Copula Entropy

Theory and Applications

MA Jian
majian.2003@tsinghua.org.cn

Contents

Lecture 2: Theory of Copula Entropy (2)

Physical Meaning of Copula Entropy

- **Laws of Thermodynamics**
  - First law: Conservation of energy
    Energy cannot be created or destroyed and only changes from one form to another.
    $$\Delta U = Q - W \tag{1}$$
  - Second law: Entropy increase
    The total disorder of an isolated system always goes up over time.
    $$\delta Q = TdS \tag{2}$$
  - Third law: Absolute zero limit
    As temperature approaches absolute zero (0 K), a system's entropy approaches a constant minimum.

Physical Meaning of Copula Entropy

- **Physical meaning of copula entropy**
  entropy reduction due to increasing order of thermodynamic systems

$$H_c(\mathbf{x}) \leq 0 \tag{3}$$

Physical Meaning of Copula Entropy

- **Systems that keep order**
  - Earth
    constantly receiving energy from Sun to keep earth system evolving
  - Society
    receiving energy from natural environment to keep society running
  - Life
    getting food and energy from outside to keep alive

Copula Entropy: Estimation

- **Parametric** Estimation by Definition

$$H_c(\mathbf{x}) = -E(\log c(\mathbf{u}; \alpha)). \tag{4}$$

1. estimating empirical copula density $\mathbf{u}$ based on rank statistic

$$u_i = F_i(x_i) = \frac{1}{T}\sum_{t=1}^{T} \mathbf{1}(x_t^i \leq x_i) \tag{5}$$

2. estimating the parameter α of copula density function $c(\cdot)$ with maximum likelihood method
3. calculating according to (4)

Copula Entropy: Estimation

- **Non-Parametric** Estimation Method[1]
  1. estimating empirical copula density $\mathbf{u}$ based on rank statistic

$$u_i = F_i(x_i) = \frac{1}{T}\sum_{t=1}^{T} \mathbf{1}(x_t^i \leq x_i) \tag{6}$$

  2. estimating copula entropy with the kNN entropy estimation method
- Advantages
  - distribution-free, non-parametric
  - tuning-free, insensitive to parameters
  - good convergence
  - easy to implement
  - low computation burden

[1] Jian Ma and Zengqi Sun. "Mutual Information Is Copula Entropy". In: *Tsinghua Science & Technology* 16.1 (2011). See also arXiv preprint arXiv:0808.0845, 2008, pp. 51–54.

Copula Entropy: Estimation

- **Conditional Mutual Information** Estimation

can be done with copula entropy estimators according to (7).

$$I(x; y|z) = H_c(x, z) + H_c(y, z) - H_c(x, y, z). \quad (7)$$

1. estimating the CE terms in (7) with CE estimator
2. calculating the result

Theory Growth

Classical Information Theory

- CE based MI estimation, such as GCMI[2], MICE[3], Vector Copula based MI estimation[4], Vine copula based MI and CMI estimation[5], semi-parametric MI estimation[6], kernel based CE estimation[7], non-parametric CE estimation[8]

---

[2] Robin A.A. Ince et al. "A statistical framework for neuroimaging data analysis based on mutual information estimated via a gaussian copula". In: *Human Brain Mapping* 38.3 (2017), pp. 1541–1573.

[3] Adrien Cortes. "Accélération des techniques d' analyse de facteurs par méthodes d' apprentissage statistique et application à un GSE". Mémoire. Institut des Sciences Financiéres et d' Assurance, 2023.

[4] Yanzhi Chen et al. "Neural Mutual Information Estimation with Vector Copulas". In: *Advances in Neural Information Processing Systems*. 2025.

[5] 王洪安. "上肢运动的肌间耦合分析方法研究". 硕士学位论文. 杭州电子科技大学, 2021, Houman Safaai. "Amortized Vine Copulas for High-Dimensional Density and Information Estimation". In: *arXiv preprint arXiv:2604.20568* (2026). arXiv: 2604.20568 [cs.LG]. URL: https://arxiv.org/abs/2604.20568.

[6] M. Mohammadi, M. Hashempour, and M. Emadi. "Semiparametric estimation of mutual information for elliptical copulas". In: *The 7th Seminar on Copula Theory and its Applications*. 2023, p. 44.

[7] 林哲緯. "基於 Copula Entropy 與核密度估計的變數選取方法". 硕士学位论文. 国立政治大学, 2025.

[8] Baby A.K. and Rajesh G. "Nonparametric Estimation of Copula Entropy for Left Truncated Right Censored Data". In: *International Journal of Statistics and Reliability Engineering* 13.1 (2026), pp. 10–17.

Theory Growth

Classical Information Theory

- asymmetric MI estimation[9], CE^2[10], Bayesian dependence measure[11], and tail dependence measure[12];

[9] Soumik Purkayastha and Peter X.K. Song. "Asymmetric predictability in causal discovery: an information theoretic approach". In: *arXiv preprint arXiv:2210.14455* (2022).

[10] Lin Liu, Cong Hu, and Xiao-Jun Wu. "CE^2: A Copula Entropic Mutual Information Estimator for Enhancing Adversarial Robustness". In: *Pattern Recognition and Computer Vision*. Springer Singapore, 2024, pp. 163–174.

[11] Guillaume Marrelec and Alain Giron. "An Inferential Measure of Dependence Between Two Systems Using Bayesian Model Comparison". In: *IEEE Transactions on Systems, Man, and Cybernetics: Systems* 55.3 (2025), pp. 1671–1683. DOI: 10.1109/TSMC.2024.3509025.

[12] Omid M. Ardakani and Rawan Ajina. "Tail risks in household finance". In: *Finance Research Letters* 69 (2024), p. 106065. ISSN: 1544-6123. DOI: https://doi.org/10.1016/j.frl.2024.106065. URL: https://www.sciencedirect.com/science/article/pii/S154461232401095X.

Theory Growth

Classical Information Theory

- CE based information decomposition[13];
- CE based maximal entropy[14] and maximal Tsallis CE[15].

[13] Aleksander Wieczorek and Volker Roth. "Causal Compression". In: *arXiv preprint arXiv:1611.00261* (2016), Dinu Johannes Kaufmann. "Semi-parametric Gaussian Copula Models for Machine Learning". PhD thesis. University of Basel, 2017.

[14] 進藤晋. "Copula Entropy の最大化". In: 日本 *OR* 学会 *2011* 年度秋季研究Ⓕ表会. 2011, pp. 362–363, Milan Bubák and Mirko Navara. "Fitting Copulas with Maximal Entropy". In: *Entropy* 27.1 (2025), p. 87. ISSN: 1099-4300. DOI: 10.3390/e27010087. URL: https://www.mdpi.com/1099-4300/27/1/87, Seyedeh Azadeh Fallah Mortezanejad, Gholamreza Mohtashami Borzadaran, and Bahram sadeghpour Gildeh. "Joint dependence distribution of data set using optimizing Tsallis copula entropy". In: *Physica A: Statistical Mechanics and its Applications* 533 (2019), p. 121897. ISSN: 0378-4371. DOI: https://doi.org/10.1016/j.physa.2019.121897. URL: https://www.sciencedirect.com/science/article/pii/S0378437119311161, Seyedeh Azadeh Fallah Mortezanejad et al. "Non-Parametric Multivariate Control Chart Using Copula Entropy". In: *Sankhya B* 87 (2025), pp. 468–518. URL: https://doi.org/10.1007/s13571-025-00374-y.

[15] Seyedeh Azadeh Fallah Mortezanejad, Gholamreza Mohtashami Borzadaran, and Bahram sadeghpour Gildeh. "Joint dependence distribution of data set using optimizing Tsallis copula entropy". In: *Physica A: Statistical Mechanics and its Applications* 533 (2019), p. 121897. ISSN: 0378-4371. DOI: https://doi.org/10.1016/j.physa.2019.121897. URL: https://www.sciencedirect.com/science/article/pii/S0378437119311161.

Theory Growth

Graph Theory

- MI based graph similarity[16].

[16] Francisco Escolano et al. "The mutual information between graphs". In: *Pattern Recognition Letters* 87 (2017), pp. 12–19. DOI: https://doi.org/10.1016/j.patrec.2016.07.012. URL: https://www.sciencedirect.com/science/article/pii/S016786551630174X, Manuel Curado Navarro. "Structural Similarity: Applications to Object Recognition and Clustering". PhD thesis. Universitat d'Alacant, 2018.

Theory Growth

Information Bottleneck

- CE based information bottleneck computation[17].

[17] Mélanie Rey and Volker Roth. "Meta-Gaussian information bottleneck". In: *Advances in Neural Information Processing Systems* 25 (2012), Mélanie Rey. "Copula Models in Machine Learning". PhD thesis. University of Basel, 2015.

Theory Growth

Partial Information Decomposition

- unique information estimation[18], synergy estimation[19], and Ω information estimation[20].

[18] Ari Pakman et al. "Estimating the Unique Information of Continuous Variables in Recurrent Networks". In: *Advances in Neural Information Processing Systems* 34 (2021), pp. 20295–20307.

[19] Vlad-Bogdan Coroian, Pedro Mediano, and Tolga Birdal. *Scaling up synergy estimators*. Tech. rep. Imperial College London, 2024.

[20] Laouen Belloli and Rubén Herzog. *THOI: An efficient library for higher order interactions analysis based on Gaussian copulas enhanced by batch-processing*. GitHub. URL: https://github.com/Laouen/THOI. 2024. URL: https://pypi.org/project/thoi/.

Theory Growth

Information Geometry

- copula variational Bayes algorithm[21].

[21] Robert Zimmerman. *Copulas and Information Geometry: Strange Bedfellows?* URL:https://rob-zimmerman.ca/files/projects/STA4531_Copulas_IG_Strange_Bedfellows.pdf. 2025.

Theory Growth

Copula Theory

- copula function estimation[22], Vine copula model selection[23] and circular-linear copulas[24].

[22] Longxia Qian et al. "A New Estimation Method for Copula Parameters for Multivariate Hydrological Frequency Analysis With Small Sample Sizes". In: *Water Resources Management* 36.4 (2022), pp. 1141–1157. ISSN: 1573-1650. URL: https://doi.org/10.1007/s11269-021-03016-w.

[23] Fadhah Amer Alanazi. "Truncating Regular Vine Copula Based on Mutual Information: An Efficient Parsimonious Model for High-Dimensional Data". In: *Mathematical Problems in Engineering* 2021 (2021), p. 4347957. URL: https://doi.org/10.1155/2021/4347957, Rafael Calsaverini and Renato Vicente. "An information-theoretic approach to statistical dependence: Copula information". In: *EPL (Europhysics Letters)* 88.6 (2009), p. 68003. DOI: 10.1209/0295-5075/88/68003. URL: https://doi.org/10.1209/0295-5075/88/68003, 王念鸽. "基于互信息的 Vine Copula 模型的高频数据投资组合风险测度研究". 硕士学位论文. 浙江财经大学, 2023, Jin Liu et al. "Vine-Copula Seismic Functionality Evaluation Method of Medical Room Systems Based on Shaking Table Tests". In: *Earthquake Engineering & Structural Dynamics* 54.6 (2025), pp. 1499–1519, Edith Kovács and Tamás Szántai. "Vine copulas as a mean for the construction of high dimensional probability distribution associated to a Markov Network". In: *4th Workshop on Vine Copula Distributions and Applications*. See also arXiv preprint arXiv:1105.1697. 2011, Dániel Pfeifer. "Matrix-based vine copula building algorithms". MA thesis. Budapest University of Technology and Economics, 2022.

[24] Florian H. Hodel and John R. Fieberg. "cylcop: An R Package for Circular-Linear Copulae with Angular Symmetry". In: *bioRxiv* (2021), p. 2021.07.14.452253, Najibullah Karimi1 and Ali Dastbaravarde. "Circular-Linear Copulas: Properties and Applications". In: *15th Seminar on Probability and Stochastic Processes, University of Kurdistan*. 2025, pp. 200–211.

Theory Growth

Causal Analysis

- causal compression[25], causal structure learning[26], LiNGAM-MMI[27], and time series causal network construction[28].

[25] Aleksander Wieczorek and Volker Roth. "Causal Compression". In: *arXiv preprint arXiv:1611.00261* (2016).

[26] Leonie Bossemeyer. "Machine Learning for Causal Discovery with Applications in Economics". MA thesis. Ludwig-Maximilians-Universität München, 2021, Marvin Lasserre, Régis Lebrun, and Pierre-Henri Wuillemin. "Learning Continuous High-Dimensional Models using Mutual Information and Copula Bayesian Networks". In: *Thirty-Fifth AAAI Conference on Artificial Intelligence, AAAI 2021*. AAAI Press, 2021, pp. 12139–12146. URL: https://ojs.aaai.org/index.php/AAAI/article/view/17441, Marvin Lasserre. "Apprentissages dans les réseaux bayésiens à base de copules non-paramétriques". PhD thesis. Sorbonne Université, 2022. URL: https://tel.archives-ouvertes.fr/tel-03647090.

[27] Joe Suzuki and Tian-Le Yang. "Generalization of LiNGAM that allows confounding". In: *2024 IEEE International Symposium on Information Theory (ISIT)*. IEEE. 2024, pp. 3540–3545.

[28] 陈琦琦. "基于 Copula 熵和偏秩相关系数的时序因果网络构建算法及其应用". 硕士学位论文. 合肥工业大学, 2023, Jing Yang and Xinzhi Rao. "Copula Entropy Based Causal Network Discovery from Non-stationary Time Series". In: *Pattern Recognition*. Springer Cham, 2025, pp. 115–131.

Theory Growth

Machine Learning

- structure learning[29], clustering[30], nonlinear PCA[31], decision trees[32], reinforcement learning[33].

[29] Luiz Desuó Neto et al. "Learning Continuous Decomposable Models Using Mutual Information and Statistical Copulas". In: *Entropy* 28.3 (2026), p. 293. ISSN: 1099-4300. DOI: 10.3390/e28030293. URL: https://www.mdpi.com/1099-4300/28/3/293, Luiz Desuó Neto. "Untying the Gordian knot of Bayesian inference in Markov networks". PhD thesis. Universidade de São Paulo, 2025.

[30] 刘磊 et al. "基于非线性相关性和复杂网络的径流相似性分区". In: 水科学进展 33.3 (2022), pp. 442–451, Francesca Condino. "La divergenza di Jensen-Shannon nell'algoritmo di clustering dinamico per dati descritti da distribuzioni multivariate". PhD thesis. Università degli Studi di Napoli Federico II, 2009.

[31] Yingpeng Wei and Li Wang. "Copula entropy-based PCA method and application in process monitoring". In: *2022 4th International Conference on Intelligent Information Processing (IIP)*. 2022, pp. 61–64. DOI: 10.1109/IIP57348.2022.00019.

[32] Qingsong Shan and Qianning Liu. "Binary Trees for Dependence Structure". In: *IEEE Access* 8 (2020), pp. 150989–150998. DOI: 10.1109/ACCESS.2020.3017529, 罗良清 et al. "中国贫困治理经验总结：扶贫政策能够实现有效增收吗？". In: 管理世界 38.2 (2022), pp. 70–83.

[33] Yu-Xin Tian and Chuan Zhang. "A multimodal deep reinforcement learning framework for multi-period inventory decision-making under demand uncertainty". In: *Fuzzy Optimization and Decision Making* 24.4 (Dec. 2025), pp. 723–750. ISSN: 1573-2908. URL: https://doi.org/10.1007/s10700-025-09462-0.

Theory Growth

Neural Networks and Deep Learning

- graph neural networks[34], IEGAIN[35], neural network pruning[36], and tree-copula variational autoencoders[37].

[34] 陈佳雷 et al. "一种基于多要素注意力时空图卷积网络的径流预报方法". Pat. CN117151285A. CN117151285A. 2023, 汤宇飞. "基于脉搏波的糖尿病和高血压诊断算法研究". 硕士学位论文. 中国矿业大学, 2023.

[35] 武昊. "基于深度学习的化工过程软测量建模方法研究". 博士学位论文. 北京化工大学, 2023.

[36] Zihao Jiang et al. "CoPruning: Exploring the Parameter-Gradient Nonlinear Correlation for Neural Network Pruning Using Copula Function". In: *OpenReview.net* (2024). URL: https://openreview.net/forum?id=Yqqa9aNwB0.

[37] Yarden Rachamim and Shai Fine. "Beyond Mean-Field: Tree-Copula Variational Autoencoders for Structured Latent Dependencies". In: *Proceedings of the 42nd Conference on Uncertainty in Artificial Intelligence*. 2026, pp. 5601–5622.

Theory Growth

Control Theory

- model prediction controller[38] and sliding mode controller[39].

[38] Zhiwei Li et al. "Passenger-centric model predictive control for indoor temperature to facilitate energy-efficient airport terminal". In: *Building and Environment* 283 (2025), p. 113344. ISSN: 0360-1323. DOI: https://doi.org/10.1016/j.buildenv.2025.113344. URL: https://www.sciencedirect.com/science/article/pii/S0360132325008200.

[39] Ting Yang et al. "Cooperative Voltage Control in Distribution Networks Considering Multiple Uncertainties in Communication". In: *Sustainable Energy, Grids and Networks* 39 (2024), p. 101459. DOI: 10.1016/j.segan.2024.101459.

Theory Growth

Optimization

- CE based evolutionary algorithm[40], CE based estimation of vine copula distribution algorithm[41], and CE based grey wolf optimization[42].

[40] Stuart William Card. "Towards an Information Theoretic Framework for Evolutionary Learning". PhD thesis. Syracuse University, 2011.

[41] 王筱萍. "基于分导 Copula 函数的分布估计算法研究". 博士学位论文. 兰州理工大学, 2013.

[42] Hongyu Pan, Shanxiong Chen, and Hailing Xiong. "A high-dimensional feature selection method based on modified Gray Wolf Optimization". In: *Applied Soft Computing* 135 (2023), p. 110031. ISSN: 1568-4946. DOI: https://doi.org/10.1016/j.asoc.2023.110031. URL: https://www.sciencedirect.com/science/article/pii/S1568494623000492.

Theory Growth

Approximation Theory

- B-spline approximation[43].

[43] 張🄵棠. "基於 Copula Entropy 的變數選取方法與節點選取方法". 硕士学位论文. 国立政治大学, 2024.

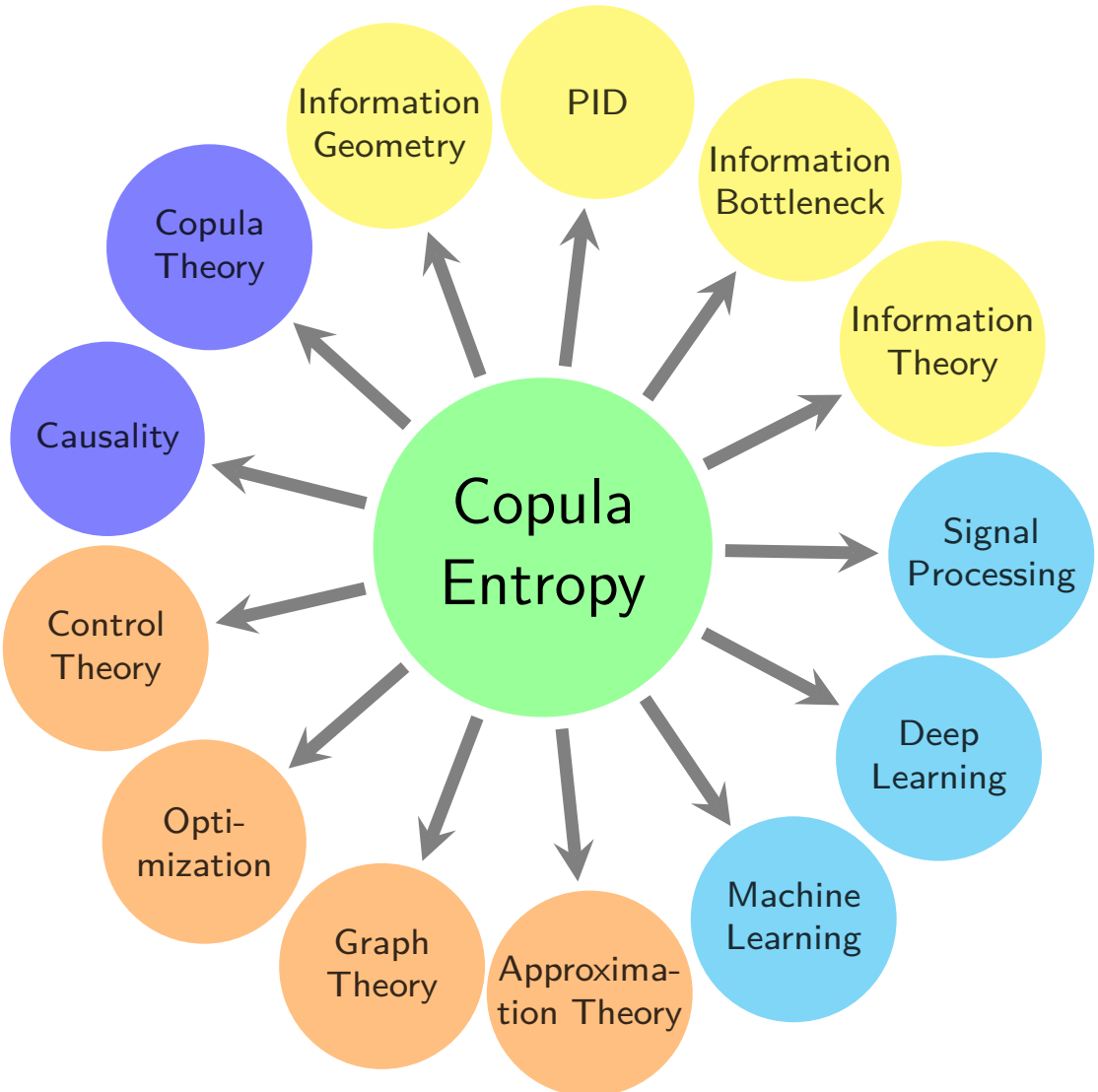
Theory Growth

Signal Processing

- copula component analysis[44].

[44] Paola Palmitesta and Corrado Provasi. "Copula component analysis for dependence modelling". In: *Quaderni di Statistica* 14 (2012), pp. 185–188.

Theory Growth



Monograph

Jian Ma. "Copula Entropy: Theory and Applications". In: *arXiv preprint arXiv:2025.18168* (2025)

Thanks!

Q&A