

ニューラルネット

カップリングに基づく可逆神経回路網は 微分同相写像の万能近似器である

手嶋毅志^{1 2}, 石川 勲^{3 2}, 東條 広一², 大野 健太¹, 池田 正弘², 杉山 将^{2 1}

¹東京大学 ²理化学研究所 ³愛媛大学

2020年11月26日

第23回情報論的学習理論ワークショップ (IBIS2020)

概要

本研究の問い	可逆神経網は十分な表現力を持つか?
問いの重要性	モデルの表現力保証はその実応用において重要な理論的土台.
主結果	カップリングに基づく可逆神経網は万能近似器.

本研究はNeurIPS2020に採択され, Oral発表に選出された.



問い

可逆神経網(INN)は十分な表現力を持つか？



問いの重要性

- モデルの**表現力保証**は実応用において**重要**
- 広い応用を持つにも関わらずINNには表現力保証が無い場合が多かった

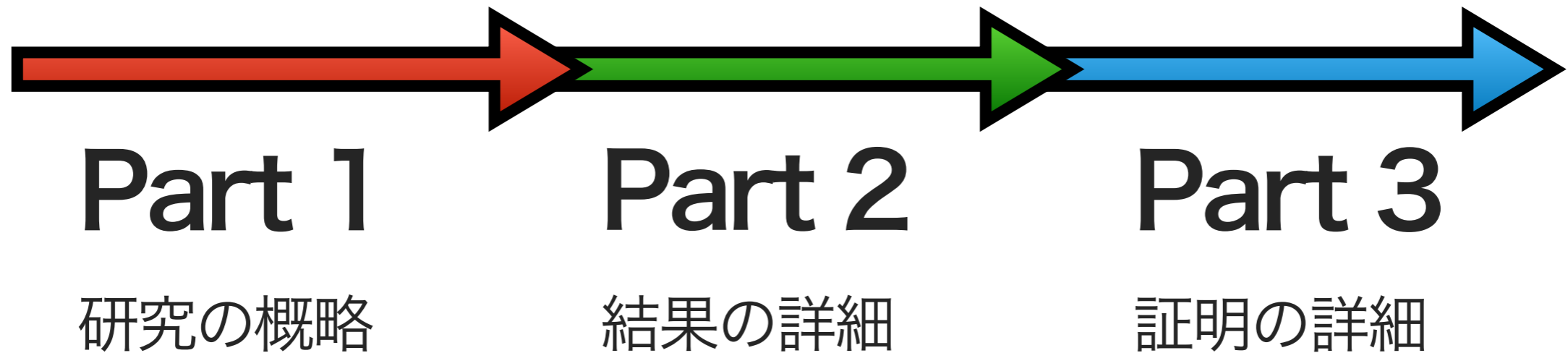


主結果

カップリングに基づくINNは非常に単純な構造ながら**万能近似器**

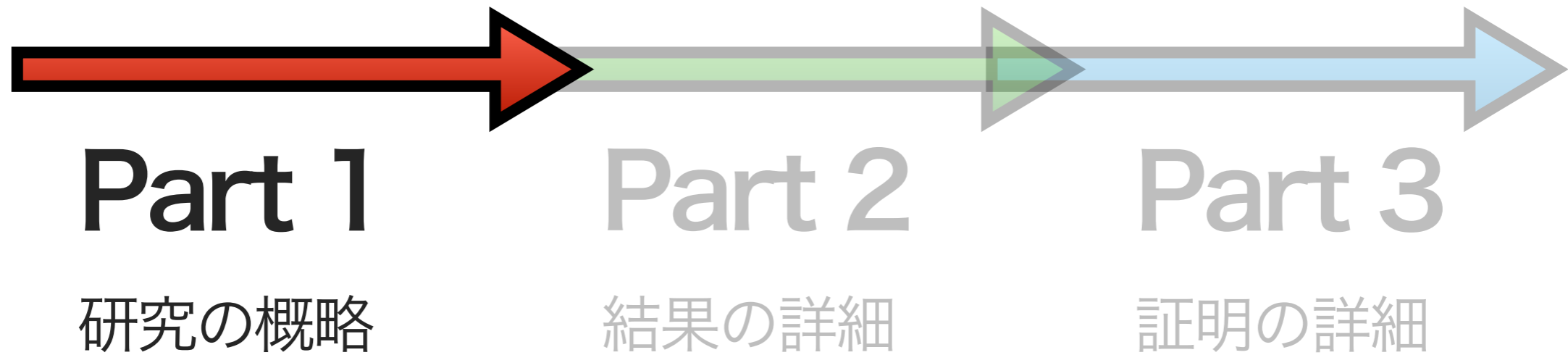
メッセージ

「カップリングに基づくINN」は可逆関数や確率分布を表現するモデルとして十分な表現力を持つ



注意

本研究の内容はarXiv: 2006.11469にまとめられている. 本スライドの定義等の正確な表記については  Sec. ?として論文の箇所を明記する.



- 本研究の分析対象のモデルとリサーチクエスチョン

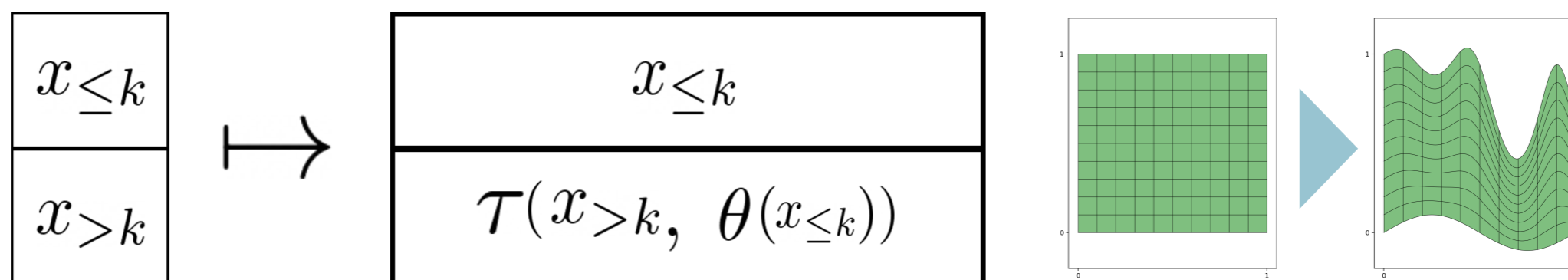
定義

Sec. 2.1

可逆神経回路網 (invertible neural network, INN) とは
アフィン変換層と「可逆フロー」層の有限個の合成のこと

例 (カップリングレイヤー) [DKB14, PNRML19, KPB19]

Sec. 2.1



カップリングフローの特徴: 一部の成分をそのまま次の層へ渡す.

CF-INN=カップリングに基づく可逆神経回路網

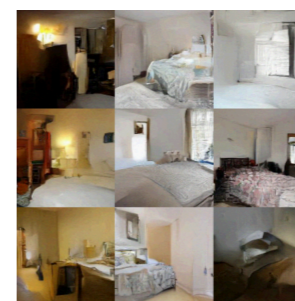
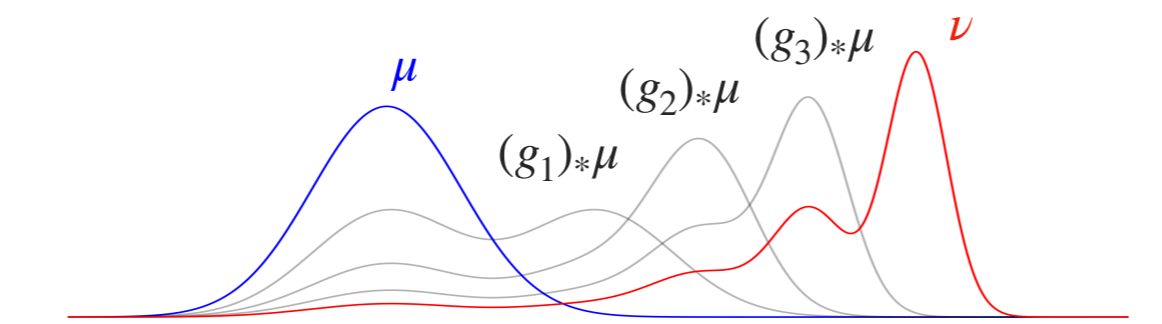
問い

CF-INNsは十分な表現力を持つか？

CF-INNsの用途

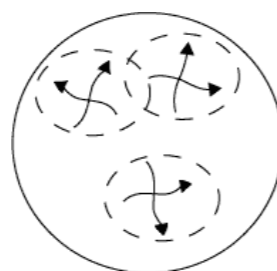
Sec. 1

- 確率分布の近似 (normalizing flow).



[KD18]

- 可逆写像の近似 (特徴抽出 & 特徴操作).



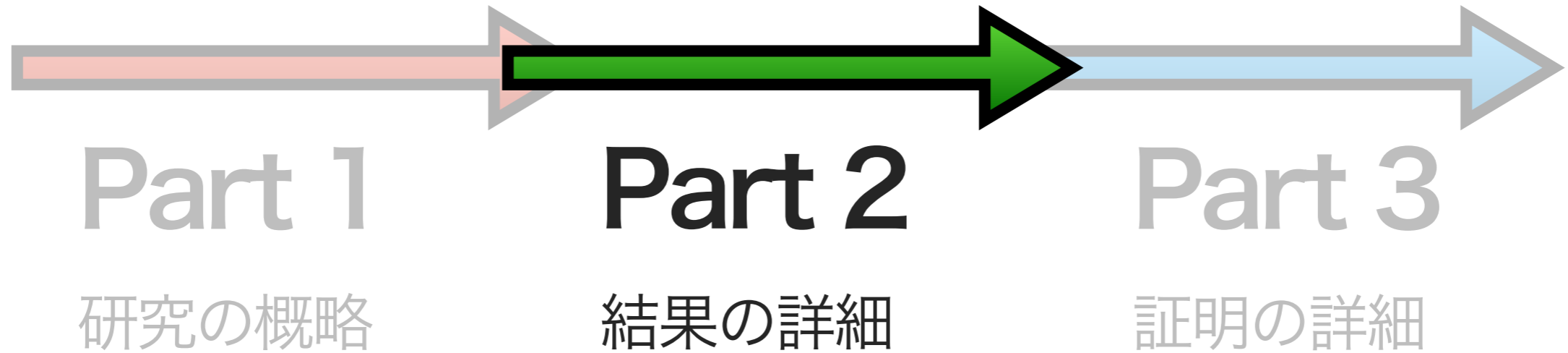
[DSB17]

問い

CF-INNsは十分な表現力を持つか？

答え

はい。



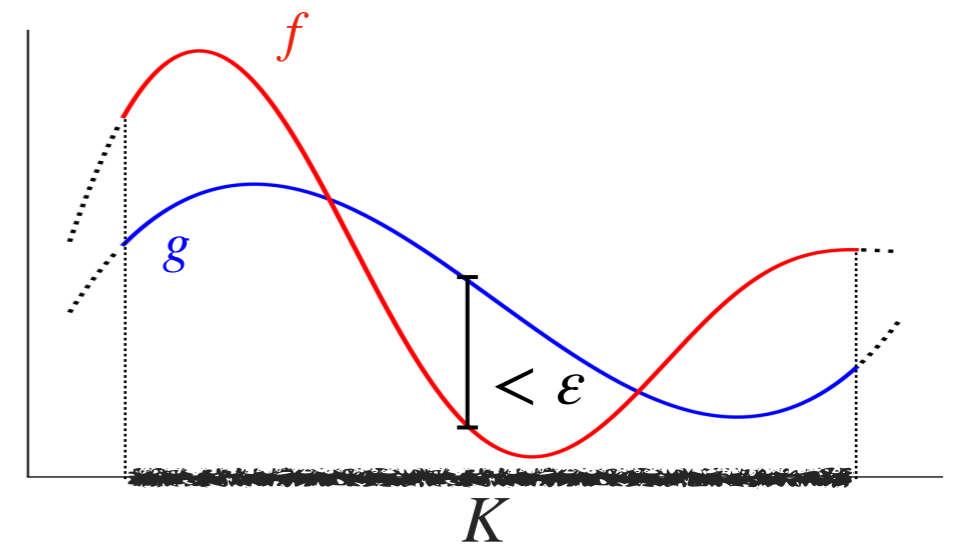
- 主結果の紹介とその意味について

表現力 = 万能近似能力 (Universality)

定義 [C89,HSW89]

sup- (L^p) -万能近似器とは近似対象関数を (sup- (L^p) -ノルムの意味で)、与えられたコンパクト集合上で任意精度で近似できるモデルのこと

Sec. 2.2



本研究で考える近似対象 \mathcal{D}^2

滑らかな \mathbb{R}^d から \mathbb{R}^d への可逆写像の集合

↑ かなり大きい

Sec. 3.1



arXiv: 2006.11469

定理

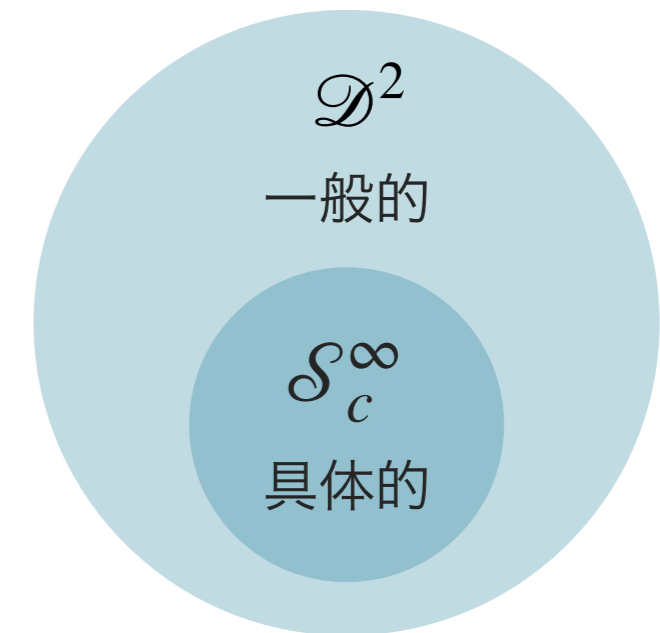
📄 Sec. 3.1, Theorem 1

あるINNモデルが

\mathcal{S}_c^∞ に対するsup- (L^p -)万能近似器

$\implies \mathcal{D}^2$ に対するsup- (L^p -)万能近似器

$\mathcal{S}_c^\infty := \{\tau : \text{コンパクト台 } \tau(\mathbf{x}, y) = (y, u(\mathbf{x}, y))\}$



応用

この定理から既存のCF-INNモデルである

- Sum-of-squares polynomial flow (SoS) [JSY19]
- Deep sigmoidal flow (DSF) [HKLC18]

が \mathcal{D}^2 に対するsup-万能近似器であることが従う

↑ [JSY19, HKLC18]で示された結果より**ずっと強い表現力保証**

定義 [DKB14,DSB17,KD18]

📄 Sec. 2.1

(単成分) **アフィンカップリングフロー** (ACFs) とは次式で定義されるカップリングフローの1種:

$$\Psi_{s,t}(\mathbf{x}, y) := (\mathbf{x}, e^{s(\mathbf{x})}y + t(\mathbf{x}))$$

ACFの重要性について

- 広範な応用
 - 生成モデル [DSB17,KD18,OLB+18,KLSKY19,ZMWN19]
 - 確率的推論 [BM19,WSB19,LW17,AKRK19]
 - 半教師付き学習 [IKFW20]
 - 転移学習 [TSS20]



[DSB17]

- 最も簡便なカップリングレイヤー

→ACFの理論保証は、より複雑なカップリング層にも自動的に適用

📄 = arXiv: 2006.11469

定理

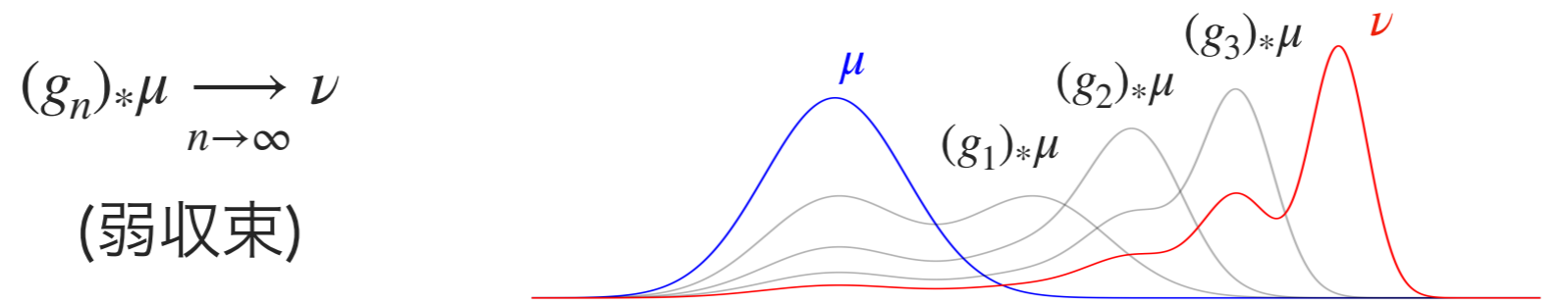
Sec. 3.2, Theorems 2, 3

ACFに基づく可逆神経回路網は \mathcal{D}^2 に対する L^p -万能近似器である。
さらに、そこから分布万能近似器であることも従う。

定義

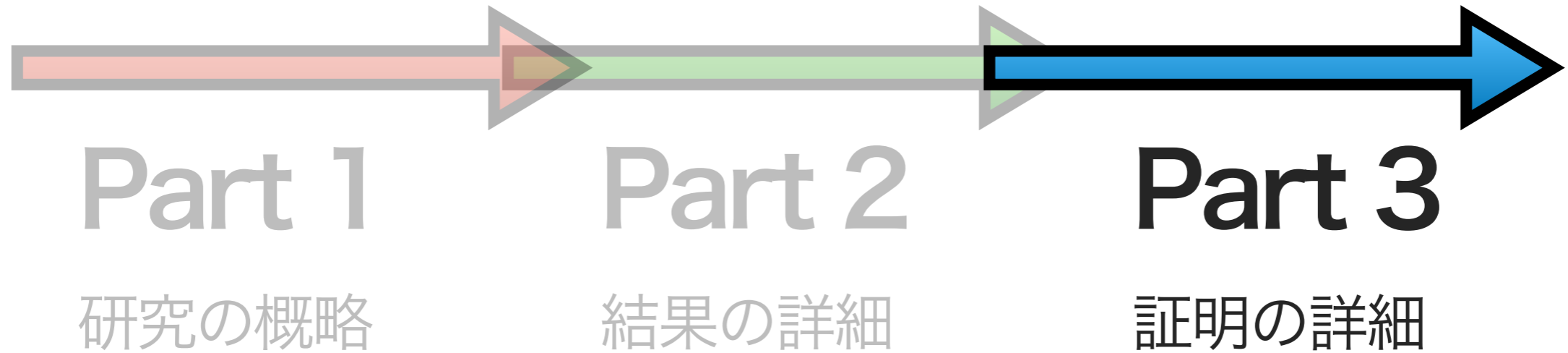
Sec. 2.2

あるモデルが分布万能近似器であるとは所与の分布を別の任意の分布に（弱収束の意味で）任意に近づけられること



定理の系

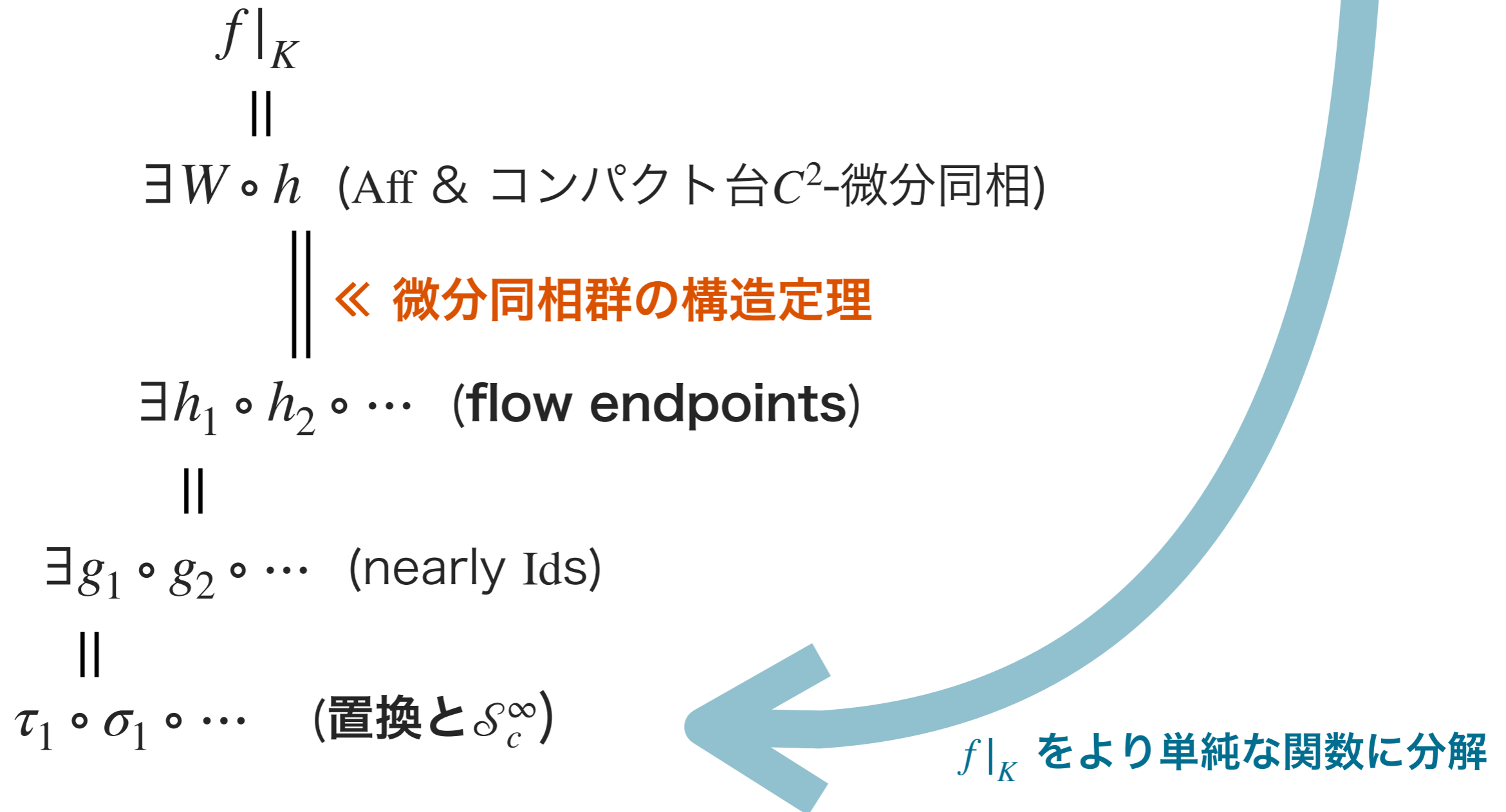
- もしあるINNのフロー層がACFを特別な場合として含めば、そのINNは (L^p -分布) 万能近似器である
- ACF-INNの分布近似能力に関する未解決問題を肯定的に解決



- 主結果1の証明の概略

主結果1の証明の概略

$f \in \mathcal{D}^2$: 近似対象, $K \subset U_f$: コンパクト





問い

可逆神経網(INN)は十分な表現力を持つか？



問いの重要性

- モデルの**表現力保証**は実応用において**重要**
- 広い応用を持つにも関わらずINNには表現力保証が無い場合が多かった



主結果

カップリングに基づくINNは非常に単純な構造ながら**万能近似器**

メッセージ

「カップリングに基づくINN」は可逆関数や確率分布を表現するモデルとして十分な表現力を持つ

- [C89] Cybenko, G. (1989).
Approximation by superpositions of a sigmoidal function.
Mathematics of Control, Signals, and Systems, 2, 303–314.
- [HSW89] Hornik, K., Stinchcombe, M., & White, H. (1989).
Multilayer feedforward networks are universal approximators.
Neural Networks, 2(5), 359–366.
- [JSY19] Jaini, P., Selby, K. A., & Yu, Y. (2019).
Sum-of-squares polynomial flow.
Proceedings of the 36th International Conference on Machine Learning, 97, 3009–3018.
- [HKLC18] Huang, C.-W., Krueger, D., Lacoste, A., & Courville, A. (2018).
Neural autoregressive flows.
Proceedings of the 35th International Conference on Machine Learning, 80, 2078–2087.
- [KD18] Kingma, D. P., & Dhariwal, P. (2018).
Glow: Generative flow with invertible 1x1 convolutions.
In Advances in Neural Information Processing Systems 31 (pp. 10215–10224).
- [PNRML19] Papamakarios, G., Nalisnick, E., Rezende, D. J., Mohamed, S., & Lakshminarayanan, B. (2019).
Normalizing flows for probabilistic modeling and inference.
ArXiv:1912.02762 [Cs, Stat].
- [KPB19] Kobyzev, I., Prince, S., & Brubaker, M. A. (2019).
Normalizing flows: An introduction and review of current methods.
ArXiv:1908.09257 [Cs, Stat].

-
- [DKB14] Dinh, L., Krueger, D., & Bengio, Y. (2014).
NICE: Non-linear independent components estimation.
ArXiv:1410.8516 [Cs.LG].
- [DSB17] Dinh, L., Sohl-Dickstein, J., & Bengio, S. (2017).
Density estimation using Real NVP.
Fifth International Conference on Learning Representations (ICLR)
- [AKRK19] Ardizzone, L., Kruse, J., Rother, C., & Köthe, U. (2019).
Analyzing inverse problems with invertible neural networks.
7th International Conference on Learning Representations.
- [BM19] Bauer, M., & Mnih, A. (2019).
Resampled priors for variational autoencoders.
In Proceedings of machine learning research, 89, 66–75.
- [LW17] Louizos, C., & Welling, M. (2017).
Multiplicative normalizing flows for variational Bayesian neural networks.
In Proceedings of the 34th International Conference on Machine Learning, 70, 2218–2227.
- [NMT+19] Nalisnick, E. T., Matsukawa, A., Teh, Y. W., Görür, D., & Lakshminarayanan, B. (2019).
Hybrid models with deep and invertible features.
In Proceedings of the 36th International Conference on Machine Learning, 97, 4723–4732.
- [IKFW20] Izmailov, P., Kirichenko, P., Finzi, M., & Wilson, A. G. (2020).
Semi-supervised learning with normalizing flows.
Proceedings of the 37th International Conference on Machine Learning.

-
- [KD18] Kingma, D. P., & Dhariwal, P. (2018).
Glow: Generative flow with invertible 1x1 convolutions.
In *Advances in Neural Information Processing Systems* 31, 10215–10224.
- [OLB+18] Oord, A., Li, Y., Babuschkin, I., Simonyan, K., Vinyals, O., Kavukcuoglu, K., Driessche, G., Lockhart, E., Cobo, L., Stimberg, F., Casagrande, N., Grewe, D., Noury, S., Dieleman, S., Elsen, E., Kalchbrenner, N., Zen, H., Graves, A., King, H., ... Hassabis, D. (2018).
Parallel WaveNet: Fast high-fidelity speech synthesis.
Proceedings of the 35th International Conference on Machine Learning, 80, 3918–3926.
- [TSS20] Teshima, T., Sato, I., & Sugiyama, M. (2020).
Few-shot domain adaptation by causal mechanism transfer.
Proceedings of the 37th International Conference on Machine Learning.
- [KLSKY19] Kim, S., Lee, S.-G., Song, J., Kim, J., & Yoon, S. (2019).
FloWaveNet: A generative flow for raw audio.
In *Proceedings of the 36th International Conference on Machine Learning*, 97, 3370–3378.
- [ZMWN19] Zhou, C., Ma, X., Wang, D., & Neubig, G. (2019).
Density matching for bilingual word embedding.
Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers), 1588–1598.
- [WSB19] Ward, P. N., Smofsky, A., & Bose, A. J. (2019).
Improving exploration in soft-actor-critic with normalizing flows policies.
ArXiv:1906.02771 [Cs, Stat].