizovani pomoću filtra propusnika visokih učestanosti dizajniranog za biomedicinske signale [19], kako bi se osigurala stacionarnost signala. Analizirali smo četiri grupe podataka: 6 NRM pacova snimanih u osnovnim uslovima i u toku izlaganja prvom stresu ograničavanja, 6 NRM pacova snimanih u osnovnim uslovima i u toku izlaganja prvom stresu trešenja, 6 BHR pacova snimanih u osnovnim uslovima i u toku izlaganja prvom stresu ograničavanja, i 6 BHR snimanih u osnovnim uslovima i u toku izlaganja prvom stresu trešenja. Srednje vrednosti sistolnog krvnog pritiska i impulsnog intervala za sve grupe ekperimentalnih podataka prikazane su u Tabeli 1. Sve eksperimentalne procedure u ovim studijama su izvođene u skladu sa direktivom Saveta Evropske zajednice od 24.11.1986. god. (86/609/ECC) i Upustvom za eksperimente nad životinjama Medicinskog fakulteta u Beogradu.

B. Metode

XSampEn [15] meri međusobnu asinhronost dva vremenska niza: master vremenski niz je $x_i \in X$, $i=1, \dots, N$ a sledbenik je $y_j \in Y$, $j=1, \dots, N$, gde je N dužina niza. Vremenski nizovi moraju biti normalizovani i centralizovani da bi njihovo poređenje uopšte bilo moguće. U ovom istraživanju koristili smo SBP i PI signale.

U prvom koraku nizovi se dele na preklapajuće vektore dužine m (m je obično 2, 3 ili 4):

- master vektor

$$\mathbf{X}_m^{(i)} = [x_i, x_{i+1}, \dots, x_{i+m-1}], \quad i = 1, \dots, N-m + 1;$$

- sledbenik vektor

$$\mathbf{Y}_m^{(j)} = [y_j, y_{j+1}, \dots, y_{j+m-1}], \quad j = 1, \dots, N-m + 1.$$

U drugom koraku računa se rastojanje između dva vektora. Rastojanje se računa kao maksimalno apsolutno rastojanje:

$$d(\mathbf{X}_m^{(i)}, \mathbf{Y}_m^{(j)}) = \max_{k=0, \dots, m-1} |x_{i+k} - y_{j+k}|, \quad i, j = 1, \dots, N-m + 1$$

U trećem koraku, rastojajne se poredi sa unapred definisanim pragom r . Ako je rastojanje manje ili jednako od r , vektori su "slični".

$$\hat{p}_i^{(m)}(r) = \Pr\{d(\mathbf{X}_m, \mathbf{Y}_m) \leq r | \mathbf{X}_m = \mathbf{X}_m^{(i)}, \mathbf{Y}_m \in \mathbf{Y}, r > 0\} = \frac{1}{N-m} \sum_{j=1}^{N-m} \mathbb{I}\{d(\mathbf{X}_m^{(i)}, \mathbf{Y}_m^{(j)}) \leq r\}. \quad (1)$$

U (1) “ \wedge ” označava procenu, dok je $\mathbb{I}\{\cdot\}$ indikatorska funkcija koja opisuje proceduru prebrojavanja jer je jednaka jedinici ako je uslov koji indicira, tj. $d(\mathbf{X}_m^{(i)}, \mathbf{Y}_m^{(j)}) \leq r$, ispunjen, u suprotnom jednaka je nuli. Izraz (1) predstavlja relativnu učestanost vektora sličnih master vektoru $\mathbf{X}_m^{(i)}$.

Sabirak $\hat{\Phi}^{(m)}$ formira se kao logaritam suma verovatnoća procenjenih u (1) što ujedno predstavlja i najveću izmenu u poređenju sa *XApEn*:

$$\hat{\Phi}^{(m)}(N, r) = \ln \left(\sum_{i=1}^{N-m} \hat{p}_i^{(m)}(r) \right) \quad (2)$$

Postupak se ponavlja za vektore dužine $m+1$. Nakon procene sabirka $\hat{\Phi}^{(m+1)}$, unakrsna entropija uzorka procenjuje se kao:

$$XSampEn = \hat{\Phi}^{(m)}(r, N) - \hat{\Phi}^{(m+1)}(r, N). \quad (3)$$

C. Integralna transformacija verovatnoće

Ova transformacija konvertuje slučajnu promenljivu u promenjuju uniformne raspodele u opsegu od [0, 1]. Kada se konvertuju dva signala, oba dobijaju uniformnu raspodelu ali vremenski niz odmeraka ostaje isti, tako da se i međusobna zavisnost održava. Za transformaciju se koristi kumulativna funkcija raspodele slučajne promenljive r , $t = F_r(r)$

$$F_r(t) = \Pr\{t \leq t_0\} = \Pr\{F_r(r) \leq t_0\} = \Pr\{r \leq F_r^{-1}(t_0)\} = F_r(F_r^{-1}(t_0)) = t_0. \quad (4)$$

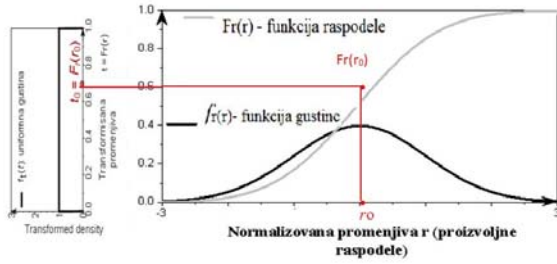
Na Sl. 1. ilustrativno je prikazana transformacija slučajne promenljive r funkcije gustine raspodele $f_r(r)$ u slučajnu promenljivu $t = F_r(r)$ sa uniformnom funkcijom gustine raspodele $f_t(t)$.

PIT kao transformacija se najviše koristila na početku razvoja simulacionih sistema, jer je pomoću nje bilo moguće generisati bilo koju drugu raspodelu ukoliko postoji generator uniformnih brojeva u opsegu od 0 do 1. Ta njena korisna osobina je u međuvremenu zaboravljena jer svaki softverski alat poseduje ugrađene generatore bezmalo svih poznatih

TABELA I. SISTOLNI KRVNI PRITISAK I PULSNI INTERVAL U OSNOVNIM USLOVIMA I USLOVIMA STRESA

STRES	RASA	SBP [mmHg]		PI [ms]	
		OSNOVNI	STRES	OSNOVNI	STRES
TREŠENJA	NRM	116.28±6.82	125.55±8.89	167.8±13.95	157.97±10.98
TREŠENJA	BHR	134.87±9.75	149.54±6.97	186.56±22.64	175.82±17.39
OGRANIČAVANJA	NRM	107.08±7.83	114.37±24.92	179.29±13.66	131.46±5.71
OGRANIČAVANJA	BHR	133.26±13.5	148.89±15.55	182.27±10.87	151.71±16.67

Rezultati su prikazani kao srednja vrednost ± standardna devijacija.



Slika 1. Transformacija gustine verovatnoće

raspodela.

D. Pouzdana procena verovatnoća-Jeruhimovi kriterijumi

Verovatnoće (1) predstavljaju osnovu *XSampEn* metode, pa nepouzdanost u njihovoj proceni utiče na krajnju nekonzistentnost u proceni entropije. Pošto verovatnoće odgovaraju binarnom događaju - master i sledbenik mogu biti ili slični ili različiti - njihova pouzdanost može da se proveri Jeruhimovim kriterijumima [20]. Kriterijumi definišu minimalnu dužinu signala sa kojom bi se postigla pouzdanost u proceni (interval pouzdanosti 95%). Teorijska analiza izvedena u [20] potvrdila je tradicionalno inženjersko pravilo da dužina signala potrebna za pouzdanu procenu binarnih događaja treba da bude od $10/p_i^m(r)$ do $100/p_i^m(r)$ [20]. Ako su raspodele signala X i Y nesaglasne, verovatnoća pojavljivanja sličnih segmenata neće biti procenjene na dovoljnom broju sličnih uzoraka. Nepouzdana procena verovatnoća, uticaće na nepouzdanost u proceni *XSampEn*. Ako su amplitude raspodele uniformno raspoređene, što nam omogućava PIT transformacija, verovatnoća ispunjenosti Jeruhimovih kriterijuma raste i procena entropije je više pouzdana.

Na Sl. 2 su posebno označeni optimalni pragovi, dobijeni metodom izvedenom u [17]:

$$r(m) = 0.15 \cdot (m + 1) + r_T(m) \quad [17] \quad (5)$$

$$r_T(m) = r_A(m) + \begin{cases} -0.02 + \frac{0.023 \cdot \sqrt{(\sigma_{DX} + \sigma_{DY})/2}}{\sqrt[4]{N/1000}}, & m=2 \\ -0.006 + \frac{0.043 \cdot \sqrt{(\sigma_{DX} + \sigma_{DY})/2}}{\sqrt[4]{N/1000}}, & m=3 \end{cases} \quad [17] \quad (6)$$

$$r_A(m) = \begin{cases} \frac{-0.02 + 0.23 \cdot \sqrt{(\sigma_{DX} / \sigma_Y)}}{\sqrt[4]{N/1000}}, & m=2 \\ \frac{-0.06 + 0.43 \cdot \sqrt{(\sigma_{DX} / \sigma_Y)}}{\sqrt[4]{N/1000}}, & m=3 \end{cases} \quad [10] \quad (7)$$

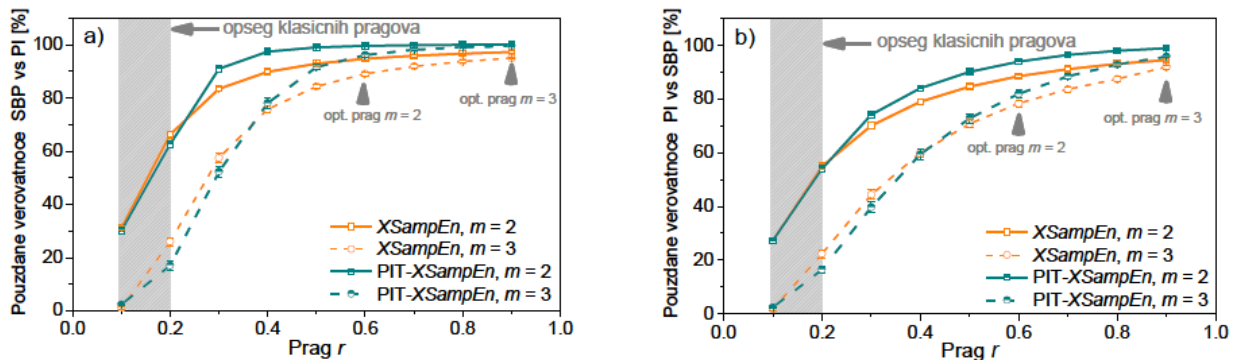
U izrazima (5), (6) i (7) σ_X je standardna devijacija master niza \mathbf{X} (u kros-entropijskim procenama je uvek jednaka jedinici) a σ_{DX} i σ_{DY} su procenjene standardne devijacije diferencijalnih nizova $x(i)-x(i-1)$ i $y(i)-y(i-1)$.

III. REZULTATI I DISKUSIJA

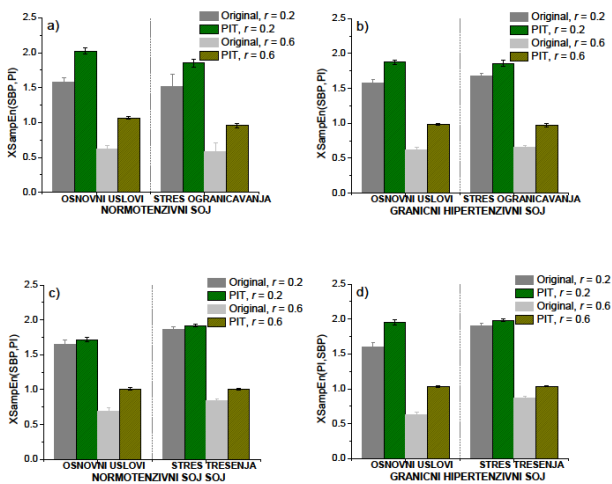
Analizirali smo pouzdanost procenjenih verovatnoća $\hat{p}_i^m(r)$ (1) za niz pragova r - od 0.1 do 0.9 - i za dužine vektora $m=2$ i $m=3$. Na Sl. 2a narandžastom linijom prikazan je procenat pouzdanost procenjenih verovatnoća za *XSampEn*(SBP,PI), gde je SBP odabran kao master vremenski niz a PI kao njegov sledbenik. Zelenom linijom prikazan je procenat pouzdanost procenjenih verovatnoća za SBP i PI signal nakon PIT transformacije. Na Sl. 2b prikazani su rezultati istog eksperimenta, ali PI je odabran kao master vremenski niz. Sivom bojom označen je spektar pragova r koji se uobičajeno preporučuje u literaturi [15].

Sl. 3 predstavlja procenjene vrednosti *XSampEn* za uobičajen prag $r=0.2$ i za vrednost $r=0.6$ koja je optimalna za dužine vektora $m=2$ i $N=3000$ [17], za sve četiri grupe eksperimentalnih podataka.

Nepouzdanost u proceni uslovnih verovatnoća $\hat{p}_i^m(r)$ predstavlja jedan od osnovnih uzroka nekonzistentnosti u proceni *XSampEn*, što je potvrđeno i za *XApEn* u [17]. Na Sl. 2 je pokazano da pouzdanost u proceni uslovnih verovatnoća raste sa porastom kriterijuma sličnosti r i sa smanjenjem dužine



Slika 2. Procenat pouzdanost procenjenih $\hat{p}_i^m(r)$ a) SBP-master, PI-sledbenik b) SBP -sledbenik, PI-master. Korišćeni signali dužine $N=3000$.



Slika 3. $XSampEn(SBP,PI)$ procenien na originalnim SBP-PI parovima signala i na PIT transformisanim parovima signala pri različitim tipovima stresa. (a) NRM pacovi u stresu ograničavanja; (b) pacovi u stresu ograničavanja; (c) NRM pacovi u stresu trešenja; (d) NRM pacovi u stresu trešenja. Sva četiri panela prikazuju i procenu entropije u uslovima bez stresa, a dužine nizova su $N=3000$.

segmenta m , tj. raste sa povećanjem broja sličnih segmenata na osnovu kojih se procenjuju $\hat{p}_i^m(r)$. Takođe se uočava da se konzistentnost procene verovatnoća povećava primenom PIT na oba signala. Budući da se iste verovatnoće koriste i za procenu $XApEn$, i za procenu $XSampEn$, konzistentnost njihove procene se povećava i sa porastom praga, čime su potvrđeni i rezultati iz [17]. Sl. 3 pokazuje da su promene u vrednostima entropije $XSampEn$ i $PIT-XSampEn$ izraženije za životinje izložene stresu trešenja što je u skladu sa medicinskim razmatranjima datim u [18] a takođe se uočava da su rezultati procene entropije nad originalnim i nad transformisanim signalima uporedivi.

Jedna od prednosti PIT transformacije je mogućnost konzistentne procene entropije kada su raspodele signala asimetrične i međusobno se na velikim segmentima ne preklapaju. Međutim, gustine raspodela normalizovanih i centralizovanih SBP i PI signala su simetrične i u istom opsegu tako da ova prednost PIT transformacije nije izražena. Njena prednost je izrazitija kada se PIT koristi za transformaciju signala kao što su inter-spajk intervali čija je raspodela bliska eksponencijalnoj.

LITERATURA

- [1] C.E. Shannon, "Communications in the presence of noise," Proc. IRE 1949, vol. 37, pp. 10–21, Jan. 1949.
- [2] D. Gustavo, S. Bernd, Information Dynamics: Foundations and Applications, 1st ed., Springer-Verlag: New York, USA, 2012, pp.101.
- [3] P. Grassberger, I. Procaccia, "Estimation of the Kolmogorov entropy from a chaotic signal," Phys. Rev. A, vol. 28, pp. 2591–2593, May 1983.
- [4] J. P., Eckmann, D. Ruelle, "Ergodic theory of chaos and strange attractors," Rev. Mod. Phys., vol. 57, pp. 617–656, July 1985.
- [5] S. M. Pincus, "Approximate entropy as a measure of system complexity," Proc. Natl Acad. Sci. U.S.A., vol. 88, pp. 2297–2301, Mar. 1991.

- [6] X. Li, S. Yu, H. Chen, C. Lu, K. Zhang, et. al., "Cardiovascular autonomic function analysis using approximate entropy from 24-h heart rate variability and its frequency components in patients with type 2 diabetes," J. Diabetes. Investig., vol. 6, pp. 227–235, Mar. 2015.
- [7] G. Krstacic, D. Gamberger, A. Krstacic, T. Smuc, D. Milicic, "The chaos theory and non-linear dynamics in heart rate variability in patients with heart failure," Proc Comput Cardiol, pp. 957–959, Sep 2008.
- [8] R. J. Storella, H. W. Wood, K. M. Mills, J. K. Kanters, M. V. Højgaard, et. al., "Approximate entropy and point correlation dimension of heart rate variability in healthy subjects," Integr Physiol and Behav Sci, vol. 33, pp. 315–320, Oct. 1998.
- [9] M. P. Tulppo, T. H. Makikallio, T. E. Takala, T. Seppanen, H. V. Huikuri, "Quantitative beat-to-beat analysis of heart rate dynamics during exercise," Am J Physiol Heart Circ Physiol, vol. 271, pp. H244–H252, Jul. 1996.
- [10] S. Lu, X. Chen, J.K. Kanters, I. C. Solomon, and K.H.Chon, "Automatic selection of the threshold value r for approximate entropy," IEEE Trans Biomed Eng, vol.55, pp. 1966–1972, Aug 2008.
- [11] K. Chon, C. Scully, and S. Lu, "Approximate entropy for all signals," IEEE Eng Med Biol Mag, vol. 28, no. 6, pp. 18–23, Nov. 2009.
- [12] P. Castiglioni and M. Di Rienzo, "How the threshold "r" influences approximate entropy analysis of heart-rate variability," Proc Comput Cardiol, pp.561–564, Oct. 2008.
- [13] A. Boskovic, T. Loncar-Turukalo, O. Sarenac, N. Japundzic Zigon, and D. Bajic, "Unbiased entropy estimates in stress: A parameter study," Comput. Biol. Med., vol.42, pp. 667–679, Jun. 2012.
- [14] J. M. Yentes, N. Hunt, K. K. Schmid, J. P. Kaipust, D. McGrath, et al., "The appropriate use of approximate entropy and sample entropy with short data sets," Ann Biomed Eng, vol.41, pp.349–365, Feb. 2013.
- [15] J. S. Richman and J. R. Moorman, "Physiological time-series analysis using approximate entropy and sample entropy," Am J Physiol-Heart Circ Physiol, vol. 278, pp. H2039–H2049, Jun. 2000.
- [16] J. E. Angus, "The Probability integral transform and related results," SIAM Review, vol. 36, pp. 652–654, 1994.
- [17] T. Skoric, O. Sarenac, B. Milovanovic, N. Japundzic-Zigon, D. Bajic, "On consistency of cross-approximate entropy in cardiovascular and artificial environments," Complexity, pp. 1–15, Aug. 2017.
- [18] O. Sarenac, M. Lozic, S. Drakulic, D. Bajic, J. F. Paton, et al., "Autonomic mechanisms underpinning the stress response in borderline hypertensive rats," Exp Physiol, pp. 1–16, Jun. 2011.
- [19] M.P. Tarvainen, P.O. Ranta-aho, P.A. Karjalainen, "An advanced detrending approach with application to HRV analysis," IEEE Trans. Biomed. Eng., vol. 42, pp. 172–174, Aug. 2002.
- [20] M. C. Jeruchim, "Techniques for estimating the bit error rate in the simulation of digital communication systems," IEEE J. Sel. Areas in Commun., vol. 2, pp. 153–170, Jan. 1984.

ABSTRACT

Cross-Sample Entropy ($XSampEn$) quantifies mutual predictability of two time series, recorded from parallelly recorded related sources. The aim of our paper is to reduce the unreliability of $XSampEn$ estimation caused by possibly mismatching distributions. Reduction is achieved by transformation of observed time series by probability integral transform, PIT, to obtain uniform distribution for both time series. The proposed methods is validated on time series systolic blood pressure and pulse interval recorded from normotensive rats and borderline hypertensive rats exposed to two types of stress.

CROSS-SAMPLE ENTROPY OF TRANSFORMED CARDIOVASCULAR DATA: ESTIMATION RELIABILITY

Dragana Bajić, Tamara Škorić, Nina Japundžić-Žigon