

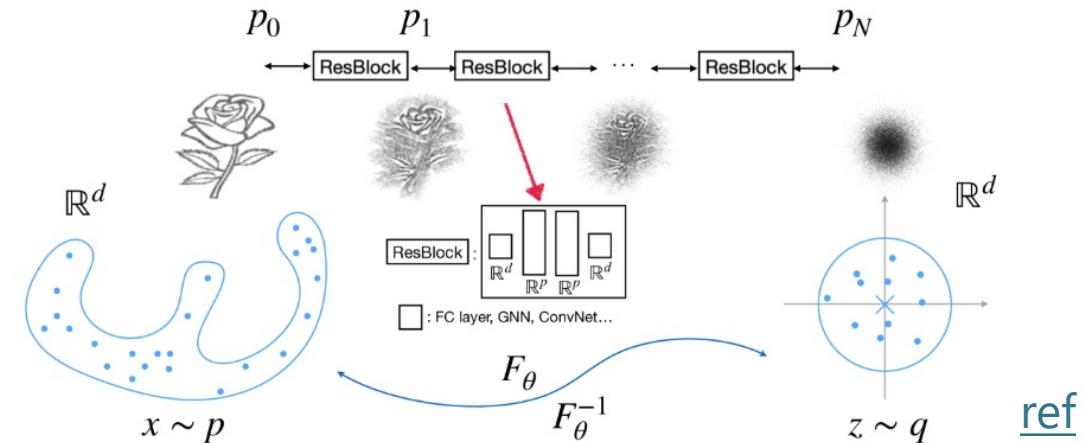
セミナーの内容

- 拡散モデルおよび確率過程の利用の動機
- 確率過程の簡単な説明
- 拡散モデルの学習と生成
- より一般的な問題「Schrödinger Bridge」
- 条件付き問題に適した「Stochastic Interpolant」
- 微気象超解像への応用

データ生成と微分方程式

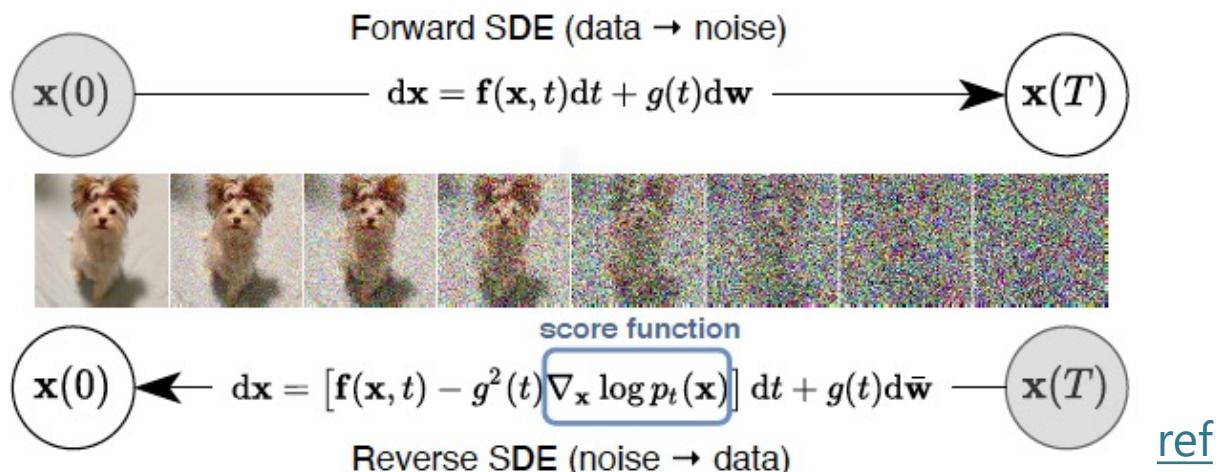
■ データ生成

- 画像や文章などを ML モデルで作成
 - 現代の AI 技術の基礎
- 既知データに変換を施す
 - 複雑なデータの生成は変換も複雑



■ 微分方程式の応用

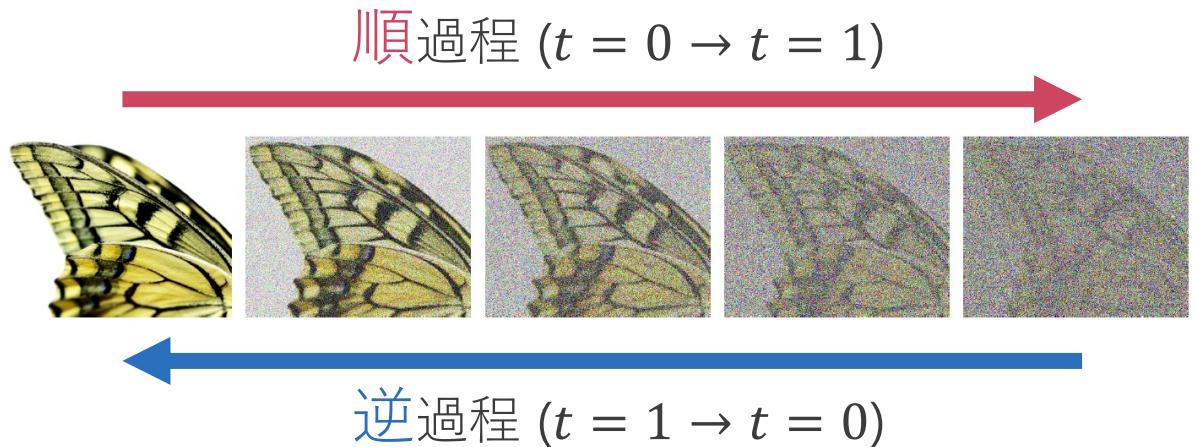
- 微小変換の積み重ねでデータを生成
- 微分方程式の数値積分と類似
 - 正規化流, Neural ODE など
- 拡散モデルは SDE でデータ生成 ([ref](#))
(SDE: 確率微分方程式)



拡散モデルのデータ生成の概要

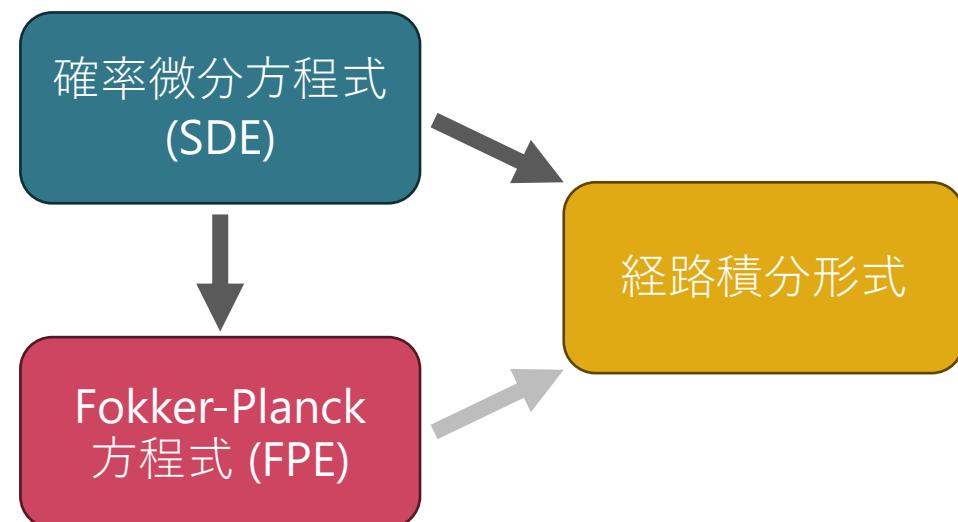
■ 順過程と逆過程の利用 (ref)

- 順過程: 拡散過程
 - データをノイズに変換
- 逆過程: 順過程の逆回し
 - ノイズからデータを生成



■ その背後にある数理: 確率解析 (ref)

- SDE: ランダム系の時系列の発展
- FPE: アンサンブルの分布の時間発展
- 経路積分形式: 時系列 1 本の確率



順過程: データからノイズへ

■ 拡散モデルの原理における主役

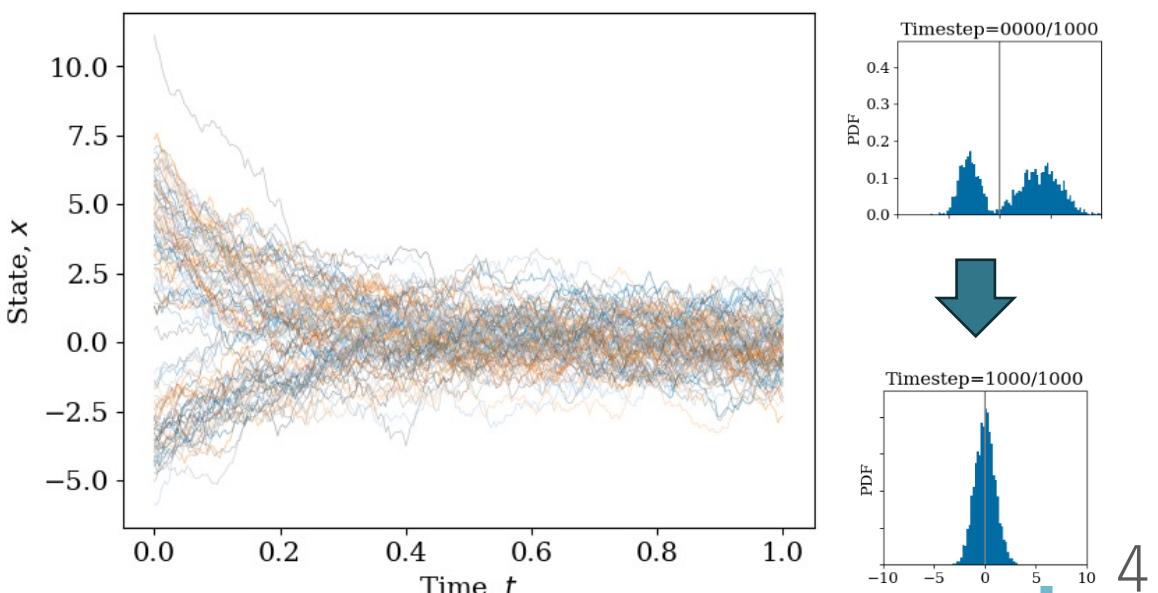
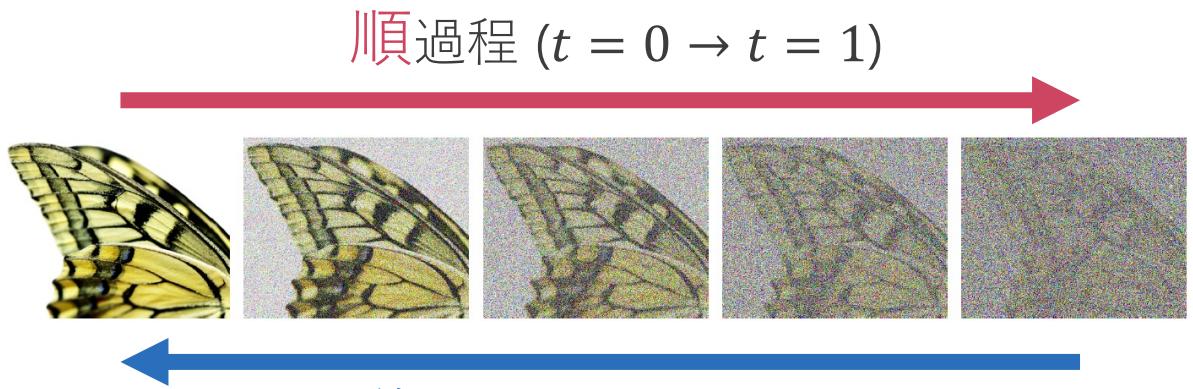
- 確率微分方程式 (SDE)
- x : 画像などの 1 サンプル
- $dx = -x dt + dW$

■ ノイズ (第二項) がなければ

- $x_t = x_0 e^{-t}$
- 初期条件を 0 へと **指数的**に減衰させる

■ ノイズ (第二項) があるため

- 減衰しつつ、単峰の正規ノイズになる



逆過程: ノイズからデータへ

■ 逆 SDE は閉じていない

- 不可逆な過程を無理やり逆転したから
 - 確率分布 $p(x, t)$ が必要
 - 理論的に逆 SDE は導出可能 ([ref](#))
 - ML は直接は関係ない

■ 逆 SDE は何をしているのか？

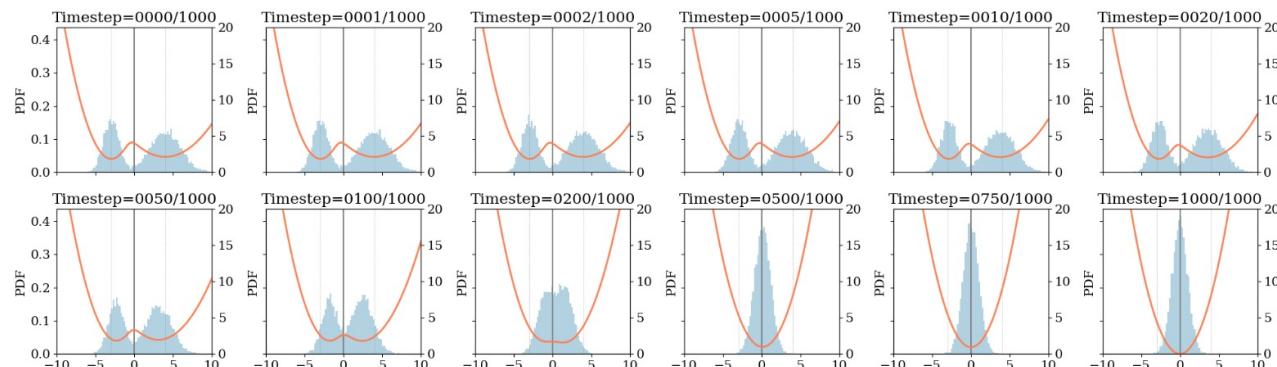
- 拡散過程の逆回し
 - 確率の高い状態へ緩和
(ポテンシャルの低い状態)
- $t = 0$ でデータ分布を再現
 - ポテンシャルは段々複雑になってく
 - $t = 0$ でデータ分布のポテンシャル

スコア関数

$$dx = \left[x + \frac{\partial}{\partial x} \ln p(x, t^R) \right] dt^R + dW$$

$$dx = \left[x - \frac{\partial}{\partial x} U(x, t^R) \right] dt^R + dW$$

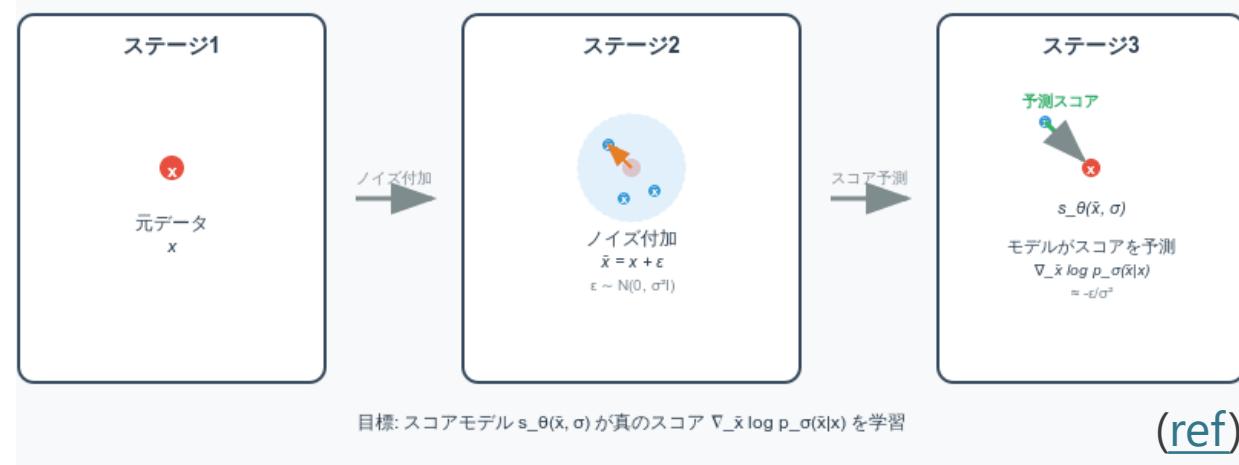
$$p(x, t) \propto \exp[-U(x, t)]$$



拡散モデルの学習と生成

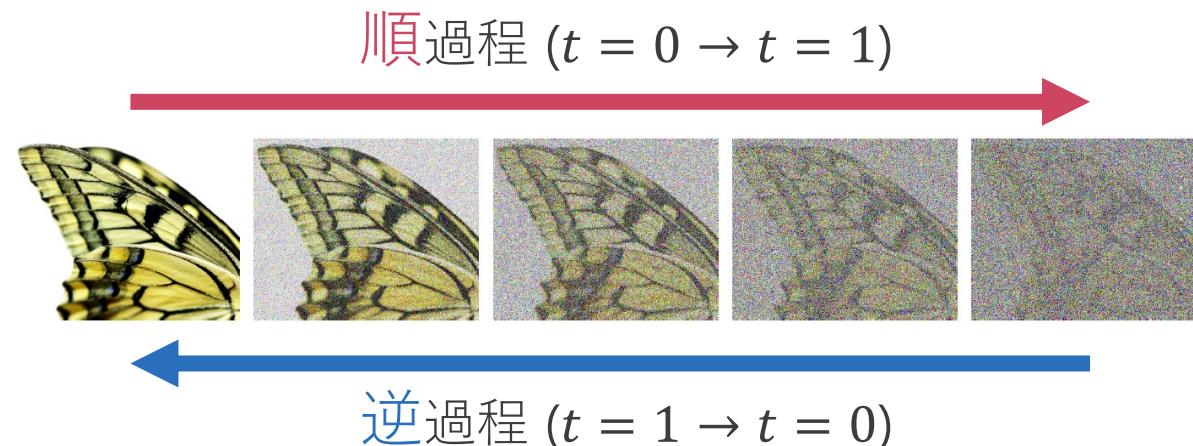
■ 学習: スコア関数の推定 (ref1, ref2)

- 正規ノイズがスコア関数に比例
- ノイズの推定問題を解く
- ノイズ (摂動) の平均的な傾向の推定
→ スコア関数 (感度) の推定



■ 生成: 逆 SDE の積分

- 初期値を正規分布からサンプリング
- 逆 SDE を数値的に解く
- 初期値毎に異なるサンプルが得られる

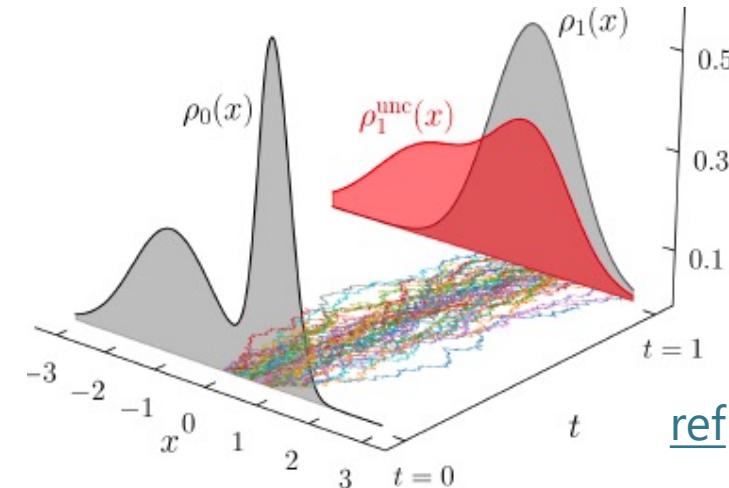


Schrödinger Bridge (SB) 問題: 任意の分布の間の変換

■ 正規ノイズから生成は無駄?

- 変換「元」があることが多い
 - 超解像: LR \rightarrow HR
- 任意の分布の間の変換問題

→ Schrödinger Bridge (SB) 問題 ([ref](#))



■ 4D-Var に似た最適化問題へ

- 順 SDE の速度 u (ドリフト) の推定問題
 - 順 SDE でデータ生成 (or 変換)
- 推定問題は「エネルギー」の最小化
 - 確率分布 p は連続の式 (FPE) を満たし
 - かつ, 端点条件を満たす
(物理方程式と観測値への対応)

順過程の SDE: $dx = u(x, t)dt + dW$

$$\min_{u,p} \frac{1}{2} \int_0^1 u(x, t)^2 p(x, t) dx dt$$

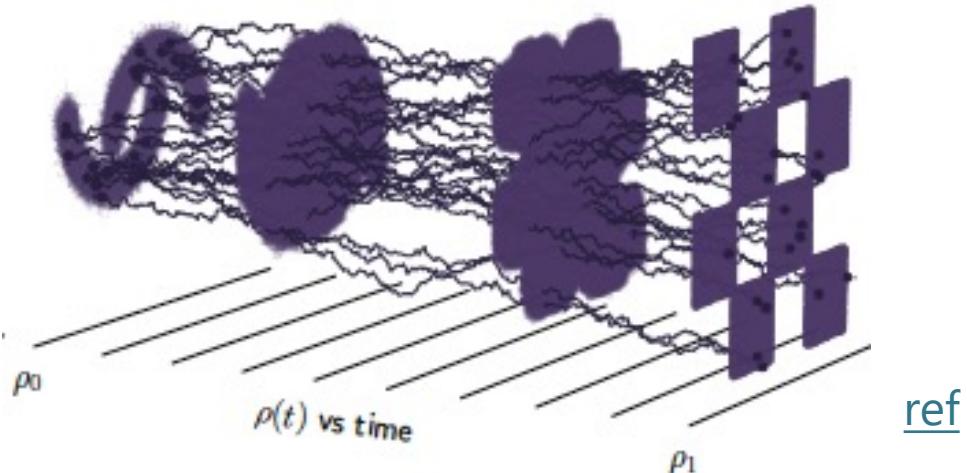
$$\frac{\partial}{\partial t} p(x, t) = -\frac{\partial}{\partial x} [u(x, t)p(x, t)] + \frac{1}{2} \frac{\partial^2}{\partial x^2} p(x, t)$$

$$p(x, 0) = \rho_0(x) \quad \text{and} \quad p(x, 1) = \rho_1(x) \quad 7$$

最適化問題の解法

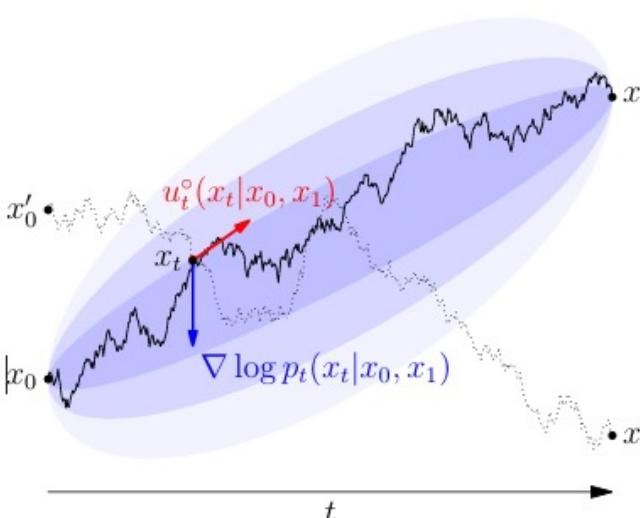
■ 随伴方程式による解法 ([ref1](#), [ref2](#))

- 順方向に連続の式を解く
- 逆方向にその随伴方程式を解く
- 繰り返すと「大域」最適解へ収束
 - SDE による変換を行う 1 つの理由



■ その他の解法 (ML を利用)

- Iterative Proportional Fitting ([ref](#))
- Forward-Backward SDE ([ref](#))
- Flow Matching ([ref](#))



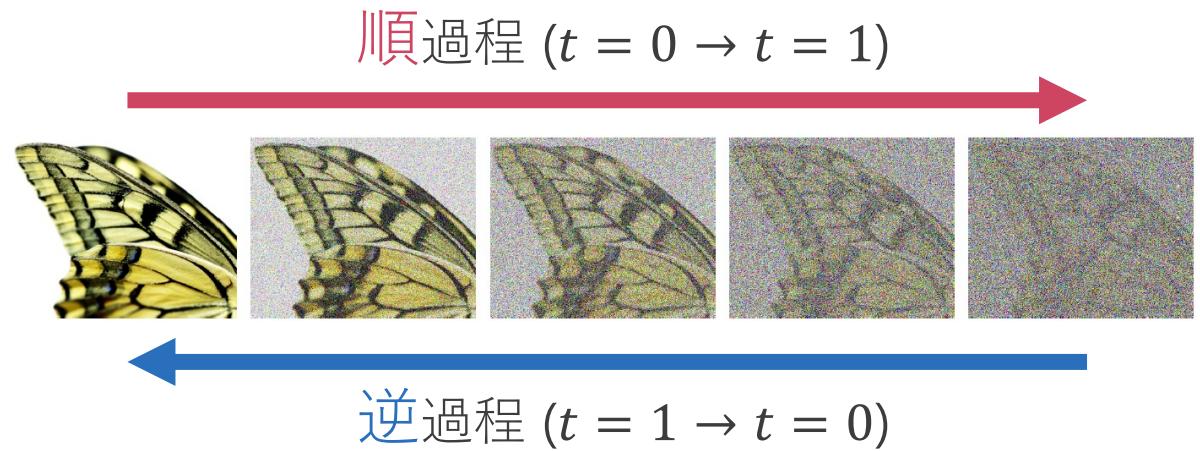
Flow Matching による学習

1. 輸送計画を求めて端点のペアを構築
2. ペア間の変換速度を Probability Flow ODE でドリフトとスコアに分解
3. ドリフトとスコアを学習

Schrödinger Bridge (SB) と拡散モデルと関係

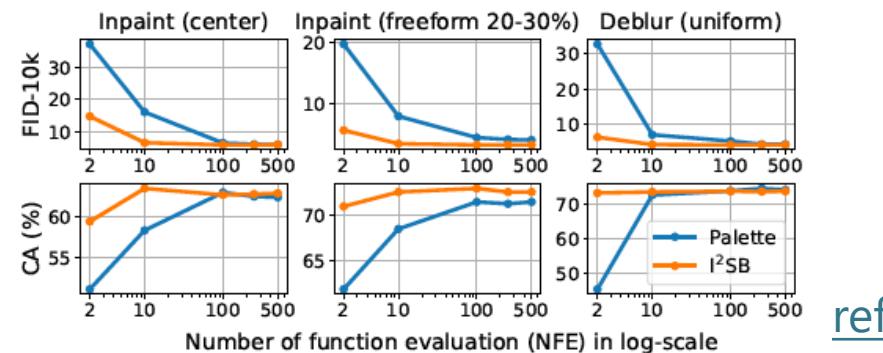
■ 拡散モデルは SB の近似解 ([ref](#))

- 拡散モデル: 端点条件を外している
 - 強い摩擦とノイズを加えることで $u(x, t) \equiv -x$ を(自明な)解とする
→ 反復法で最適化問題を解く必要ない



■ 逆過程によるデータ生成

- スコア関数で SDE を時間反転できるのは一般的な理論的結果 ([ref](#))
- 定式化によっては SB 問題も逆過程を利用して生成を行う ([ref](#))

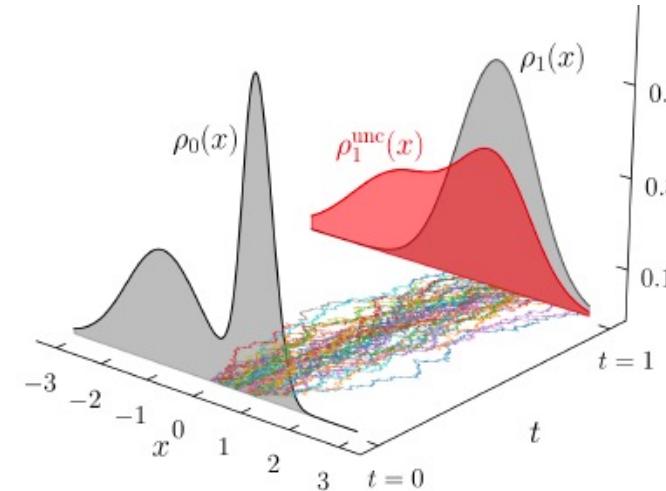


拡散モデルは、制約として多段ステップが必要
強い(指数的)減衰を数値積分でよく再現するため

より「条件付き問題」に適したアルゴリズムへ

■ SB 問題は「条件付き」構造を「積極的」には利用しない

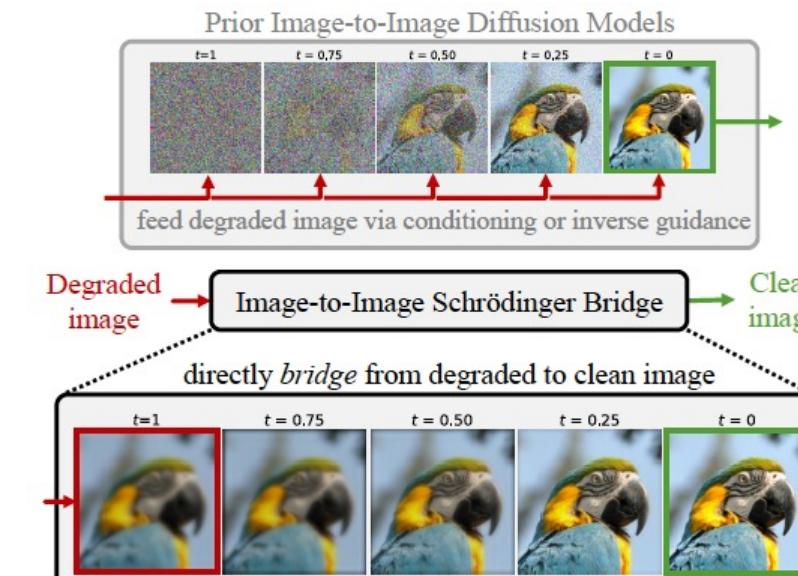
- 確率分布間の変換を学習できるし始点と終点の結合分布も得られる
- (e.g.,) 予測問題は現在→未来が欲しい
 - 大域最適性は常に必要ではない



始点と終点の
どの点同士が対応?
→ 最適輸送問題を
解いて分かる

■ 条件付き構造を利用した枠組み

- I²SB ([ref1](#), [ref2](#))
- Stochastic Interpolant ([ref1](#), [ref2](#), [ref3](#))
 - 拡散モデル, SB 問題, Flow Matchingなどを含みうる枠組み



条件付けを積極的に
利用し最適化の反復
を必要なしに

Stochastic Interpolant

■ 学習と生成 ([ref1](#), [ref2](#), [ref3](#))

- ペア (x_0, x_1) のノイズ付き補間
- その補間の速度を学習
- 学習した速度を SDE に組み込む
- SB の解になりうる

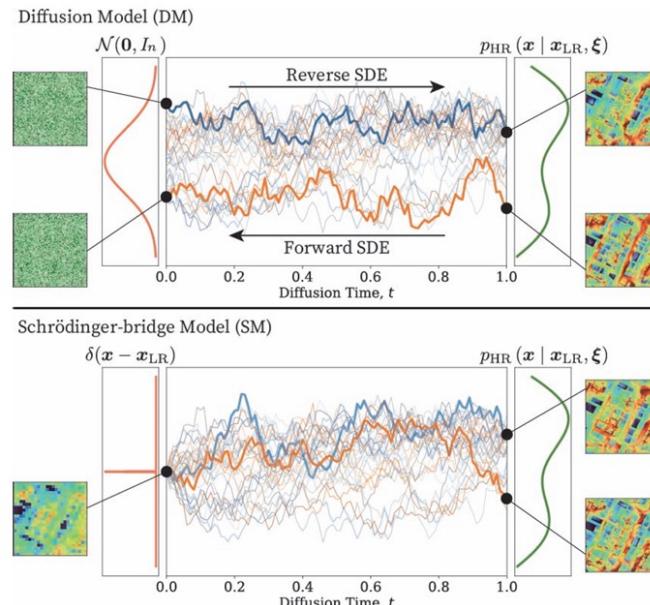
$$I_t = \alpha_t x_0 + \beta_t x_1 + \gamma_t z \quad t \in [0, 1],$$

$$L_b(\hat{b}) = \int_0^1 \mathbb{E} \left[|\hat{b}_t(I_t)|^2 - 2\dot{I}_t \cdot \hat{b}_t(I_t) \right] dt,$$

$$dX_t^F = b_t(X_t^F)dt - \epsilon_t \gamma_t^{-1} g_t(X_t^F)dt + \sqrt{2\epsilon_t} dW_t,$$

■ Stochastic Interpolant の特殊形 ([ref3](#))

- 点分布から条件付き分布への生成
- SB 問題の解だが、反復問題を解かず
に学習が可能
- 拡散モデルも同様に条件付き分布を
近似するが、有限のデータではノイズ
からの変換を学習する方が大変 ([ref](#))

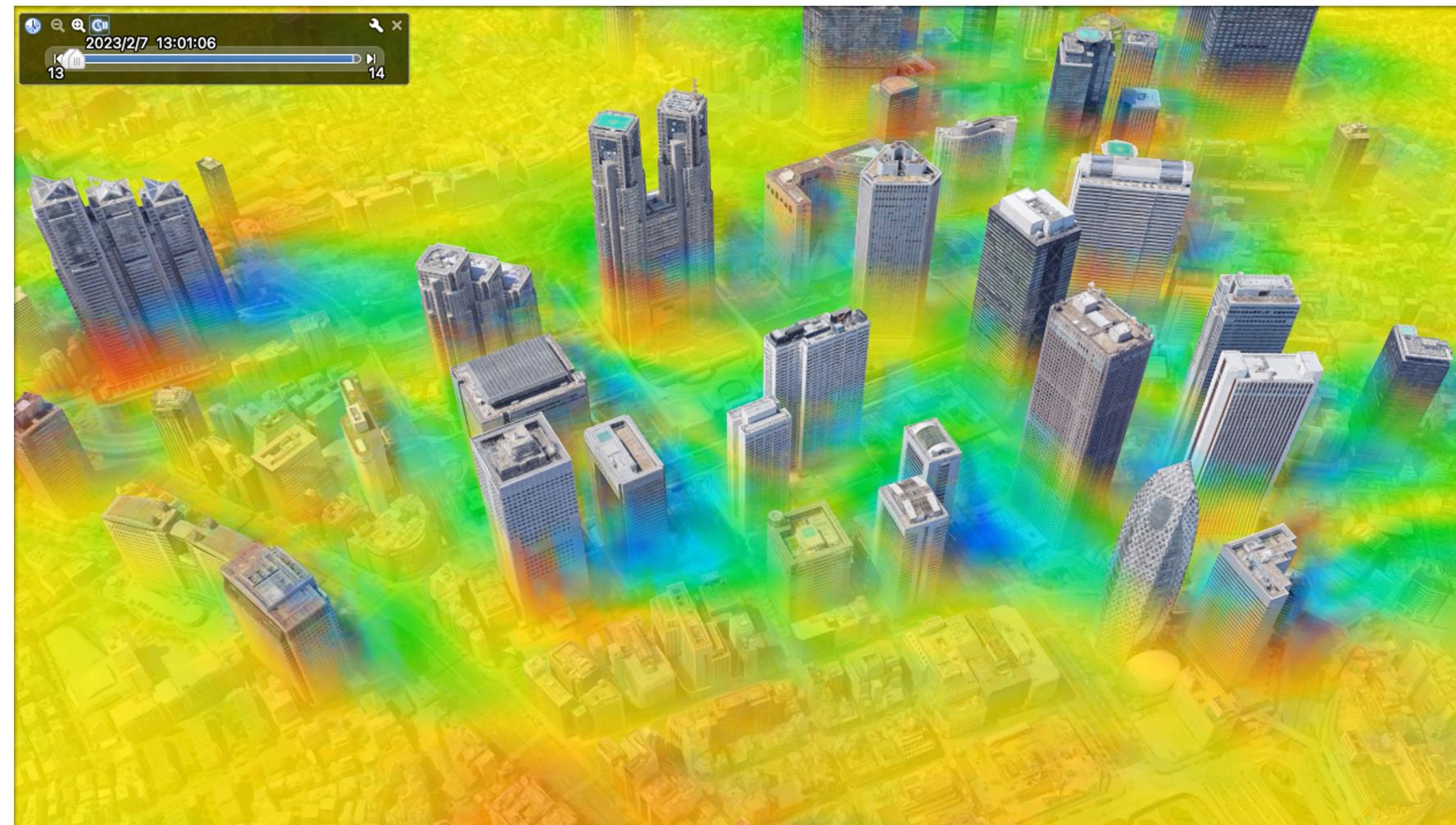


拡散モデルと同様に
条件付き分布を学習可能
しかし、変換元のサンプルが
異なる

([ref](#))

微気象超解像への応用

Prediction with Our Hybrid Method (Tokyo, 2023/02/07 13:00)

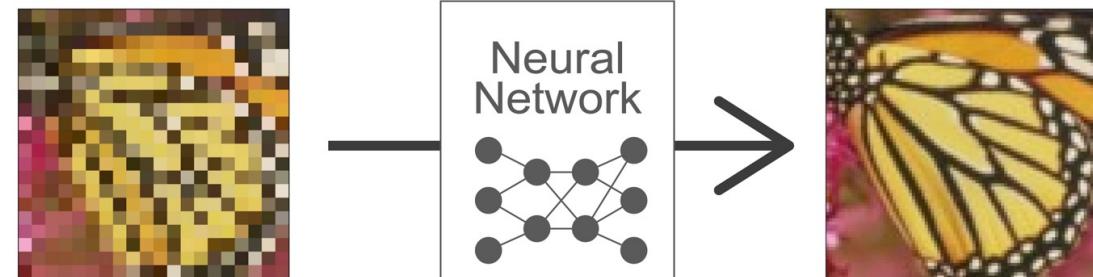


Wind Speed [m/s] 0.0 2.5 5.0 7.5 10 13

Hybrid Acceleration Using Super-Resolution

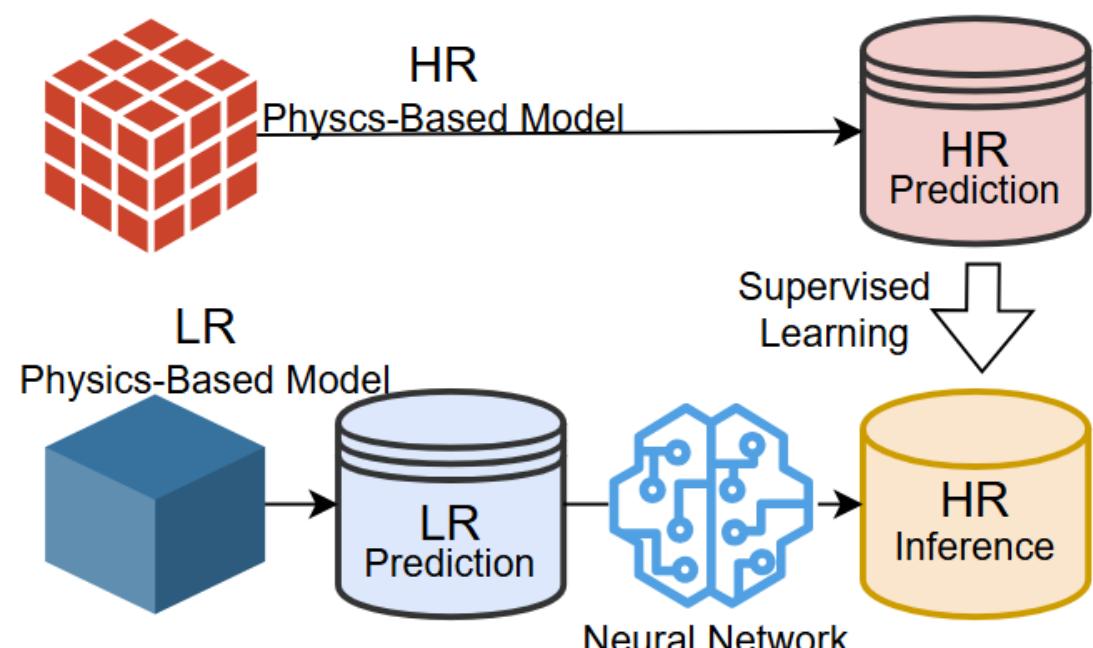
■ Super-Resolution (SR)

- Enhancing image resolution
- Studied with deep learning



■ Hybrid acceleration method ([Onishi+19](#))

- Training phase
 - learning from HR and LR predictions
- Inference (or testing) phase
 - super-resolving LR predictions
 - without costly HR computation



HR: High Resolution, LR: Low Resolution

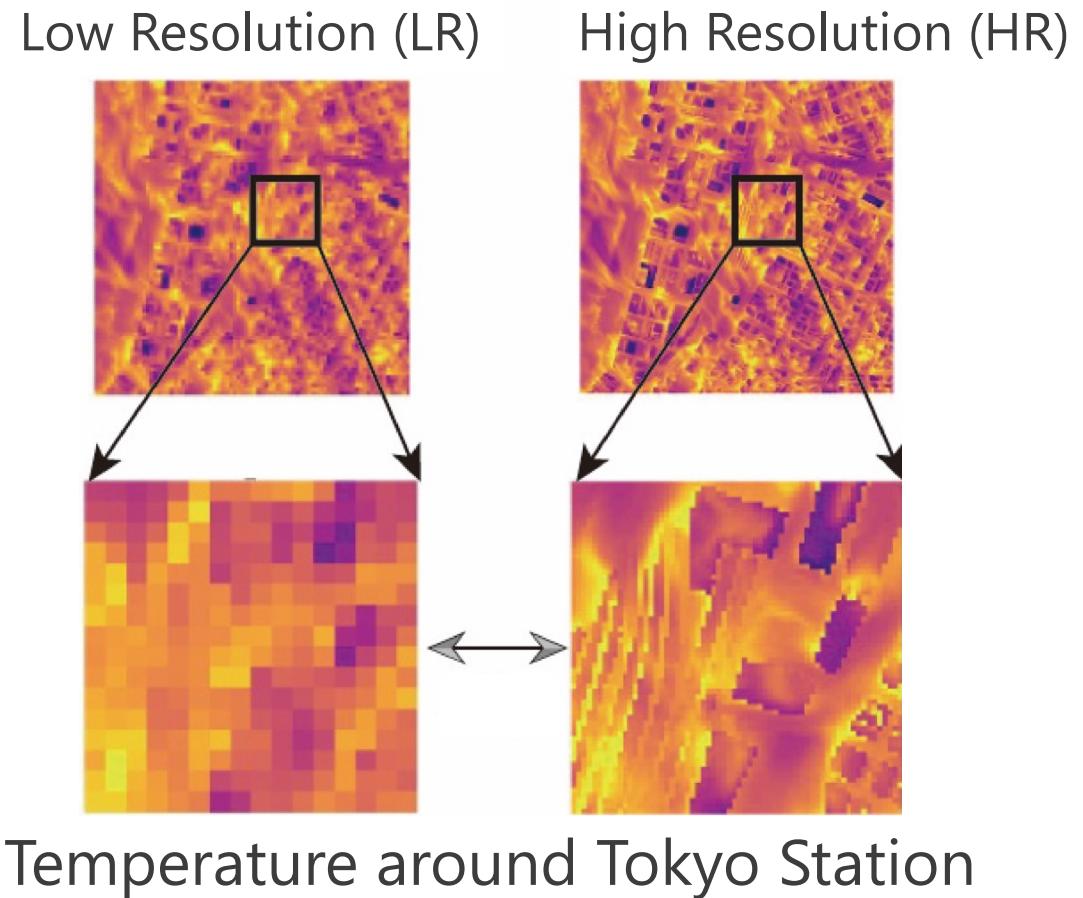
A Mechanism for SR and its Advantages

■ Dictionary-type SR

- Use of **local-pattern pairing**
 - learning from this relation
 - inferring using the learned relation

■ Advantages of SR

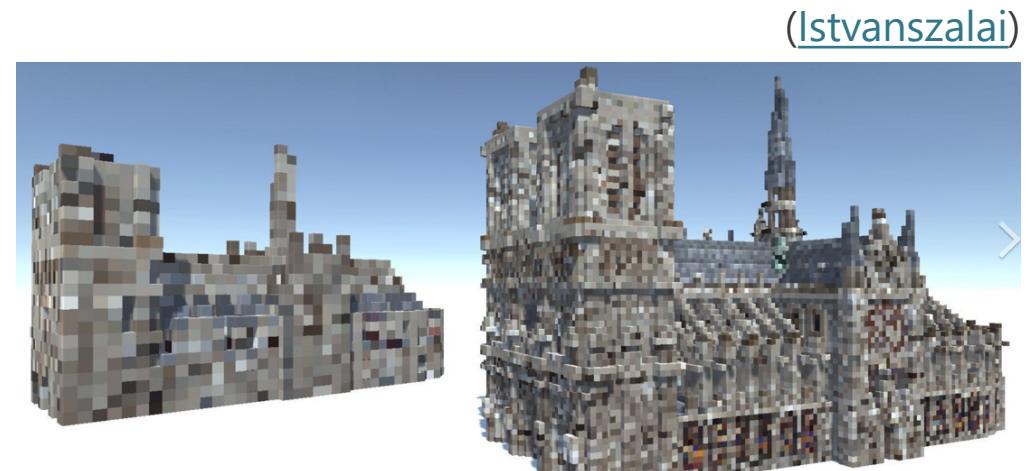
- Learnable with **small** dataset ([Dong+14](#))
 - increase in *effective* data size
- **Stability** for time evolution ([Yasuda+Onishi25](#))
 - time evolution: LR physics model
 - spatial interpolation (SR): neural net



Difficulty: Resolution-Dependent Building Shapes

■ Representation of buildings

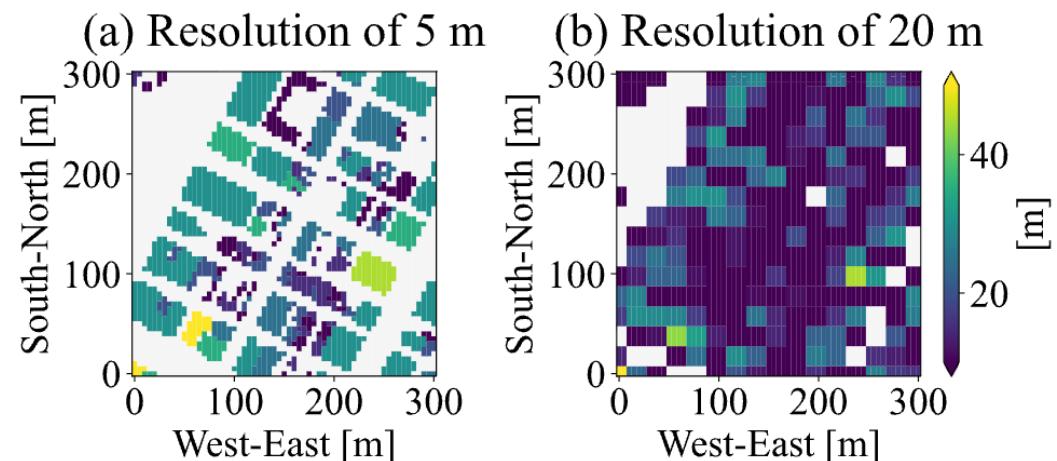
- Using **voxels** (many 3D blocks)
- Our target: flows outside buildings
- buildings acting as **obstacles**



■ Representation of narrow streets

- Cannot be represented in LR
- Must be **reconstructed** during SR

Building Height Distributions (near Tokyo St.)



Solution with Image Inpainting Techniques [\(Yasuda+23\)](#)

■ Image inpainting

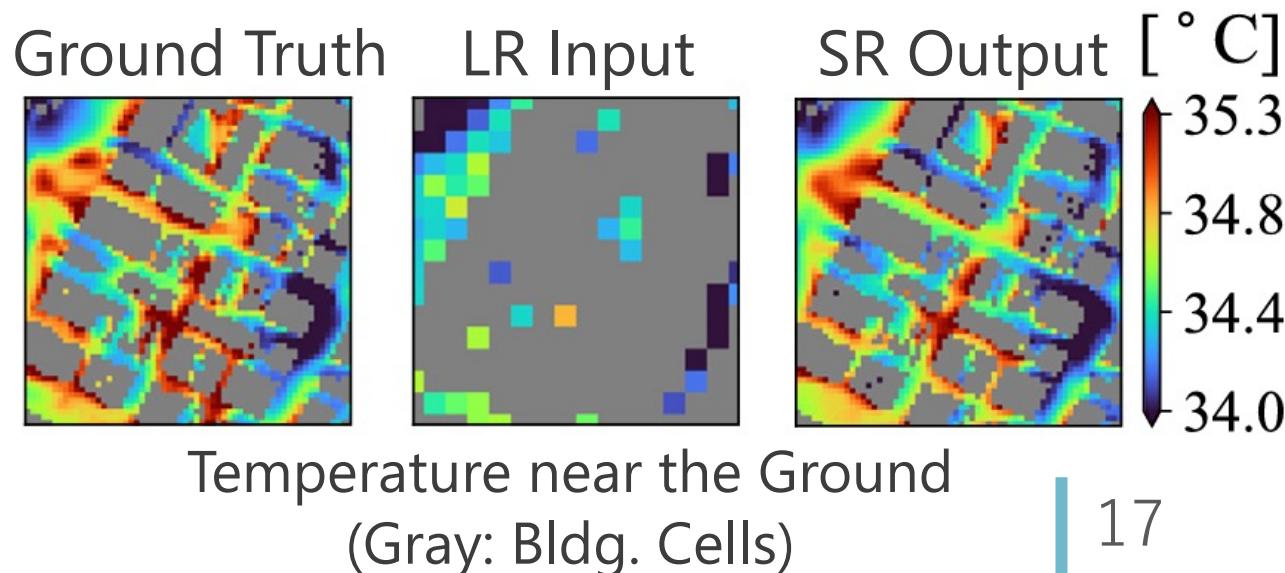
- Possible with deep learning [\(Liu+18\)](#)
 - by using surrounding information
- Applied for various apps
 - e.g., magic eraser



[\(Liu+18\)](#)

■ Application for micrometeorology

- Successfully reconstructing temperature and wind **between bldgs.** [\(Yasuda+23\)](#)
- Using **pattern matching** in neural net
 - upper flow → lower flow patterns

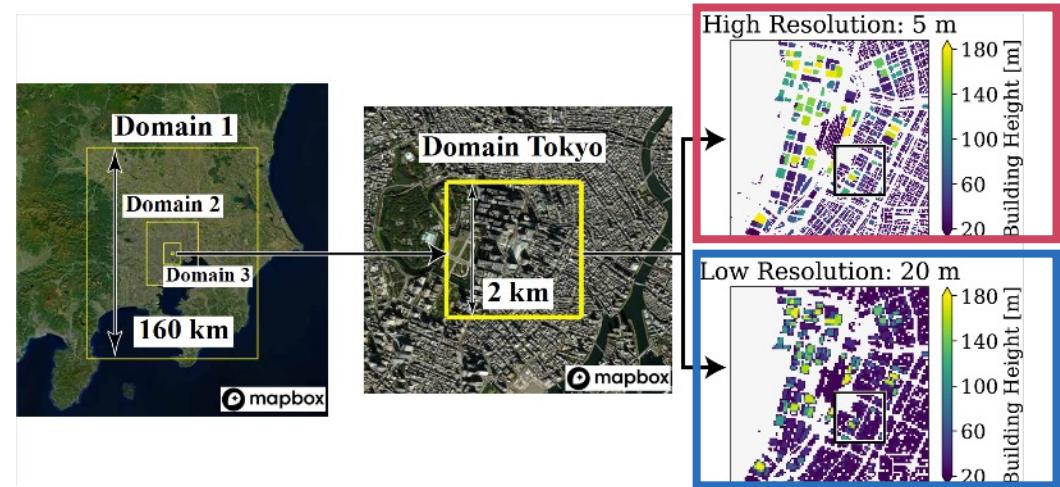


Data

■ Separate data generation of HR/LR

- from physics simulations (LESs)
 - **High resolution: 5 m** (for 3D)
 - **Low resolution: 20 m** (for 3D)

*LES: Large Eddy Simulation

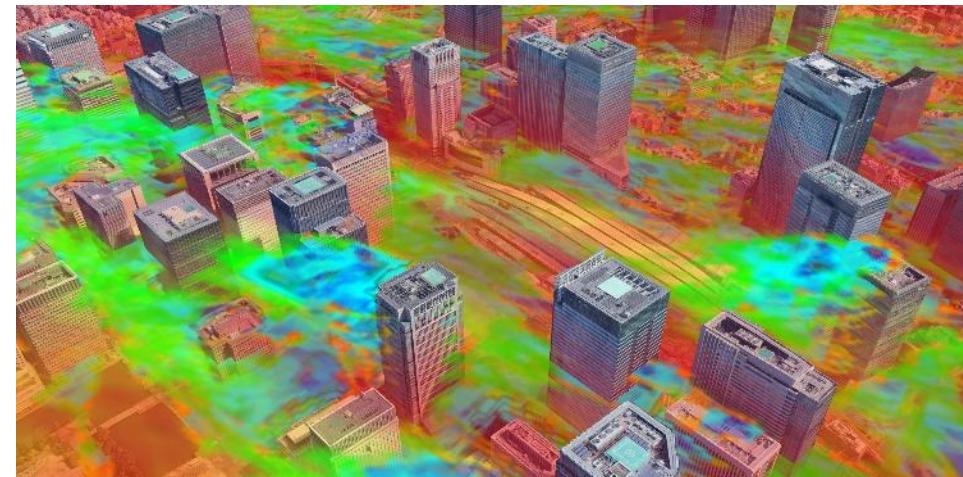


Ground
truth

Input

■ Simulation settings

- Datetimes in 2013-2020
- 1-min averaged data (3,420 sets)
 - Temperature: T
 - 3D wind velocity: $\mathbf{V} = (u, v, w)$



3D Wind Speed $\|\mathbf{V}\|$ (Tokyo Station)

Methods

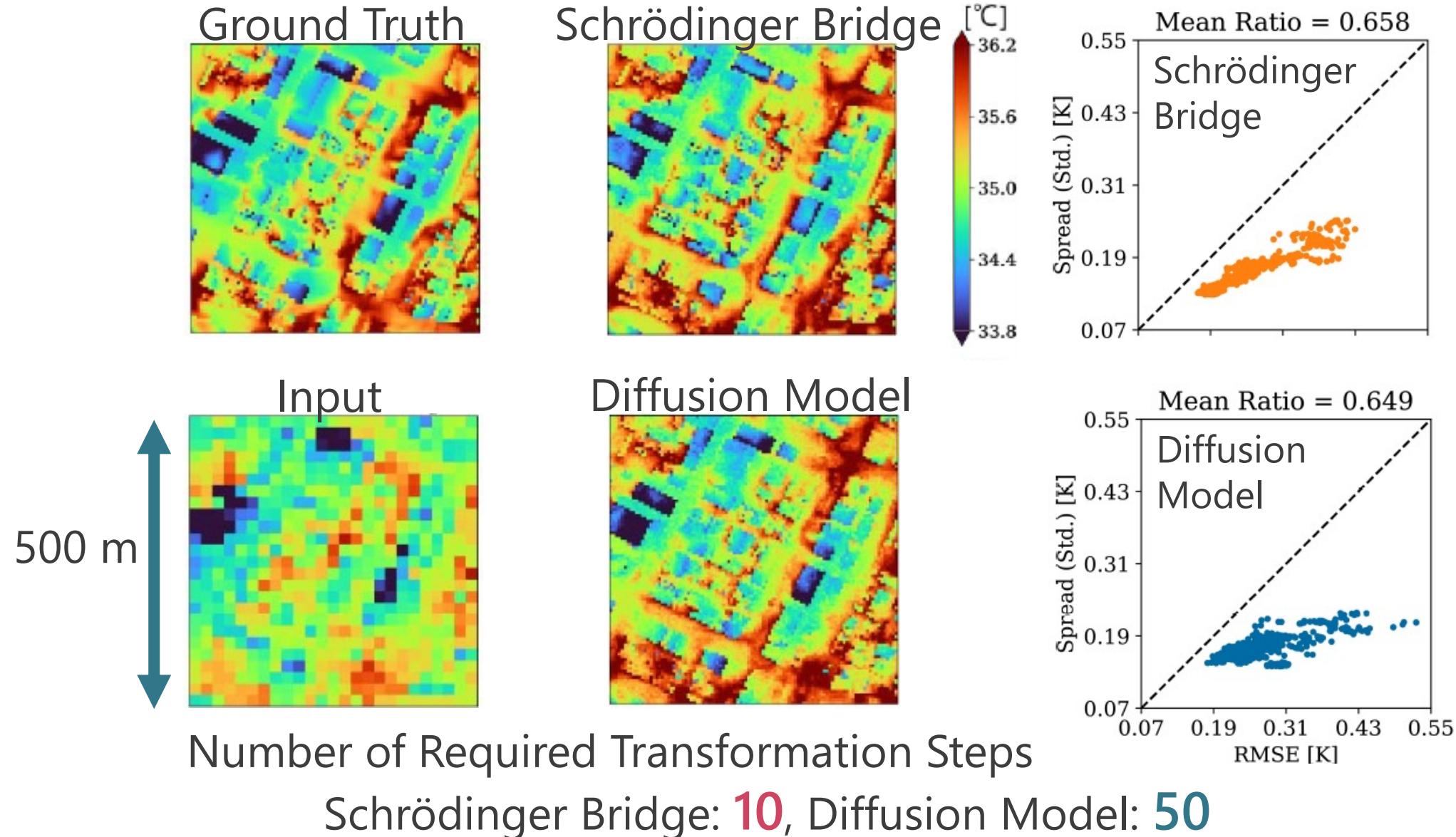
■ 2D super-resolution

- Objective: to show the **efficiency** of a Schrödinger Bridge (SB)
 - for 2m-height temperature
 - Comparing the SB and diffusion model
- Computationally **expensive**

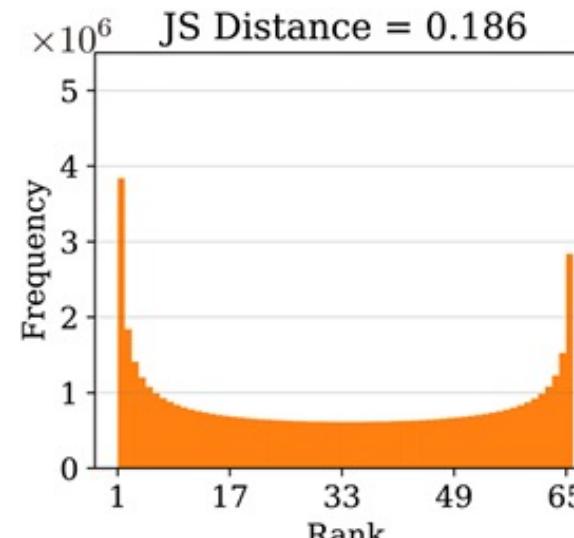
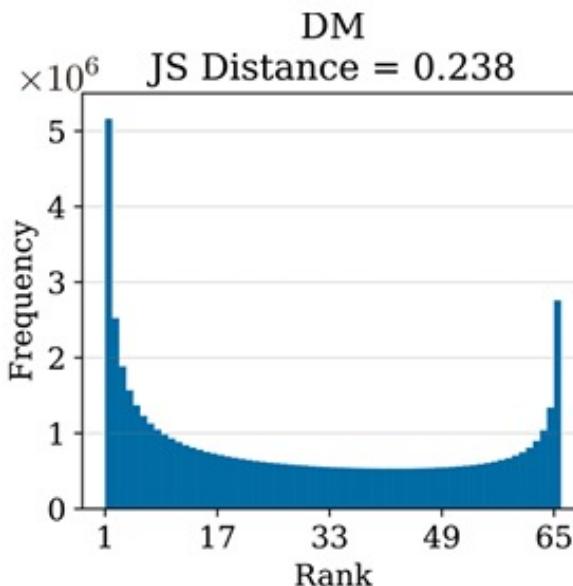
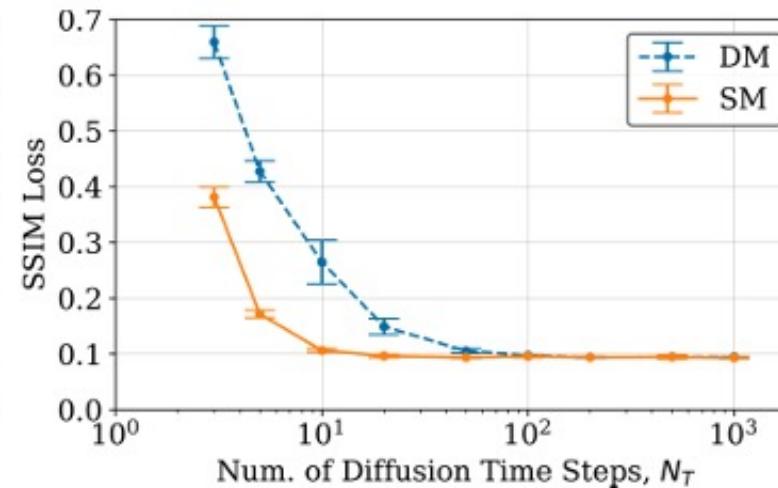
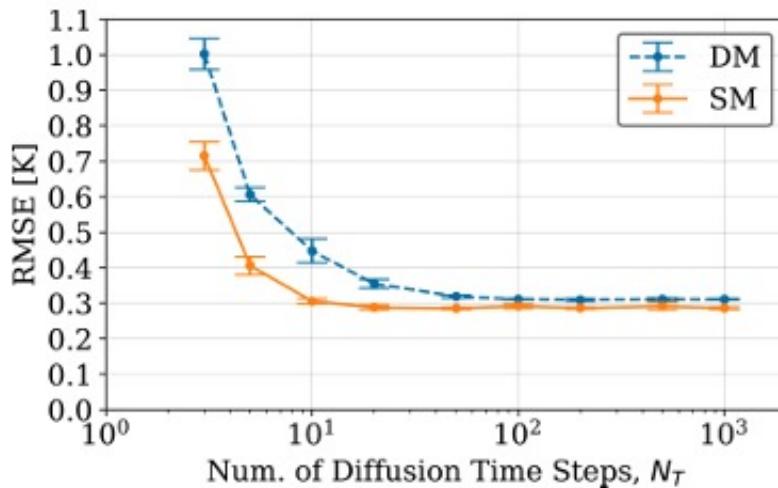
■ 3D super-resolution

- Objective: to show the **inference capability** of the SB
 - for temperature and velocity
- Extending the SB theory (Chen+24) from 2D to 3D

2D Results: Comparable Accuracy, but More Efficient



2D の追加の結果 (ref)



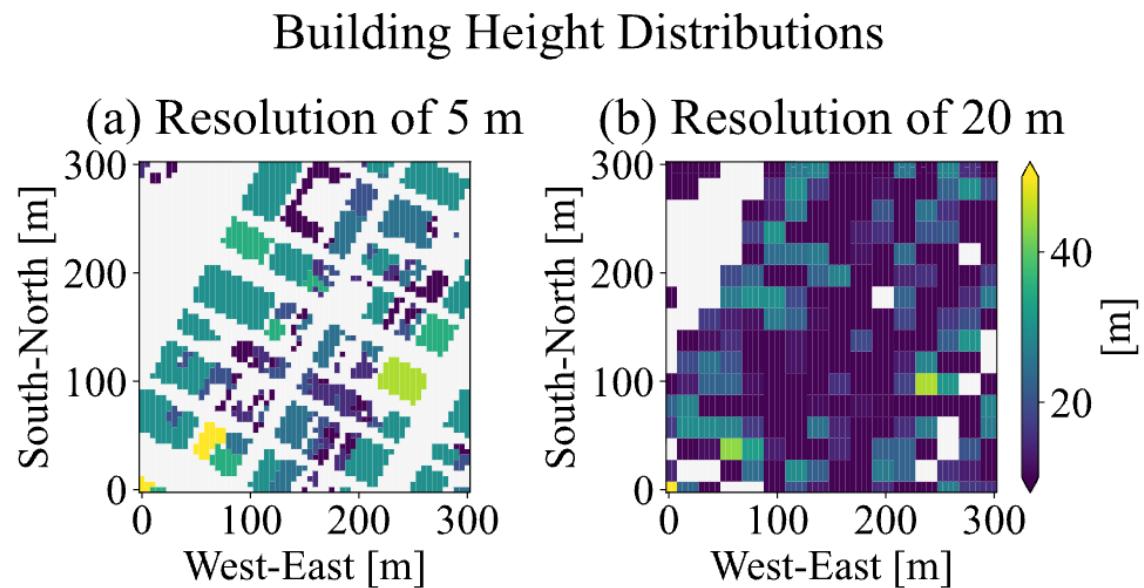
From 2D to 3D

■ To a composite task

- Super-resolution
- **Inpainting** (Yasuda+23)
- Recovering **narrow** street canyons

■ Extension of Chen+24's theory

- Using **mask** $m(\xi)$ of HR bldgs. ξ
- Showing an **optimality**.
(i.e., data transformation is
an **optimal transport**)



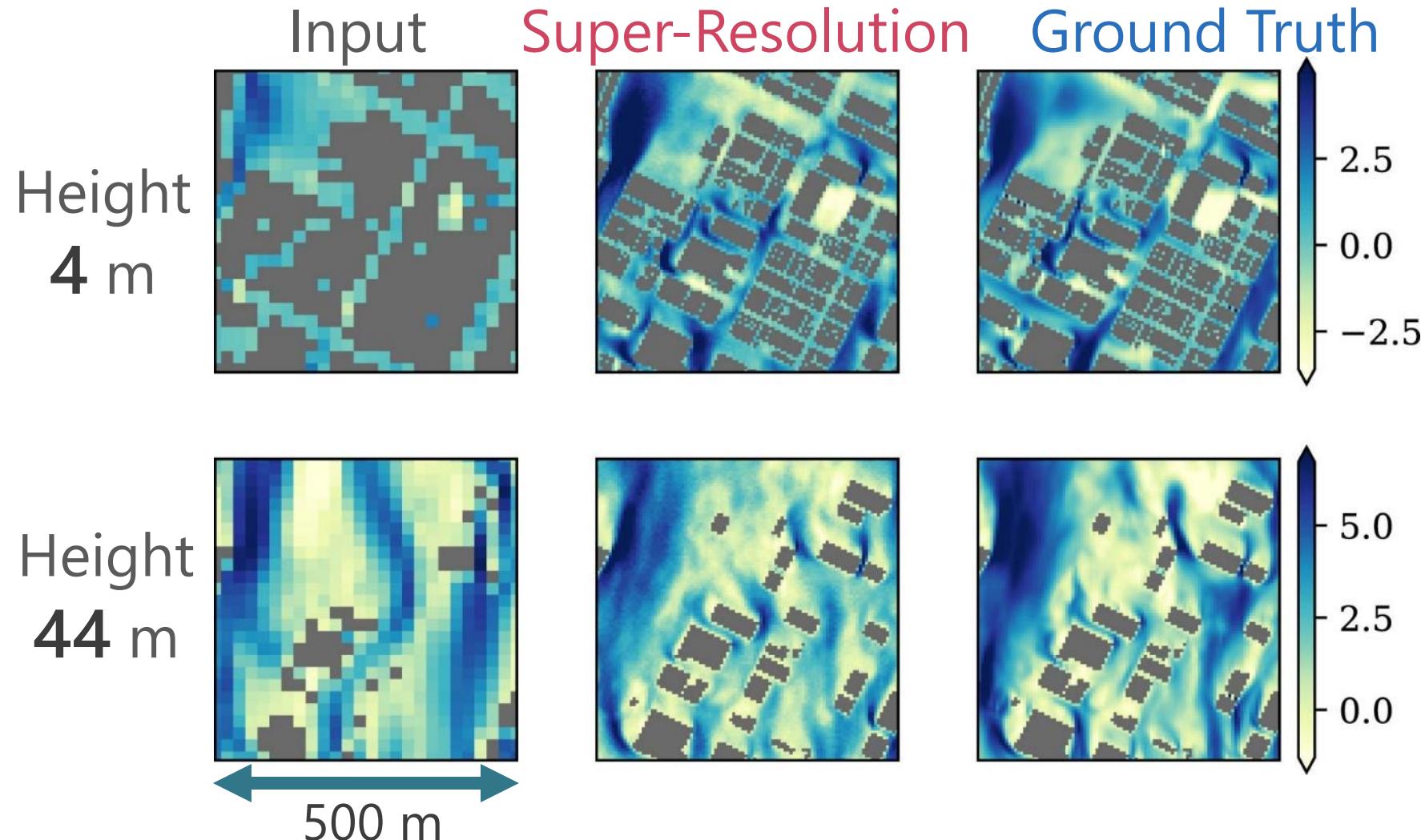
Diffusion Process in Schrödinger Bridge

$$dx = \alpha_B \partial_x \ln p(x, \xi, t) dt + m(\xi) \sqrt{2\beta_B} dW$$

α_B depends on m

No noise at bldgs.

3D Results: Northward Component of Winds v



Wall clock time **3.5** s for 1 shot of (T, u, v, w) [1.6 km \times 1.6 km \times 160 m]

まとめ

■ SDE (or ODE) を用いた生成モデルのレビュー

- 正規化流, Neural ODE → 拡散モデル → Schrödingerブリッジ
- Schrödingerブリッジ問題は 4D-Var ライクな最適化問題に帰着
- データの条件付け構造を積極的に利用する枠組みの利用

■ 微気象超解像への応用

- DM と同等の精度を $1/5$ の計算コストで実現
- 背景の数理を考えることは応用の場面でも重要