

Modelling Gas-Phase Reaction Kinetics with Guided Particle Diffusion Sampling

Andrew Millard
Zheng Zhao
Henrik Pedersen

Division of Statistics and Machine Learning (STIMA), Linköping University, Sweden
STIMA, Linköping University, Sweden
Division of Chemistry, Linköping University, Sweden

Abstract

Physics-guided sampling with diffusion priors has recently shown strong performance in solving complex systems of partial differential equations (PDEs) from sparse observations. However, these methods are typically evaluated on benchmark problems that do not fully demonstrate their ability to generate temporally consistent solutions of time-dependent PDEs, often focusing instead on reconstructing a single snapshot. In this work, we apply these methods to gas-phase reaction kinetics problems governed by the advection-reaction-diffusion (ARD) equation, providing a setting that more closely reflects realistic laboratory experiments. We demonstrate that guided sampling can be used to reconstruct full spatiotemporal trajectories, rather than isolated states. Furthermore, we show that these methods generalise to previously unseen parameter regimes, highlighting their potential for real-world applications.

Proceedings of the 2nd International Conference on Probabilistic Numerics (ProbNum) 2026, Lappeenranta, Finland. PMLR Volume 341. Copyright 2026 with the author(s)

1 Introduction and Relevant Work

Studies of chemical kinetics [Atkinson and Carter, 1984], i.e., the rates at which chemical reactions proceed, are fundamental to understanding both chemical processes and their underlying mechanisms. By quantifying how fast reactions occur and how they respond to environmental conditions, kinetic models enable the prediction and control of complex chemical systems. A well-known example is the role of reaction kinetics in explaining the depletion of the ozone layer and the formation of ozone holes over the polar regions. Chemical kinetics is also central to catalysis and has played a key role in enabling large-scale processes that support global food production. More broadly, it underpins a wide range of applications, from modelling atmospheric chemistry [Herrmann, 2003] to simulating chemical vapour deposition (CVD) processes [Danielsson et al., 2020], a widely used technique for producing thin films in electronic devices and protective coatings.

Given a set of reaction mechanisms and their constituent steps, the rate of each participating species (molecule, atom, ion, or radical) can be expressed, via kinetic rate theory, as a function of species concentrations and reaction rate constants [Atkins et al., 2023]. The resulting set of reaction rate expressions forms a system of differential equations, the solution of which describes the temporal evolution of species concentrations. Classical numerical methods for solving partial differential equations (PDEs), such as finite element methods [FEMs, Monk, 2003, Kreiss and Scherer, 1974], finite difference methods [FDMs, LeVeque, 2007] and spectral methods [Trefethen, 1996] provide reliable approximations to these systems, but can become computationally expensive as the dynamics grow more complex or the dimensionality increases. Therefore, application of these methods to environments requiring high-resolution simulations or studies requiring exploration of large parameter spaces

can become prohibitively expensive.

This motivates the exploration of machine learning methods as potential surrogates for classical numerical solvers. Machine learning approaches have been used to model reaction rates [Meuwly, 2021], photochemical reactions Staszak [2023], and reaction networks [Stocker et al., 2020]. However, these approaches typically focus on modelling specific components of gas-phase chemistry, rather than providing a holistic simulation of the full chemical reaction process. End-to-end machine learning approaches have, however, seen significant success in other domains such as structural biology [Jumper et al., 2021, Abramson et al., 2024], where models take an amino acid sequence as input and directly predict the corresponding three-dimensional protein structure. The success of such approaches, particularly the inclusion of generative components to model complex, high-dimensional structures, as incorporated in Abramson et al. [2024], suggests that similar techniques may be promising for end-to-end modelling of PDE-governed systems.

Recent advances in deep generative models, including diffusion models [Erichson et al., 2025, Yang and Sommer, 2023], have demonstrated strong capability in approximating the spatiotemporal structure of complex PDE systems. Furthermore, guided sampling approaches [Huang et al., 2024, Millard et al., 2026, Yao et al., 2025, Jacobsen et al., 2025] incorporate physical constraints such as the governing PDE, directly into the sampling process by incorporating it as part of a likelihood term, thereby steering generated solutions toward physically consistent states. However, most existing work in this area focuses on idealised benchmarks Millard et al. [2026], Huang et al. [2024], Yao et al. [2025]. It is yet unclear whether they can apply to practical and realistic laboratory scenarios, and this is our main focus here.

Our contributions to this field are threefold:

- We present a series of experiments designed to evaluate whether guided sampling techniques generalise effectively to chemistry scenarios by evaluating our methods on simulations closer to real-world scenarios.
- We demonstrate that guided diffusion-based methods remain effective under these more complex and physically relevant conditions, and generalise well to previously unseen parameter regimes.
- We show that guided sampling from diffusion priors can capture complex chemical reaction dynamics across a broad range of temporal scales and accurately reconstruct spatiotemporal chemical species concentration fields from sparse observations.

The remainder of this paper adopts a mathematical structure similar to that of Millard et al. [2026], with closely aligned notation and presentation.

2 Background

We begin by a generic class of time-dependent PDEs and then in this section we explain how a generative diffusion model can be applied to numerically solve them.

For a spatial domain Ω and time horizon $[0, T]$, a general time-dependent PDE can be written as

$$\begin{aligned} f(u, c, \tau, a) &= 0, \quad c \in \Omega, \quad \tau \in [0, T], \\ u(c, \tau) &= g(c, \tau), \quad c \in \partial\Omega, \quad \tau \in [0, T], \\ u(c, 0) &= a(c), \end{aligned} \quad (1)$$

where c denotes the spatial coordinate, τ denotes physical (PDE) time (distinct from diffusion time t), and $u \in \mathcal{U}$ is the PDE solution field. The function $a \in \mathcal{A}$ specifies the initial condition (or, more generally, fixed PDE parameters), and g denotes the boundary conditions.

Throughout this work, the abstract solution field $u(c, \tau)$ represents a spatiotemporal physical field. In later sections 3 we denote this field as a concentration field $\mathcal{C}(c, \tau)$ governed by a reaction-transport PDE. The notation $\mathcal{C}(\cdot, \tau)$ refers to the spatial concentration field over Ω at a fixed PDE time τ .

Sampling for Generative Diffusion Models

Generative diffusion models [Sohl-Dickstein et al., 2015, Song et al., 2021, Ho et al., 2020, Cao et al., 2024, Karras et al., 2022] provide a probabilistic framework for sampling from complex, high-dimensional data distributions $p_{\text{data}}(x)$.

The forward process progressively corrupts $p_{\text{data}}(x)$ with Gaussian noise according to a predetermined noise schedule $\sigma_t \in [0, \sigma_{\text{max}}]$ [Song et al., 2021, Karras et al., 2022, Ho et al., 2020], $t \in [0, T]$, yielding a reference distribution $p(x_T) \approx \mathcal{N}(0, \sigma_{\text{max}}^2 I)$. Sampling proceeds by drawing x_T from this reference distribution and reversing the noising process to obtain $x_0 \sim p_{\text{data}}(x)$.

The probability flow ordinary differential equation (ODE) describes how a sample drawn from the reference distribution can be gradually denoised in reverse time [Karras

et al., 2022, Song et al., 2021] in order to recover a sample from the data distribution:

$$dx_t = -\dot{\sigma}_t \sigma_t \nabla_{x_t} \log p_t(x_t, \sigma_t) dt, \quad t \in [0, T], \quad (2)$$

where $\log p_t(x_t, \sigma_t)$ is the score function. Karras et al. [2022] propose learning a denoising function $\delta_\theta(x_t, \sigma_t)$ so that the score function can be estimated as

$$\nabla_{x_t} \log p_t(x_t, \sigma_t) \approx \nabla_{x_t} \log p_t^\theta(x_t, \sigma_t) = \frac{x_t - \delta_\theta(x_t, \sigma_t)}{\sigma_t^2}, \quad (3)$$

where δ_θ is a neural network trained to denoise the sample at a given noise level σ_t . This reverse-time denoising process can also be simulated using a corresponding stochastic differential equation (SDE) which recovers the same marginal distributions as Equation (2) [Song et al., 2021]:

$$dx_t = -2\dot{\sigma}_t \sigma_t \nabla_x \log p_t(x_t, \sigma_t) dt + \sqrt{2\dot{\sigma}_t \sigma_t} dW_t, \quad (4)$$

where W_t is a Brownian motion and $\dot{\sigma}_t$ denotes the time derivative of the noise schedule.

Guided Sampling

We consider the setting in which physical parameters are fixed and known, and a neural network is trained over realisations of the PDE solution field. This neural network represents the diffusion model prior $p_\theta(x)$, and diffusion samples therefore take the form

$$x \in \mathcal{X} = \mathcal{U}, \quad (5)$$

where x represents a discretised realisation of the PDE solution field, which in this work corresponds to a concentration field at a single physical time. Sampling from the diffusion model by solving Equation (2) or (4) corresponds to sampling from the prior distribution without incorporating additional information, such as the governing PDE equation or sparse observations.

To incorporate such information, we condition the diffusion process on auxiliary data y . In this case, we aim to sample from the posterior distribution

$$p_\theta(x | y) \propto p(y | x) p_\theta(x). \quad (6)$$

When using the reverse-time SDE, sampling from this distribution can be achieved by modifying Equation (4) as

$$\begin{aligned} dx_t &= -2\dot{\sigma}_t \sigma_t \nabla_{x_t} \log p_t^\theta(x_t, \sigma_t) dt \\ &\quad - 2\dot{\sigma}_t \sigma_t \nabla_{x_t} \log p_t^\theta(y | x_t, \sigma_t) dt + \sqrt{2\dot{\sigma}_t \sigma_t} dW_t, \end{aligned} \quad (7)$$

where the additional gradient term is known as the *guidance term*. This term encourages samples to be consistent with the conditional information, which in our setting includes sparse observations and the governing PDE dynamics.

Sequential Monte Carlo

Sequential Monte Carlo (SMC) methods [e.g., Chopin and Papaspiliopoulos, 2020, Naesseth et al., 2019] provide a principled method for sampling from a sequence of distributions and therefore, are well suited for the

sequential denoising process associated with generative diffusion models. Previous work explores the use of SMC as a framework for conditional sampling of diffusion priors [Wu et al., 2023, Cardoso et al., 2024, Stevens et al., 2025, Dou and Song, 2024, Zhao et al., 2025] for problems such as inpainting [Kelvinius et al., 2025], motif-scaffolding for proteins [Trippe et al., 2023] and many others. In previous subsections, we have used t when discussing continuous time processes. In the following sections, we discuss simulating these processes in discrete time and therefore denote this discrete time as k .

SMC simulates propagating an ensemble of N particles $\{x_k^{(i)}\}_{i=1}^N$ across $k = 0, 1, \dots, K$ iterations/time steps. In this setting, the particles are first sampled from the diffusion prior M_K and weighted according to an initial potential G_K :

$$x_K^{(i)} \sim M_K = p(x_K) \approx \mathcal{N}(0, I), \quad w_K^{(i)} \propto G_K(x_K^{(i)}). \quad (8)$$

For K iterations, samples are propagated via a Markov kernel/proposal distribution M_{k-1}

$$x_{k-1}^{(i)} \sim M_{k-1}(x_{k-1}^{(i)} | x_k^{(i)}), \quad (9)$$

and reweighted

$$w_{k-1}^{(i)} \propto w_k^{(i)} G_{k-1}(x_k^{(i)}, x_{k-1}^{(i)}). \quad (10)$$

We resample [Douc and Cappé, 2005] the particles if our effective sample size (ESS) drops below a certain threshold N_{eff} . At each iteration k , we have a weighted ensemble $\{(x_k^{(i)}, w_k^{(i)})\}_{i=1}^N$ that provides an approximation of the target distribution ν

$$\begin{aligned} \nu_k(x_{k:K}) &\propto G_K(x_K) M_K(x_K) \\ &\times \prod_{j=k+1}^K G_{j-1}(x_j, x_{j-1}) M_{j-1}(x_{j-1} | x_j). \end{aligned} \quad (11)$$

Equation (11) is known as the Feynman-Kac (FK) formula and describes the time evolution of stochastic processes. We use this SMC framework in order to sample from the posterior distribution given by Equation (6). To do this, we choose proposals and potential functions $\{M_k, G_k\}_{k=0}^K$ such that the final time *marginal* distribution is the conditional target distribution i.e. $\nu_0(x_0) = p_\theta(x_0 | y)$. Appendix A gives the pseudocode for the SMC framework.

Stochastic Sampling

We approximate the reverse process described by Equation (4) using a guided Euler-Maruyama (GEM) update [Millard et al., 2026]

$$\begin{aligned} x_{k-1} &= x_k + (\sigma_{k-1}^2 - \sigma_k^2) \frac{x_k - \delta_\theta(x_k, \sigma_k)}{\sigma_k^2} \\ &+ (\sigma_{k-1}^2 - \sigma_k^2) \nabla_{x_k} \log \tilde{p}_\theta(y | x_k) + \sqrt{\sigma_{k-1}^2 - \sigma_k^2} \zeta, \end{aligned} \quad (12)$$

with $\zeta \sim \mathcal{N}(0, I)$ and $\tilde{p}_\theta$ denotes a likelihood function whose formulation is detailed in later sections. The time interval Δt is implicitly accounted for in the noise

scheduler. The pseudocode for this can be found in Appendix B, Algorithm 2.

Twisted SMC

The GEM update in Equation (12) gives a Markov kernel $\tilde{p}_\theta(x_{k-1} | x_k, y) = \mathcal{N}(x_k | \mu_{\text{prop}}, (\sigma_{k-1}^2 - \sigma_k^2) I)$ which can be used to propagate samples within SMC. If we define our proposal distribution $M_{k-1}(x_{k-1} | x_k) = \tilde{p}_\theta(x_{k-1} | x_k, y)$ and the corresponding weighting function as

$$G_{k-1}(x_k, x_{k-1}) = \frac{\tilde{p}_\theta(y | x_{k-1})}{\tilde{p}_\theta(y | x_k)} \frac{p_\theta(x_{k-1} | x_k)}{\tilde{p}_\theta(x_{k-1} | x_k, y)}, \quad (13)$$

then this corresponds to the SMC framework used in the Twisted Diffusion Sampler [TDS, Wu et al., 2023]. We can verify that these choices of potentials and proposals ensure that we target the correct final time marginal by substituting them into the FK formula given by Equation (11):

$$\begin{aligned} \nu_{k-1}(x_{k-1:K}) &\propto \tilde{p}_\theta(y | x_{k-1}) p_\theta(x_{k-1:K}) \\ \nu_{0:K}(x_{0:K}) &= p_\theta(x_{0:K} | y). \end{aligned} \quad (14)$$

Second-Order Stochastic Sampling

Karras et al. [2022] propose using second-order stochastic sampling process. First, the noise scale is jittered and then the ODE dynamics via the denoiser are run, i.e. we run the following set of equations:

$$\hat{\sigma}_k \leftarrow \sigma_k + \gamma_k \sigma_k, \quad (15)$$

$$\hat{x}_k = x_k + \sqrt{\hat{\sigma}_k^2 - \sigma_k^2} \psi, \quad \psi \sim \mathcal{N}(0, I), \quad (16)$$

$$x_{k-1} = \hat{x}_k + (\sigma_{k-1}^2 - \hat{\sigma}_k^2) \frac{\hat{x}_k - \delta_\theta(\hat{x}_k, \hat{\sigma}_k)}{\hat{\sigma}_k^2}, \quad (17)$$

where γ_k is a constant used to reach a higher noise level and a second order correction is used to improve sampling quality. The full details of this can be found in Karras et al. [2022]. Millard et al. [2026] propose including guidance information after the second order correction such that:

$$\begin{aligned} x_{k-1} &= \hat{x}_k + (\sigma_{k-1}^2 - \hat{\sigma}_k^2) \frac{\hat{x}_k - \delta_\theta(\hat{x}_k, \hat{\sigma}_k)}{\hat{\sigma}_k^2} \\ &- \alpha \nabla_{\hat{x}_k} \log \tilde{p}_\theta(y | \hat{x}_k), \end{aligned} \quad (18)$$

where α is a guidance weighting parameter. The resulting algorithm is called the Second-Order Stochastic Guided (SOSaG) update. The pseudocode for this algorithm is given in Appendix B, Algorithm 3. Further details of this can be found in Millard et al. [2026].

pBS SMC Framework

Millard et al. [2026] use the SOSaG proposal within a different SMC framework, the so-called pseudo-bootstrap (pBS) method.

It is not obvious how to incorporate SOSaG as a proposal distribution within the TDS framework as the particles generated by it are not point-evaluable proposal density. The particles are jittered before the denoising process and therefore are not Gaussian distributed like the samples generated by GEM. Instead, they define a new FK

model $\bar{\nu}_{0:K}$ using a similar structure to the bootstrap particle filter [Gordon et al., 1993]

$$\begin{aligned}\bar{M}_{k-1}(x_{k-1} | x_k) &= \bar{p}_\theta(x_{k-1} | x_k, y), \\ \bar{G}_{k-1}(x_k, x_{k-1}) &= \frac{\tilde{p}_\theta(y | x_{k-1})}{\tilde{p}_\theta(y | x_k)},\end{aligned}$$

where $\bar{p}_\theta(x_{k-1} | x_k, y)$ is the proposal distribution of the SOSaG method given in Equation (18). This new FK model recovers a target distribution $\bar{\nu}_0(x_0) = \bar{p}_\theta(x_0 | y) p_\theta(y | x_0)$ at $k = 0$. Although this does not target the same distribution as Equation (14), it has been found to give more physically consistent results [Millard et al., 2026].

3 Methodology

We consider transport-reaction systems governed by advection-reaction-diffusion (ARD) equations [Rubio et al., 2008, Cosner, 2014]. In this setting, a diffusion model sample (at diffusion time $k = 0$)

$$x \equiv \mathcal{C}(\cdot, \tau_n)$$

represents a discretised concentration field at a given physical time $\tau_n \in [0, \mathcal{T}]$, corresponding to the solution of the underlying PDE $u(\cdot, \tau_n)$ introduced in Section 2. In this work, $\mathcal{C}$ denotes the vector of S species concentrations, given by

$$\mathcal{C} = (\mathcal{C}_1, \dots, \mathcal{C}_S).$$

We focus on an axisymmetric cylindrical reactor of length L and radius A (see, Figure 2). The concentration of species s , denoted by $\mathcal{C}_s(r, z, \tau)$, evolves according to

$$\frac{\partial \mathcal{C}_s}{\partial \tau} + u(r) \frac{\partial \mathcal{C}_s}{\partial z} = D \nabla^2 \mathcal{C}_s + R_s(\mathcal{C}), \quad (19)$$

where $r \in [0, A]$ and $z \in [0, L]$ denote the radial and axial coordinates, respectively. Here, D is the molecular diffusivity and $R_s(\mathcal{C})$ represents the reaction source term [Robert J. Kee, 2003].

The axial velocity field u is assumed to follow a laminar parabolic profile,

$$u(r) = 2U_{\text{avg}} \left(1 - \left(\frac{r}{A} \right)^2 \right),$$

where U_{avg} denotes the cross-sectional average velocity, which is taken to be identical across all species.

In our experiments, all physical parameters of the ARD system are assumed to be known and fixed. Consequently, the inference task is restricted to recovering the spatiotemporal concentration fields $\{\mathcal{C}_s\}_{s=1}^S$.

ARD Residual-Based Likelihood for Time-Series Reconstruction

A sequence of concentration fields is reconstructed by combining stochastic sampling proposals with the corresponding SMC framework [see, e.g., Wu et al., 2023, Zhao et al., 2025] with a diffusion prior applied independently at each physical time step τ_n . For each τ_n , the

diffusion process is initialised from Gaussian noise and an SMC-guided denoising procedure is used to obtain a reconstruction of $\mathcal{C}(\cdot, \tau_n)$. During this process, candidate samples are evaluated according to a likelihood that enforces both agreement with sparse observations and consistency with the underlying ARD dynamics.

The likelihood is defined through a combination of observation and PDE residual terms. Specifically, the log-likelihood is given by

$$\begin{aligned}\log p(y | x) &= -\frac{1}{\sigma_{\text{obs}}} \frac{1}{N_{\text{obs}}} \|\mathcal{C}_{\text{obs}} - \mathcal{M} \odot x\|_2^2 \\ &\quad - \frac{1}{\sigma_{\text{pde}}} \frac{1}{AL} \|\mathcal{R}_n(\mathcal{C}; \mathcal{C}^{\tau_{n-1}})\|_2^2 + \text{const.}, \quad (20)\end{aligned}$$

where $\mathcal{M}$ is a binary mask indicating observed locations, $\mathcal{C}$ denotes the reconstruction of the physical concentration field corresponding to x , $\mathcal{C}_{\text{obs}}$ represents sparse observations of the ground truth field and $\mathcal{R}_n$ is the PDE residual which is defined later in this section. The parameters σ_{obs} and σ_{pde} control the relative weighting of the observation and PDE terms, while N_{obs} denotes the number of observations.

Following Wu et al. [2023], the intermediate diffusion likelihood is approximated by evaluating the likelihood at the denoised estimate,

$$p_t^\theta(y | x_t, \sigma_t) \approx p(y | x_0 = \delta_\theta(x_t, \sigma_t)) =: \tilde{p}_\theta(y | x_t).$$

To quantify consistency with the governing dynamics, we introduce the spatial ARD operator

$$\mathcal{F}(\mathcal{C}) = u(r) \frac{\partial \mathcal{C}}{\partial z} - D \left(\frac{\partial^2 \mathcal{C}}{\partial z^2} + \frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial \mathcal{C}}{\partial r} \right) \right) - R(\mathcal{C}), \quad (21)$$

where $R(\mathcal{C})$ denotes the nonlinear reaction source term applied component-wise to the concentration vector $\mathcal{C}$. $R(\mathcal{C})$ is derived from the corresponding chemical reaction mechanism and kinetic rate laws.

The PDE residual at discrete time points $\{\tau_n\}$, can be for example modelled by a first-order finite difference

$$\begin{aligned}\mathcal{R}_n(\mathcal{C}^{\tau_n}; \mathcal{C}^{\tau_{n-1}}) &= \frac{\mathcal{C}^{\tau_n} - \mathcal{C}^{\tau_{n-1}}}{\Delta \tau_n} + \mathcal{F}(\mathcal{C}^{\tau_n}), \\ \Delta \tau_n &:= \tau_n - \tau_{n-1}, \quad n \geq 1. \quad (22)\end{aligned}$$

Here, $\mathcal{C}^{\tau_{n-1}}$ denotes the reconstructed concentration field at the previous time step. At the initial time $\tau_0 = 0$, the concentration field is set to $\mathcal{C}^{\text{init}} = 0$, corresponding to an initially empty reactor.

4 Experimental Setup and Results

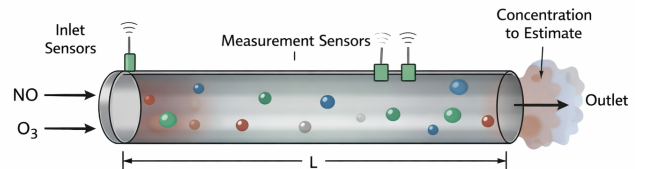


Figure 2: Diagram of the experimental set up being simulated.

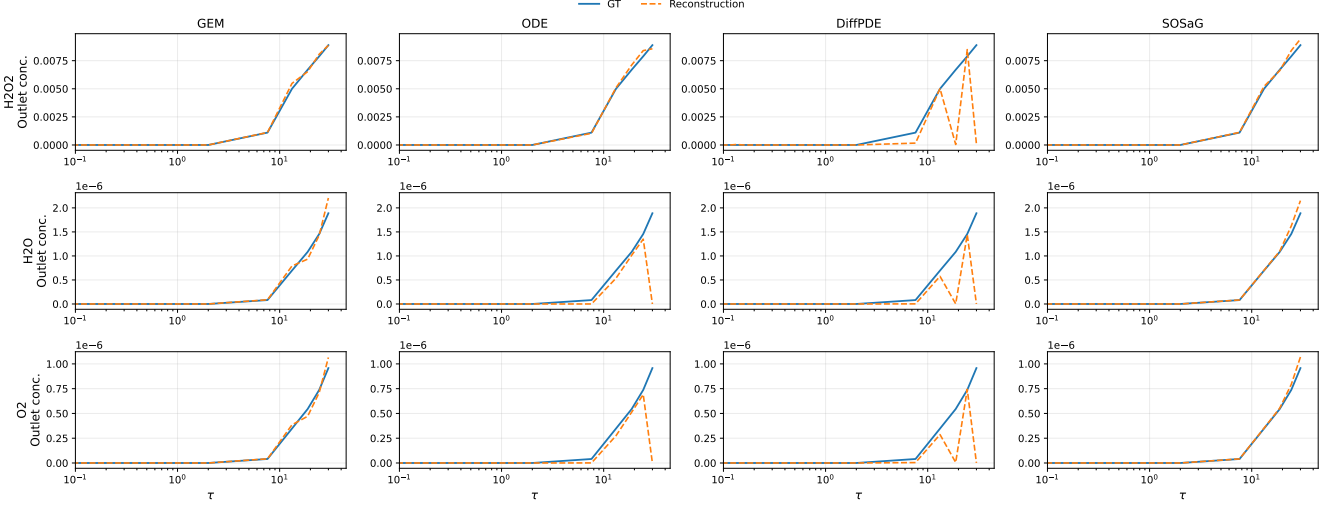


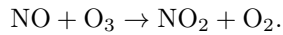
Figure 1: Species outlet comparison for H_2O_2 Decomposition.

The aim of these experiments is to simulate gas-phase reaction kinetics within a cylindrical tube reactor. Inlet gases enter from the left-hand side of the reactor, where they mix and react as they flow downstream toward the outlet on the right. Along the length of the reactor, sparse sensors provide local measurements of species concentrations. A schematic of this setup is shown in Figure 2.

The objective is twofold: first, to reconstruct the spatial distribution of chemical species concentrations throughout the reactor $\{\mathcal{C}_s(r, z, \tau)\}_{s=1}^S$, and second, to predict the temporal evolution of species concentrations at the reactor outlet $\{\mathcal{C}_s(r, L, \tau_n)\}_{s=1, \dots, S}^{n=0, \dots, T}$. We compare against the first and second order ODE and stochastic methods mentioned previously. Both stochastic methods are used within an SMC framework.

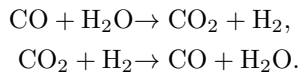
We consider six reduced reaction mechanisms spanning atmospheric chemistry, catalytic reaction systems, and simplified combustion kinetics.

NO- O_3 to NO_2 . This mechanism represents a minimal atmospheric reaction system describing ozone consumption by nitric oxide:

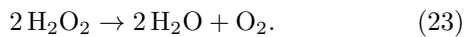


It provides a simple benchmark for nonlinear reactive transport with a single irreversible reaction.

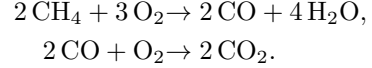
Water-gas shift. This mechanism models the reversible water-gas shift reaction:



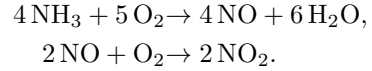
Hydrogen peroxide decomposition. This mechanism describes the decomposition of hydrogen peroxide into water and oxygen:



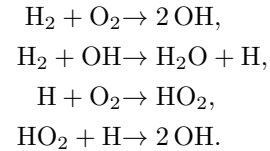
Two-step methane oxidation. This mechanism approximates methane combustion using a reduced two-step description:



Ammonia oxidation. This mechanism models a reduced form of ammonia oxidation:



Hydrogen oxidation subset. This mechanism is a reduced radical subset of hydrogen combustion chemistry:



Reaction Rates

A crucial coefficient in our ARD equation is the reaction rate constants. They are computed using Arrhenius kinetics,

$$\kappa_l(\mathcal{K}) = A_l \exp\left(-\frac{E_{a,l}}{R_{\text{gas}}\mathcal{K}}\right), \quad (24)$$

where $\mathcal{K}$ is the sampled temperature, $\kappa_l(\mathcal{K})$ is the temperature dependent rate constant for reaction l , A_l is the pre-exponential factor, $E_{a,l}$ is the activation energy, and R_{gas} is the universal gas constant.

Given a vector of species concentrations $\mathcal{C} = (\mathcal{C}_1, \dots, \mathcal{C}_S)$, the mass-action kinetics defines the reaction rate of reaction j as

$$\chi_l(\mathcal{C}) = \kappa_l(\mathcal{K}) \prod_{i=1}^S \mathcal{C}_i^{\alpha_{il}},$$

where α_{il} denotes the stoichiometric coefficient of species i as a reactant in reaction l . The nonlinear reaction

Table 1: Gas-phase reaction kinetics full-field and outlet absolute errors across all experiments and methods (non-transformed / physical space). We give the mean $\pm$ one standard deviation, the best results are bolded.

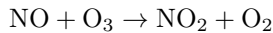
Method	RMSE	MAE	Outlet RMSE	Outlet MAE	Time (s)
H₂O₂ Decomposition					
GEM	$(1.49 \pm 0.58) \times 10^{-4}$	$(3.19 \pm 1.57) \times 10^{-5}$	$(1.22 \pm 0.39) \times 10^{-4}$	$(2.39 \pm 1.02) \times 10^{-5}$	$(1.55 \pm 0.01) \times 10^3$
SOSaG	$(1.47 \pm 0.57) \times 10^{-4}$	$(3.10 \pm 1.49) \times 10^{-5}$	$(1.09 \pm 0.32) \times 10^{-4}$	$(2.17 \pm 0.81) \times 10^{-5}$	$(1.55 \pm 0.01) \times 10^3$
ODE	$(3.29 \pm 1.66) \times 10^{-3}$	$(5.30 \pm 3.38) \times 10^{-4}$	$(2.28 \pm 1.20) \times 10^{-3}$	$(3.98 \pm 2.48) \times 10^{-4}$	$(1.46 \pm 0.02) \times 10^3$
DiffPDE	$(3.09 \pm 1.43) \times 10^{-3}$	$(5.42 \pm 3.05) \times 10^{-4}$	$(2.35 \pm 1.18) \times 10^{-3}$	$(4.05 \pm 2.62) \times 10^{-4}$	$(2.12 \pm 0.09) \times 10^3$
NO + O₃ → NO₂					
GEM	$(2.11 \pm 1.41) \times 10^{-4}$	$(6.04 \pm 4.50) \times 10^{-5}$	$(1.49 \pm 1.03) \times 10^{-4}$	$(4.19 \pm 3.45) \times 10^{-5}$	$(1.59 \pm 0.00) \times 10^3$
SOSaG	$(2.03 \pm 1.39) \times 10^{-4}$	$(5.87 \pm 4.40) \times 10^{-5}$	$(1.57 \pm 1.26) \times 10^{-4}$	$(4.77 \pm 4.61) \times 10^{-5}$	$(2.36 \pm 0.01) \times 10^3$
ODE	$(6.67 \pm 4.11) \times 10^{-3}$	$(1.18 \pm 0.95) \times 10^{-3}$	$(2.78 \pm 2.50) \times 10^{-3}$	$(6.48 \pm 5.96) \times 10^{-4}$	$(1.39 \pm 0.01) \times 10^3$
DiffPDE	$(3.32 \pm 2.62) \times 10^{-3}$	$(9.07 \pm 7.77) \times 10^{-4}$	$(1.73 \pm 1.36) \times 10^{-3}$	$(5.10 \pm 4.21) \times 10^{-4}$	$(2.07 \pm 0.01) \times 10^3$
Ammonia Oxidation					
GEM	$(3.48 \pm 2.17) \times 10^{-4}$	$(7.84 \pm 5.93) \times 10^{-5}$	$(4.14 \pm 2.97) \times 10^{-4}$	$(8.61 \pm 7.01) \times 10^{-5}$	$(1.58 \pm 0.00) \times 10^3$
SOSaG	$(3.88 \pm 2.44) \times 10^{-4}$	$(9.39 \pm 7.39) \times 10^{-5}$	$(4.35 \pm 3.14) \times 10^{-4}$	$(9.67 \pm 8.57) \times 10^{-5}$	$(2.40 \pm 0.01) \times 10^3$
ODE	$(3.77 \pm 2.63) \times 10^{-3}$	$(1.01 \pm 0.77) \times 10^{-3}$	$(2.94 \pm 2.21) \times 10^{-3}$	$(6.98 \pm 5.47) \times 10^{-4}$	$(1.43 \pm 0.00) \times 10^3$
DiffPDE	$(3.77 \pm 2.65) \times 10^{-3}$	$(9.91 \pm 7.57) \times 10^{-4}$	$(2.82 \pm 2.09) \times 10^{-3}$	$(6.60 \pm 5.10) \times 10^{-4}$	$(2.10 \pm 0.01) \times 10^3$
Hydrogen Oxidation Subset					
GEM	$(2.51 \pm 3.51) \times 10^{-4}$	$(3.59 \pm 4.35) \times 10^{-5}$	$(1.59 \pm 1.98) \times 10^{-4}$	$(3.24 \pm 4.79) \times 10^{-5}$	$(1.65 \pm 0.00) \times 10^3$
SOSaG	$(2.21 \pm 2.85) \times 10^{-4}$	$(3.65 \pm 4.40) \times 10^{-5}$	$(1.99 \pm 3.20) \times 10^{-4}$	$(4.11 \pm 7.63) \times 10^{-5}$	$(2.43 \pm 0.01) \times 10^3$
ODE	$(2.72 \pm 1.75) \times 10^{-3}$	$(5.81 \pm 5.34) \times 10^{-4}$	$(1.92 \pm 1.62) \times 10^{-3}$	$(3.87 \pm 3.78) \times 10^{-4}$	$(1.47 \pm 0.01) \times 10^3$
DiffPDE	$(2.65 \pm 1.59) \times 10^{-3}$	$(5.49 \pm 4.48) \times 10^{-4}$	$(1.78 \pm 1.22) \times 10^{-3}$	$(3.83 \pm 3.38) \times 10^{-4}$	$(2.14 \pm 0.01) \times 10^3$
Two-step Methane Oxidation					
GEM	$(1.09 \pm 0.83) \times 10^{-4}$	$(2.40 \pm 1.75) \times 10^{-5}$	$(0.95 \pm 1.05) \times 10^{-4}$	$(1.93 \pm 1.98) \times 10^{-5}$	$(1.59 \pm 0.00) \times 10^3$
SOSaG	$(1.17 \pm 0.80) \times 10^{-4}$	$(2.58 \pm 1.66) \times 10^{-5}$	$(9.49 \pm 6.06) \times 10^{-5}$	$(1.85 \pm 1.24) \times 10^{-5}$	$(2.41 \pm 0.00) \times 10^3$
ODE	$(3.95 \pm 2.57) \times 10^{-3}$	$(9.59 \pm 7.02) \times 10^{-4}$	$(2.61 \pm 2.10) \times 10^{-3}$	$(5.92 \pm 4.68) \times 10^{-4}$	$(1.43 \pm 0.00) \times 10^3$
DiffPDE	$(3.41 \pm 2.78) \times 10^{-3}$	$(9.30 \pm 7.31) \times 10^{-4}$	$(2.37 \pm 1.94) \times 10^{-3}$	$(5.42 \pm 4.17) \times 10^{-4}$	$(2.11 \pm 0.00) \times 10^3$
Water Gas Shift					
GEM	$(7.47 \pm 5.25) \times 10^{-5}$	$(1.76 \pm 1.47) \times 10^{-5}$	$(6.09 \pm 6.19) \times 10^{-5}$	$(1.48 \pm 1.63) \times 10^{-5}$	$(1.59 \pm 0.00) \times 10^3$
SOSaG	$(7.61 \pm 4.96) \times 10^{-5}$	$(1.76 \pm 1.39) \times 10^{-5}$	$(6.50 \pm 5.43) \times 10^{-5}$	$(1.50 \pm 1.35) \times 10^{-5}$	$(2.37 \pm 0.00) \times 10^3$
ODE	$(5.96 \pm 4.08) \times 10^{-4}$	$(7.93 \pm 6.29) \times 10^{-5}$	$(3.26 \pm 3.50) \times 10^{-4}$	$(5.92 \pm 5.62) \times 10^{-5}$	$(1.40 \pm 0.00) \times 10^3$
DiffPDE	$(5.66 \pm 3.51) \times 10^{-4}$	$(7.25 \pm 4.90) \times 10^{-5}$	$(4.02 \pm 4.52) \times 10^{-4}$	$(6.68 \pm 6.98) \times 10^{-5}$	$(2.10 \pm 0.00) \times 10^3$

source term $R(\mathcal{C})$ in the ARD system (19) is then obtained by combining the reaction rates with the corresponding product and reactant stoichiometry.

For example, in the NO-O₃ mechanism with species

$$\mathcal{C} = (\mathcal{C}_{\text{NO}}, \mathcal{C}_{\text{O}_3}, \mathcal{C}_{\text{NO}_2}, \mathcal{C}_{\text{O}_2}),$$

we consider a single reaction, and therefore write $r_1 = r$. The reaction



has reaction rate

$$\chi(\mathcal{C}) = \kappa(\mathcal{K}) \mathcal{C}_{\text{NO}} \mathcal{C}_{\text{O}_3},$$

which yields the source term

$$R(\mathcal{C}) = (-\chi(\mathcal{C}), -\chi(\mathcal{C}), \chi(\mathcal{C}), \chi(\mathcal{C})).$$

Since only a single reaction is considered, NO and O₃ are consumed while NO₂ and O₂ are produced, so that the source term reflects the corresponding gain and loss of each species.

Data Generation

Training and test data are generated by numerically solving the resulting ARD system in an axisymmetric cylindrical reactor on a structured (r, z) grid of size $N_r \times N_z = 64 \times 64$. For the advection and diffusion we discretised them via finite difference and the reaction terms were solved separately using operator splitting and explicit time integration.

We leverage the axisymmetric structure of the reactor geometry and model only the radial and axial directions. Species concentration are fixed at the inlet $z = 0$ with a zero-gradient condition imposed at the outlet $z = L$ to allow the species to exit freely. Symmetry is imposed at the centre line $r = 0$ and a no flux condition is applied at the reactor wall $r = A$.

For each simulated experiment, we sample the temperature $\mathcal{K}$, average axial velocity U_{avg} , a base diffusivity D_{base} , and inlet concentrations for the feed species. Species diffusivities are derived as $D_s = D_{\text{base}} f_s$, where f_s is a mechanism-specific diffusion factor. Product-species inlet concentrations are fixed to zero unless produced by reaction.

For the training data we generate 2,000 simulations, each of which have 32 time snapshots meaning we have a total of 64,000 images to train our model on. Our test dataset consists of another 50 simulations. Further details of the dataset generation and model training are provided in Appendix D.

As some of the species concentrations are negligible compared to the more dominate species, the data has to be transformed before training the diffusion model in order to be able to capture the dynamics of the spatiotemporal structure for all of the species. To do this, a geometric mean of all strictly positive values is used to calculate a per species constant $c_{0,s}$

$$c_{0,s} = \exp \left(\frac{1}{N^+} \sum_{i: \mathcal{C}_i > 0} \log \mathcal{C}_i \right), \quad (25)$$

where N^+ is the number of positive values. Then the data is transformed

$$C_s^X = \log \left(1 + \frac{C_s}{c_{0,s}} \right). \quad (26)$$

Once the data is transformed, the data is normalised using a mean μ_X and variance σ_X that is computed over all values of all species. The diffusion model is then trained on this transformed and normalised data. The test data is transformed using the same species coefficients and normalisation statistics calculated in training and sampling is performed on this transformed and normalised test data. We can then convert the sampled data back to physical space to get the estimated species concentration using the following inverse transform

$$C_s = c_{0,s}(\exp(C_s^X \sigma_X + \mu_X) - 1). \quad (27)$$

We compare four sampling algorithms: the GEM and SOSaG methods used as proposals within a sequential Monte Carlo framework, the DiffusionPDE method [Huang et al., 2024], and a first-order guided ODE solver. Further details are provided in Appendix C. For the two stochastic methods, $N = 4$ particles were used as opposed to the $N = 1$ for the ODE samplers and all methods used $K = 400$ iterations for the discrete time process.

5 Discussion

Figure 1 gives a comparison of the predicted outlet estimates for the H_2O_2 Decomposition experiment for an example trajectory run. We can see that the stochastic methods are able to simulate and estimate the species concentration better than the alternative methods. Corresponding figures for the other experiments for an example trajectory are given in Appendix E.2.

Absolute errors for species concentrations and outlet concentrations across all timesteps, together with the average wall-clock time required to sample a full trajectory, are reported in Table 1. The stochastic methods consistently outperform the deterministic alternatives, often by approximately an order of magnitude. Despite using multiple samples, the vectorised implementation allows the stochastic methods to achieve wall-clock times comparable to those of the first and second-order deterministic solvers.

Relative percentage errors for both the transformed and physical concentration fields are reported in Table 2. Again, the stochastic methods significantly outperform the alternative approaches in both spaces. However, error metrics in physical space can be strongly influenced by species whose concentrations are very small relative to the dominant species in the system.

For example, in the H_2O_2 decomposition experiment, the concentrations of O_2 and H_2O remain several orders of magnitude smaller than that of H_2O_2 throughout the trajectory. As a result, small deviations in the transformed space can lead to very large relative errors after applying the inverse transformation to recover

physical concentrations. Consequently, relatively large percentage errors in physical space do not necessarily correspond to physically meaningful discrepancies. For instance, estimating a concentration of 1×10^{-8} when the true value is 1×10^{-10} produces a large relative error, even though such a difference would be far below the detection limits of typical laboratory instrumentation. For this reason, we set a minimum threshold ϕ when calculating the relative percentage which would correspond to a conservative estimate of the sensitivity of a real world sensor measurement. We have discussed this further in Appendix E.1.

Therefore, accurate estimation of the dominant species can still yield large relative percentage errors when low-concentration species are included in the metric. This effect is illustrated in Appendix E.3 where we provide an example trajectory reconstruction for each method over select time indices. The visual reconstructions demonstrate that the stochastic methods are typically able to recover the concentration fields from sparse observations and produce solutions that closely match the ARD dynamics in transformed space. After applying the inverse transformation, the resulting concentration fields in physical space remain visually consistent with the underlying dynamics, with minor deviations occurring primarily in low-concentration species. Nevertheless, even small discrepancies in transformed space can amplify into large relative errors when expressed in physical concentration units.

6 Conclusion

In this work we demonstrate, across multiple chemical kinetics datasets that closely reflect real-world laboratory conditions, that guided stochastic sampling methods are capable of reconstructing full spatiotemporal trajectories of time-varying PDE systems. Both qualitative and quantitative results indicate that the generated solutions are physically admissible.

Although relatively large errors are reported in Table 2, the visualisations in Appendix E.3 show that species concentrations are accurately reconstructed in the transformed space, which in turn leads to meaningful reconstructions in physical space. Furthermore, the outlet concentration profiles presented in Figure 1 and Appendix E.2 demonstrate that the proposed methods effectively capture the temporal evolution of species concentrations at the reactor outlet, as defined by the experimental setup in Figure 2.

Probabilistic Numerics Link and Uncertainty Quantification

Traditionally, the field of probabilistic numerics quantifies uncertainty by framing numerical problem-solving as a statistical inference task. This typically yields approximate solutions that carry structural uncertainty when applied to complex systems. Our work departs from this traditional approach: instead, we leverage the probabilistic setting of multimodal diffusion priors and stochastic processes to aim to produce plausible solutions to complex systems in more tractable time than

Table 2: Gas-phase reaction kinetics full-field and outlet relative percentage errors across all experiments and methods in transformed (normalised) and physical spaces. Sensitivity threshold $\phi = 10^{-8}$.

Method	Transformed Space		Physical Space	
	Rel. (%)	Outlet (%)	Rel. (%)	Outlet (%)
H₂O₂ Decomposition				
GEM	3.99 ± 1.85	0.743 ± 0.725	2.09 ± 2.06	2.38 ± 0.27
SOSaG	5.04 ± 4.85	0.619 ± 0.429	1.52 ± 0.26	2.84 ± 0.31
ODE	63.4 ± 58.8	18.5 ± 12.9	$(1.96 \pm 2.81) \times 10^4$	$(0.72 \pm 1.44) \times 10^5$
DiffPDE	43.3 ± 60.0	18.4 ± 7.2	$(2.01 \pm 1.73) \times 10^4$	$(2.21 \pm 3.00) \times 10^4$
NO + O₃ → NO₂				
GEM	7.3 ± 10.1	0.909 ± 0.668	5.18 ± 8.39	1.78 ± 0.85
SOSaG	9.2 ± 17.6	0.673 ± 0.307	4.20 ± 5.51	1.60 ± 0.86
ODE	66.9 ± 18.5	45.2 ± 25.0	$(5.01 \pm 4.94) \times 10^4$	$(2.99 \pm 3.45) \times 10^5$
DiffPDE	58.2 ± 24.2	33.2 ± 21.8	$(2.99 \pm 2.86) \times 10^4$	$(3.45 \pm 8.64) \times 10^5$
Ammonia Oxidation				
GEM	6.63 ± 3.01	4.82 ± 5.49	19.5 ± 41.9	4.40 ± 2.87
SOSaG	7.67 ± 3.69	5.30 ± 4.07	6.04 ± 8.14	3.91 ± 1.85
ODE	53.9 ± 20.7	52.7 ± 25.8	$(3.24 \pm 6.10) \times 10^3$	$(0.62 \pm 1.40) \times 10^3$
DiffPDE	55.1 ± 27.0	46.4 ± 18.3	$(4.40 \pm 6.49) \times 10^3$	$(1.03 \pm 1.78) \times 10^3$
Hydrogen Oxidation Subset				
GEM	3.54 ± 1.08	1.74 ± 1.40	$(2.05 \pm 3.07) \times 10^2$	$(0.90 \pm 2.85) \times 10^4$
SOSaG	3.17 ± 1.33	1.65 ± 2.97	$(2.95 \pm 9.04) \times 10^2$	5.93 ± 3.46
ODE	56.2 ± 31.5	68.6 ± 52.8	$(3.38 \pm 2.10) \times 10^4$	$(1.87 \pm 1.08) \times 10^5$
DiffPDE	61.3 ± 53.4	77.7 ± 83.2	$(3.80 \pm 3.25) \times 10^4$	$(2.34 \pm 1.90) \times 10^5$
2 Step Methane Oxidation				
GEM	6.35 ± 4.28	1.46 ± 1.02	20.6 ± 50.4	2.55 ± 0.95
SOSaG	9.6 ± 14.4	5.2 ± 12.0	$(2.20 \pm 6.82) \times 10^2$	$(0.70 \pm 2.22) \times 10^4$
ODE	$(1.38 \pm 1.75) \times 10^2$	71.2 ± 46.8	$(3.00 \pm 5.74) \times 10^4$	$(2.54 \pm 3.58) \times 10^5$
DiffPDE	$(1.22 \pm 1.49) \times 10^2$	52.3 ± 23.6	$(1.26 \pm 0.91) \times 10^4$	$(1.79 \pm 1.92) \times 10^5$
Water Gas Shift				
GEM	3.45 ± 1.40	1.54 ± 1.78	2.59 ± 1.07	2.40 ± 0.66
SOSaG	3.02 ± 1.11	1.84 ± 2.35	2.24 ± 0.69	2.67 ± 0.70
ODE	27.3 ± 7.9	20.0 ± 22.1	$(9.90 \pm 4.59) \times 10^4$	$(5.03 \pm 4.15) \times 10^4$
DiffPDE	27.9 ± 7.1	19.8 ± 16.2	$(9.33 \pm 7.29) \times 10^4$	$(4.51 \pm 5.46) \times 10^4$

classical solving methods allow.

Despite this difference, our work remains firmly situated within probabilistic numerics, since one of its recognized application areas is precisely the setting we address: “Typical numerical tasks to which PN may be applied include optimization, integration, the solution of ordinary and *partial differential equations*”¹. Accordingly, we believe this work contributes to the field by highlighting a lesser-known application of probabilistic numerics for tractable equation solving.

Nevertheless, understanding the errors of our methods is still an important aspect that needs to be discussed. The main sources of error come from the prior, the measurement noise and the PDE operator discretisation. To give an example of understanding the uncertainty for the latter case, we have provided standard deviation heatmaps in pixel space for the DiffusionPDE and SOSaG methods (Appendix F) for different resolutions. We can see from these that the SOSaG method gives a much lower standard deviation than the DiffusionPDE method, showing that the estimates are closer to the ground truth with a smaller deviation across runs. We can also see from the accompanying line plots that SOSaG is better at decreasing the standard deviation as the resolution increases.

Future Work

Although the ARD equation is commonly used to model reaction kinetics, it is not limited to this area. Previous work has applied them to areas such as evolutionary

biology [Lam and Lou, 2022], environmental ecosystems Cantin et al. [2021] and many others. In this work, we present a general likelihood for the ARD equation and therefore, with minor modifications, this framework can be applied to a wide range of areas outside of chemical kinetics.

Another promising direction for future work is the development of a single, generalised diffusion prior capable of learning the spatiotemporal structure of chemical kinetics systems. In this work, a separate diffusion prior is trained for each experiment. In contrast, future approaches could aim to train a single model that captures the underlying ARD dynamics across a range of parameter values. At test time, guided sampling could then be applied to real-world chemical datasets. Success in this direction would demonstrate that diffusion priors, combined with guided sampling, can generalise across entire families of PDE systems and parameter regimes, rather than requiring bespoke models for each individual chemical mixture or experimental setup.

Acknowledgements

This work was partially supported by the Swedish Research Council and the Wallenberg AI, Autonomous Systems and Software Program (WASP) funded by the Knut and Alice Wallenberg (KAW) Foundation. Computations were enabled by the supercomputing resource Berzelius provided by National Supercomputer Centre at Linköping University and the KAW foundation.

¹<https://www.probablistic-numerics.org/>

References

- J. Abramson, J. Adler, J. Dunger, R. Evans, T. Green, A. Pritzel, O. Ronneberger, L. Willmore, A. J. Ballard, J. Bambrick, et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*, 630(8016):493–500, 2024.
- P. W. Atkins, J. De Paula, and J. Keeler. *Atkins’ Physical Chemistry*. Oxford university press, 2023.
- R. Atkinson and W. P. Carter. Kinetics and mechanisms of the gas-phase reactions of ozone with organic compounds under atmospheric conditions. *Chemical Reviews*, 84(5):437–470, 1984.
- G. Cantin, A. Ducrot, and B. M. Funatsu. Mathematical modeling of forest ecosystems by a reaction–diffusion–advection system: impacts of climate change and deforestation. *Journal of Mathematical Biology*, 83(6):66, 2021.
- H. Cao, C. Tan, Z. Gao, Y. Xu, G. Chen, P.-A. Heng, and S. Z. Li. A Survey on Generative Diffusion Models. *IEEE Transactions on Knowledge and Data Engineering*, 36(7):2814–2830, 2024.
- G. Cardoso, Y. J. el idrissi, S. L. Corff, and E. Moulines. Monte Carlo guided Denoising Diffusion models for Bayesian linear inverse problems. In *The Twelfth International Conference on Learning Representations*, 2024. URL <https://openreview.net/forum?id=nHESwXvxWK>.
- N. Chopin and O. Papaspiliopoulos. *An Introduction to Sequential Monte Carlo*, volume 4. Springer, 2020.
- C. Cosner. Reaction-diffusion-advection models for the effects and evolution of dispersal. *Discrete and Continuous Dynamical Systems*, 34(5):1701–1745, 2014. ISSN 1078-0947. doi: 10.3934/dcds.2014.34.1701. URL <https://www.aims sciences.org/article/id/6fe19f9a-2394-4838-9ed0-e1883781ca39>.
- O. Danielsson, M. Karlsson, P. Sukkaew, H. Pedersen, and L. Ojamae. A Systematic Method for Predictive in silico Chemical Vapor Deposition. *The Journal of Physical Chemistry C*, 124(14):7725–7736, 2020.
- Z. Dou and Y. Song. Diffusion Posterior Sampling for Linear Inverse Problem Solving: A Filtering Perspective. In *The Twelfth International Conference on Learning Representations*, 2024.
- R. Douc and O. Cappé. Comparison of Resampling Schemes for Particle Filtering. In *ISPA 2005. Proceedings of the 4th International Symposium on Image and Signal Processing and Analysis, 2005.*, pages 64–69. IEEE, 2005.
- N. B. Erichson, V. Mikuni, D. Lyu, Y. Gao, O. Azencot, S. H. Lim, and M. W. Mahoney. FLEX: A Backbone for Diffusion-Based Modeling of Spatio-temporal Physical Systems. *CoRR*, abs/2505.17351, May 2025. URL <https://doi.org/10.48550/arXiv.2505.17351>.
- N. J. Gordon, D. J. Salmond, and A. F. Smith. Novel Approach to Nonlinear/Non-Gaussian Bayesian State Estimation. In *IEEE Proceedings F (Radar and Signal Processing)*, volume 140, pages 107–113. IET, 1993.
- H. Herrmann. Kinetics of aqueous phase reactions relevant for atmospheric chemistry. *Chemical reviews*, 103(12):4691–4716, 2003.
- J. Ho, A. Jain, and P. Abbeel. Denoising Diffusion Probabilistic Models. *Advances in Neural Information Processing Systems*, 33:6840–6851, 2020.
- J. Huang, G. Yang, Z. Wang, and J. J. Park. DiffusionPDE: Generative PDE-Solving under Partial Observation, 2024. URL <https://openreview.net/forum?id=z0I2SbjNOR>.
- C. Jacobsen, Y. Zhuang, and K. Duraisamy. CocoGen: Physically Consistent and Conditioned Score-Based Generative Models for Forward and Inverse Problems. *SIAM Journal on Scientific Computing*, 47(2):C399–C425, 2025.
- J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Židek, A. Potapenko, et al. Highly accurate protein structure prediction with AlphaFold. *Nature*, 596(7873):583–589, 2021.
- T. Karras, M. Aittala, T. Aila, and S. Laine. Elucidating the Design Space of Diffusion-Based Generative Models, 2022. URL <https://openreview.net/forum?id=k7FuTOWM0c7>.
- F. E. Kelvinius, Z. Zhao, and F. Lindsten. Solving Linear-Gaussian Bayesian Inverse Problems with Decoupled Diffusion Sequential Monte Carlo. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pages 15148–15181. PMLR, 13–19 Jul 2025.
- H.-O. Kreiss and G. Scherer. Finite Element and Finite Difference Methods for Hyperbolic Partial Differential Equations. In *Mathematical Aspects of Finite Elements in Partial Differential Equations*, pages 195–212. Elsevier, 1974.
- K.-Y. Lam and Y. Lou. *Introduction to Reaction-Diffusion Equations: Theory and Applications to Spatial Ecology and Evolutionary Biology*. Springer Nature, 01 2022. ISBN 978-3-031-20421-0. doi: 10.1007/978-3-031-20422-7.
- R. J. LeVeque. *Finite Difference Methods for Ordinary and Partial Differential Equations: Steady-State and Time-Dependent Problems*. SIAM, 2007.
- M. Meuwly. Machine Learning for Chemical Reactions. *Chemical Reviews*, 121(16):10218–10239, 2021. doi: 10.1021/acs.chemrev.1c00033. URL <https://doi.org/10.1021/acs.chemrev.1c00033>. PMID: 34097378.
- A. Millard, F. Lindsten, and Z. Zhao. Particle-guided diffusion models for partial differential equations. *arXiv preprint 2601.23262*, 2026.

- P. Monk. *Finite Element Methods for Maxwell's Equations*. Oxford University Press, 2003.
- C. A. Naesseth, F. Lindsten, and T. B. Schön. Elements of sequential Monte Carlo. *Foundations and Trends in Machine Learning*, 12(3):307–392, 2019. doi: 10.1561/22000000074. URL <https://arxiv.org/abs/1903.04797>.
- P. G. Robert J. Kee, Michael E. Coltrin. *The Conservation Equations*, pages 67–149. John Wiley & Sons, Ltd, 2003. ISBN 9780471461296. doi: <https://doi.org/10.1002/0471461296.ch3>. URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/0471461296.ch3>.
- A. Rubio, A. Zalts, and C. El Hasi. Numerical solution of the advection–reaction–diffusion equation at different scales. *Environmental Modelling & Software*, 23(1):90–95, 2008. ISSN 1364-8152. doi: <https://doi.org/10.1016/j.envsoft.2007.05.009>. URL <https://www.sciencedirect.com/science/article/pii/S136481520700093X>.
- J. Sohl-Dickstein, E. Weiss, N. Maheswaranathan, and S. Ganguli. Deep Unsupervised Learning Using Nonequilibrium Thermodynamics. In *International Conference on Machine Learning*, pages 2256–2265. PMLR, 2015.
- Y. Song, J. Sohl-Dickstein, D. P. Kingma, A. Kumar, S. Ermon, and B. Poole. Score-Based Generative Modeling through Stochastic Differential Equations, 2021. URL <https://openreview.net/forum?id=PxTIG12RRHS>.
- M. Staszak. Artificial intelligence in the modeling of chemical reactions kinetics. *Physical Sciences Reviews*, 8(1):51–72, 2023. doi: doi:10.1515/psr-2020-0079. URL <https://doi.org/10.1515/psr-2020-0079>.
- T. S. Stevens, O. Nolan, J.-L. Robert, and R. J. Van Sloun. Sequential Posterior Sampling with Diffusion Models. In *ICASSP 2025–2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1–5. IEEE, 2025.
- S. Stocker, G. Csányi, K. Reuter, and J. T. Margraf. Machine learning in chemical reaction space. *Nature communications*, 11(1):5505, 2020.
- L. N. Trefethen. *Finite Difference and Spectral Methods for Ordinary and Partial Differential Equations*. Cornell University-Department of Computer Science and Center for Applied ..., 1996.
- B. L. Trippe, J. Yim, D. Tischer, D. Baker, T. Broderick, R. Barzilay, and T. S. Jaakkola. Diffusion Probabilistic Modeling of Protein Backbones in 3D for the motif-scaffolding problem. In *The Eleventh International Conference on Learning Representations*, 2023.
- L. Wu, B. L. Trippe, C. A. Naesseth, D. Blei, and J. P. Cunningham. Practical and asymptotically exact conditional sampling in diffusion models. In *Advances in Neural Information Processing Systems 36 (NeurIPS)*, 2023. URL <https://openreview.net/forum?id=eWKqr1zcRv>.
- G. Yang and S. Sommer. A Denoising Diffusion Model for Fluid Field Prediction. *arXiv preprint arXiv:2301.11661*, 2023.
- J. Yao, A. Mammadov, J. Berner, G. Kerrigan, J. C. Ye, K. Azizzadenesheli, and A. Anandkumar. Guided Diffusion Sampling on Function Spaces with Applications to PDEs. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*, 2025. URL <https://openreview.net/forum?id=oAgwvZay2U>.
- Z. Zhao, Z. Luo, J. Sjölund, and T. Schön. Conditional sampling within generative diffusion models. *Philosophical Transactions A*, 383(2299):20240329, 2025.

A SMC Pseudocode

Algorithm 1 Sequential Monte Carlo (SMC)

Require: Number of particles N , time steps K ; proposals $M_K(x_K), M_k(x_{k-1} | x_k)$; incremental weight functions $G_K(x_K), G_k(x_{k-1}, x_k)$; re-sampling threshold $N_{eff} \in (0, 1]$.

```

1 Initialization:
2 for  $n = 1 \rightarrow N$  do
3    $x_K^{(i)} \sim M_K(x_K)$ 
4    $w_K^{(i)} = G_K(x_K)$ 
5 for  $k = K \rightarrow 1$  do
6   for  $i = 1 \rightarrow N$  do
7     Propagate:  $x_{k-1}^{(i)} \sim M_k(x_{k-1} | x_k^{(i)})$ 
8     Weight update:  $\tilde{w}_{k-1}^{(i)} = w_k^{(i)} G_k(x_{k-1}^{(i)}, x_k^{(i)})$ 
9   for  $i = 1 \rightarrow N$  do
10    Normalize:  $w_{k-1}^{(i)} = \tilde{w}_{k-1}^{(i)} / \sum_{j=1}^N \tilde{w}_{k-1}^{(j)}$ 
11  Compute ESS:  $ESS_k = \frac{1}{\sum_{i=1}^N (w_{k-1}^{(i)})^2}$ 
12  if  $ESS_k \leq \tau$  then
13    Resampling (Multinomial):
14    for  $i = 1 \rightarrow N$  do
15       $a^{(i)} \sim \text{Cat}(w_{k-1}^{(1)}, \dots, w_{k-1}^{(N)})$ 
16       $x_{k-1}^{(i)} = x_{k-1}^{(a^{(i)})}$ 
17       $w_{k-1}^{(i)} = \frac{1}{N}$ 
18 Return:  $\{(x_0^{(i)}, w_0^{(i)})\}_{i=1}^N$ 

```

B Stochastic Solvers Pseudocode

B.1 GEM

Algorithm 2 GEM Algorithm

Require: $\delta_\theta(x; \sigma)$, $\sigma_k, \sigma_{k-1}, x_k$

```

1 Sample  $\epsilon_k \sim \mathcal{N}(0, I)$ 
2  $d_k \leftarrow \frac{x_k - \delta_\theta(x_k, \sigma_k)}{\sigma_k}$ 
3  $x_{k-1} \leftarrow x_k + (\sigma_{k-1}^2 - \sigma_k^2) d_k + \sqrt{\sigma_{k-1}^2 - \sigma_k^2} \epsilon_k$ 
4  $x_{k-1} \leftarrow x_{k-1} - (\sigma_{k-1}^2 - \sigma_k^2) \nabla_{x_k} \log \tilde{p}_\theta(y | x_k)$ 
5 return  $x_{k-1}$ 

```

B.2 SOSaG

Algorithm 3 SOSaG Algorithm

Require: $\delta_\theta(x, \sigma)$, $\sigma_k, \sigma_{k-1}, x_k, \gamma_k, \alpha$

```

1 Sample  $\epsilon_k \sim \mathcal{N}(0, I)$ 
2  $\hat{\sigma}_k \leftarrow \sigma_k + \gamma_k \sigma_k$ 
3  $\hat{x}_k \leftarrow x_k + \sqrt{\hat{\sigma}_k^2 - \sigma_k^2} \epsilon_k$ 
4  $d_k \leftarrow \frac{\hat{x}_k - \delta_\theta(\hat{x}_k, \hat{\sigma}_k)}{\hat{\sigma}_k}$ 
5  $x_{k-1} \leftarrow \hat{x}_k + (\sigma_{k-1} - \hat{\sigma}_k) d_k$ 
6 if  $\sigma_k \neq 0$  then
7    $d_{k-1} \leftarrow \frac{x_{k-1} - \delta_\theta(x_{k-1}, \sigma_{k-1})}{\sigma_{k-1}}$ 
8    $x_{k-1} \leftarrow \hat{x}_k + (\sigma_{k-1} - \hat{\sigma}_k) (\frac{1}{2} d_k + \frac{1}{2} d_{k-1})$ 
9  $x_{k-1} \leftarrow x_{k-1} - \alpha \nabla_{\hat{x}_k} \log \tilde{p}_\theta(y | \hat{x}_k)$ 
10 return  $x_{k-1}$ 

```

C ODE Solvers

C.1 Guided ODE

The Euler discretisation to numerically approximate the solution to the ODE in Equation (2) is given by

$$x_{k-1} = x_k + \frac{1}{2}(\sigma_{k-1}^2 - \sigma_k^2) \frac{x_k - \delta_\theta(x_k, \sigma_k)}{\sigma_k^2} \quad (28)$$

and the corresponding guided-Euler (GE) discretisation method is given by

$$x_{k-1} = x_k + \frac{1}{2}(\sigma_{k-1}^2 - \sigma_k^2) \frac{x_k - \delta_\theta(x_k, \sigma_k)}{\sigma_k^2} + \frac{1}{2}(\sigma_{k-1}^2 - \sigma_k^2) \nabla_{x_k} \log \tilde{p}_\theta(y | x_k). \quad (29)$$

The pseudocode for Equation (29) is given in Algorithm 4

Algorithm 4 GE Algorithm

Require: $\delta_\theta(x; \sigma)$, $\sigma_k, \sigma_{k-1}, x_k$

```

1 Sample  $\epsilon_k \sim \mathcal{N}(0, I)$ 
2  $d_k \leftarrow \frac{x_k - \delta_\theta(x_k, \sigma_k)}{\sigma_k}$ 
3  $x_{k-1} \leftarrow x_k + (\sigma_{k-1}^2 - \sigma_k^2) d_k + \sqrt{\sigma_{k-1}^2 - \sigma_k^2} \epsilon_k$ 
4  $x_{k-1} \leftarrow x_{k-1} - (\sigma_{k-1}^2 - \sigma_k^2) \nabla_{x_k} \log \tilde{p}_\theta(y | x_k)$ 
5 return  $x_{k-1}$ 

```

C.2 DiffPDE

Influenced by Karras et al. [2022], Huang et al. [2024] extend the GE proposal in Equation (29) by incorporating a second-order correction term. This discretisation scheme is given in Algorithm 5.

Algorithm 5 DiffPDE Algorithm

Require: $\delta_\theta(x, \sigma)$, $\sigma_k, \sigma_{k-1}, x_k, \alpha$

```

1  $d_k \leftarrow \frac{x_k - \delta_\theta(x_k, \sigma_k)}{\sigma_k}$ 
2  $x_{k-1} \leftarrow x_k + (\sigma_{k-1} - \hat{\sigma}_k) d_k$ 
3 if  $\sigma_k \neq 0$  then
4    $d_{k-1} \leftarrow \frac{x_{k-1} - \delta_\theta(x_{k-1}, \sigma_{k-1})}{\sigma_{k-1}}$ 
5    $x_{k-1} \leftarrow x_k + (\sigma_{k-1} - \sigma_k) (\frac{1}{2} d_k + \frac{1}{2} d_{k-1})$ 
6  $x_{k-1} \leftarrow x_{k-1} - \alpha \nabla_{\hat{x}_k} \log \tilde{p}_\theta(y | \hat{x}_k)$ 
7 return  $x_{k-1}$ 

```

D Chemistry dataset generation and model training

Table 3 gives the summary of the ranges parameters were sampled from when generating the dataset. Table 4 gives the Arrhenius parameters used when computing the reaction networks and Table 5 gives the diffusion scaling factors for each experiment. We would like to note that the generated dataset therefore is the ground truth from which we compare our solutions against.

As mentioned previously, the models are trained on all of the time snapshots within each trajectory and thus has

Table 3: Summary of reduced chemical mechanisms and sampled parameter ranges. $[\cdot, \cdot]$ denote the ranges the parameters were uniformly sampled from.

Mechanism	Species	Feed Species	T (K)	U_{avg} (m/s)	$\log_{10} D_{\text{base}}$
NO-O ₃ to NO ₂	NO, O ₃ , NO ₂ , O ₂	NO, O ₃	[270, 330]	[0.01, 0.5]	$[-6.5, -5.2]$
Water-gas shift	CO, H ₂ O, CO ₂ , H ₂	CO, H ₂ O	[500, 1000]	[0.01, 0.5]	$[-6.5, -5.2]$
Hydrogen peroxide decomposition	H ₂ O ₂ , H ₂ O, O ₂	H ₂ O ₂	[290, 600]	[0.01, 0.5]	$[-6.5, -5.2]$
Two-step methane oxidation	CH ₄ , O ₂ , CO, CO ₂ , H ₂ O	CH ₄ , O ₂	[700, 1500]	[0.01, 0.5]	$[-6.5, -5.2]$
Ammonia oxidation	NH ₃ , O ₂ , NO, NO ₂ , H ₂ O	NH ₃ , O ₂	[600, 1200]	[0.01, 0.5]	$[-6.5, -5.2]$
Hydrogen oxidation subset	H ₂ , O ₂ , OH, H ₂ O, H, HO ₂	H ₂ , O ₂	[800, 1800]	[0.01, 0.5]	$[-6.5, -5.2]$

Table 4: Reaction networks and Arrhenius parameters used in the dataset generation process.

Mechanism	Reaction l	A_l	$E_{a,l}$ (J/mol)
NO-O ₃ to NO ₂	$\text{NO} + \text{O}_3 \rightarrow \text{NO}_2 + \text{O}_2$	2.5×10^6	1.2×10^4
Water-gas shift	$\text{CO} + \text{H}_2\text{O} \rightarrow \text{CO}_2 + \text{H}_2$	1.5×10^3	6.5×10^4
	$\text{CO}_2 + \text{H}_2 \rightarrow \text{CO} + \text{H}_2\text{O}$	4.0×10^2	7.0×10^4
Hydrogen peroxide decomposition	$2 \text{H}_2\text{O}_2 \rightarrow 2 \text{H}_2\text{O} + \text{O}_2$	2.0×10^2	5.5×10^4
Two-step methane oxidation	$2 \text{CH}_4 + 3 \text{O}_2 \rightarrow 2 \text{CO} + 4 \text{H}_2\text{O}$	8.0×10^3	1.2×10^5
	$2 \text{CO} + \text{O}_2 \rightarrow 2 \text{CO}_2$	2.5×10^4	8.5×10^4
Ammonia oxidation	$4 \text{NH}_3 + 5 \text{O}_2 \rightarrow 4 \text{NO} + 6 \text{H}_2\text{O}$	6.0×10^3	9.5×10^4
	$2 \text{NO} + \text{O}_2 \rightarrow 2 \text{NO}_2$	1.2×10^4	6.0×10^4
Hydrogen oxidation subset	$\text{H}_2 + \text{O}_2 \rightarrow 2 \text{OH}$	1.0×10^5	1.4×10^5
	$\text{H}_2 + \text{OH} \rightarrow \text{H}_2\text{O} + \text{H}$	8.0×10^5	2.0×10^4
	$\text{H} + \text{O}_2 \rightarrow \text{HO}_2$	2.0×10^6	1.5×10^4
	$\text{HO}_2 + \text{H} \rightarrow 2 \text{OH}$	1.5×10^6	1.0×10^4

Table 5: Species-dependent diffusion scaling factors used in the dataset generation process.

Mechanism	Diffusion factors f_s
NO-O ₃ to NO ₂	NO: 1.00, O ₃ : 0.95, NO ₂ : 0.92, O ₂ : 1.08
Water-gas shift	CO: 1.03, H ₂ O: 0.90, CO ₂ : 0.88, H ₂ : 1.35
Hydrogen peroxide decomposition	H ₂ O ₂ : 0.85, H ₂ O: 1.00, O ₂ : 1.20
Two-step methane oxidation	CH ₄ : 1.20, O ₂ : 1.05, CO: 1.00, CO ₂ : 0.88, H ₂ O: 0.92
Ammonia oxidation	NH ₃ : 1.08, O ₂ : 1.03, NO: 1.00, NO ₂ : 0.93, H ₂ O: 0.90
Hydrogen oxidation subset	H ₂ : 1.60, O ₂ : 1.05, OH: 1.25, H ₂ O: 0.90, H: 1.80, HO ₂ : 0.95

the opportunity to learn the spatiotemporal structure, including both the long and short-term dynamics that is induced by the logarithmic spacing of the chemical reactions. Each model takes approximately 4.5 hours to train (1 million image passes per model). At inference time, snapshots are reconstructed with temporal coherence enforced through a PDE residual-based likelihood, coupling consecutive time steps via the governing equations. The learned spatial prior reconstructs each time step, while the residual coupling ensures consistency with the system’s temporal dynamics. This deliberate design keeps the generative model flexible and data-driven while physical constraints in the likelihood enforce temporal consistency. Each time step takes 1.5 minutes to reconstruct meaning each trajectory takes ≈ 45 minutes to reconstruct. This further motivates work for a single model that encompasses multiple chemical reactions and also a method which can reconstruct multiple time steps at the same time to reduce the time taken for the sequential reconstruction.

E Supplementary Results

E.1 Absolute Error Tables

We report relative percentage errors defined as

$$\text{RPE}(x_{\text{pred}}, x_{\text{true}}) = 100 \times \frac{|x_{\text{pred}} - x_{\text{true}}|}{\max(|x_{\text{true}}|, \phi)},$$

where ϕ is a small positive constant introduced to ensure numerical stability in regions where the ground truth approaches zero.

Table 6 reports results using $\phi = 10^{-6}$, while Table 7 shows results with a smaller threshold $\phi = 10^{-12}$. As expected, reducing ϕ leads to substantially larger relative errors. This behaviour arises in regions where chemical species concentrations are very small in physical space, causing the relative error to be dominated by the denominator.

Importantly, these large relative errors do not correspond to significant degradation in reconstruction quality. As shown in Figure 1 and Appendices E.3 and E.2, the

methods still accurately recover both the concentration fields and outlet quantities. The elevated errors are therefore primarily an artefact of the metric in low-concentration regimes rather than a failure of the model.

Table 6: Gas-phase reaction kinetics full-field and outlet relative percentage errors across all experiments and methods in transformed (normalised) and physical spaces. $\phi = 10^{-6}$.

Method	Transformed Space		Physical Space	
	Rel. (%)	Outlet (%)	Rel. (%)	Outlet (%)
H₂O₂ Decomposition				
GEM	3.55 ± 1.65	0.743 ± 0.725	0.543 ± 0.343	0.345 ± 0.217
SOSaG	3.63 ± 1.87	0.619 ± 0.429	0.428 ± 0.097	0.362 ± 0.186
ODE	55.3 ± 58.3	18.5 ± 12.9	$(2.68 \pm 3.32) \times 10^2$	$(0.73 \pm 1.44) \times 10^3$
DiffPDE	41.7 ± 60.7	18.4 ± 7.2	$(2.47 \pm 1.93) \times 10^2$	$(4.12 \pm 8.66) \times 10^2$
NO + O₃ → NO₂				
GEM	7.3 ± 10.1	0.909 ± 0.668	1.77 ± 0.32	1.45 ± 0.69
SOSaG	9.2 ± 17.6	0.673 ± 0.307	1.65 ± 0.36	1.32 ± 0.62
ODE	66.9 ± 18.5	45.2 ± 25.0	$(8.67 \pm 9.07) \times 10^2$	$(4.97 \pm 8.44) \times 10^3$
DiffPDE	58.2 ± 24.2	33.2 ± 21.8	$(3.51 \pm 2.89) \times 10^2$	$(3.48 \pm 8.64) \times 10^3$
Ammonia Oxidation				
GEM	6.63 ± 3.01	4.82 ± 5.49	1.54 ± 0.55	2.06 ± 2.82
SOSaG	7.67 ± 3.69	5.30 ± 4.07	1.36 ± 0.55	1.83 ± 1.76
ODE	53.9 ± 20.7	52.7 ± 25.8	44.5 ± 58.4	20.9 ± 21.1
DiffPDE	55.1 ± 27.0	46.4 ± 18.3	70.6 ± 93.0	$(1.86 \pm 5.05) \times 10^2$
Hydrogen Oxidation Subset				
GEM	3.54 ± 1.07	1.74 ± 1.40	4.47 ± 4.43	92.0 ± 284.0
SOSaG	3.16 ± 1.32	1.65 ± 2.97	5.8 ± 13.0	1.77 ± 2.45
ODE	55.8 ± 30.2	68.6 ± 52.8	$(3.91 \pm 2.14) \times 10^2$	$(1.93 \pm 1.15) \times 10^3$
DiffPDE	60.3 ± 50.4	77.7 ± 83.2	$(4.42 \pm 3.53) \times 10^2$	$(2.95 \pm 1.85) \times 10^3$
2 Step Methane Oxidation				
GEM	6.17 ± 3.97	1.46 ± 1.02	1.04 ± 0.93	0.617 ± 0.722
SOSaG	9.3 ± 13.8	5.2 ± 12.0	3.52 ± 7.10	72.0 ± 221.0
ODE	$(1.28 \pm 1.51) \times 10^2$	71.2 ± 46.8	$(4.24 \pm 8.37) \times 10^2$	$(2.64 \pm 3.68) \times 10^3$
DiffPDE	$(1.11 \pm 1.23) \times 10^2$	52.3 ± 23.6	$(1.62 \pm 1.12) \times 10^2$	$(1.94 \pm 1.99) \times 10^3$
Water Gas Shift				
GEM	3.45 ± 1.39	1.54 ± 1.78	1.03 ± 0.44	0.946 ± 0.614
SOSaG	3.02 ± 1.11	1.84 ± 2.35	0.947 ± 0.380	1.06 ± 0.72
ODE	27.3 ± 7.9	20.0 ± 22.1	$(1.12 \pm 0.54) \times 10^3$	$(5.07 \pm 4.16) \times 10^2$
DiffPDE	27.9 ± 7.1	19.8 ± 16.2	$(1.06 \pm 0.81) \times 10^3$	$(4.57 \pm 5.45) \times 10^2$

Table 7: Gas-phase reaction kinetics full-field and outlet relative percentage errors across all experiments and methods in transformed (normalised) and physical spaces. $\phi = 10^{-12}$.

Method	Transformed Space		Physical Space	
	Rel. (%)	Outlet (%)	Rel. (%)	Outlet (%)
H₂O₂ Decomposition				
GEM	3.99 ± 1.85	0.743 ± 0.725	(5.11 ± 0.98) × 10³	(1.63 ± 0.28) × 10⁴
SOSaG	5.04 ± 4.85	0.619 ± 0.429	(6.21 ± 0.59) × 10 ³	(1.98 ± 0.31) × 10 ⁴
ODE	63.4 ± 58.8	18.5 ± 12.9	(1.49 ± 2.22) × 10 ⁸	(0.72 ± 1.44) × 10 ⁹
DiffPDE	43.3 ± 60.0	18.4 ± 7.2	(1.70 ± 1.48) × 10 ⁸	(1.92 ± 2.24) × 10 ⁸
NO + O₃ → NO₂				
GEM	7.3 ± 10.1	0.909 ± 0.668	(2.77 ± 8.30) × 10 ⁴	(1.97 ± 0.63) × 10²
SOSaG	9.2 ± 17.6	0.673 ± 0.307	(2.00 ± 5.35) × 10⁴	(2.27 ± 0.48) × 10 ²
ODE	66.9 ± 18.5	45.2 ± 25.0	(3.08 ± 2.66) × 10 ⁸	(2.52 ± 2.75) × 10 ⁹
DiffPDE	58.2 ± 24.2	33.2 ± 21.8	(2.80 ± 2.90) × 10 ⁸	(3.38 ± 8.67) × 10 ⁹
Ammonia Oxidation				
GEM	6.63 ± 3.01	4.82 ± 5.49	(1.69 ± 4.23) × 10 ⁵	(2.17 ± 0.40) × 10 ⁴
SOSaG	7.67 ± 3.69	5.30 ± 4.07	(4.27 ± 8.32) × 10⁴	(1.86 ± 0.37) × 10⁴
ODE	53.9 ± 20.7	52.7 ± 25.8	(3.11 ± 6.14) × 10 ⁷	(0.60 ± 1.40) × 10 ⁷
DiffPDE	55.1 ± 27.0	46.4 ± 18.3	(2.76 ± 4.19) × 10 ⁷	(0.84 ± 1.79) × 10 ⁷
Hydrogen Oxidation Subset				
GEM	3.54 ± 1.08	1.74 ± 1.40	(1.75 ± 2.98) × 10⁶	(0.90 ± 2.85) × 10 ⁸
SOSaG	3.17 ± 1.33	1.65 ± 2.97	(2.14 ± 6.59) × 10 ⁶	(3.63 ± 0.87) × 10⁴
ODE	56.2 ± 31.5	68.6 ± 52.8	(2.92 ± 1.93) × 10 ⁸	(1.80 ± 1.08) × 10 ⁹
DiffPDE	61.3 ± 53.4	77.7 ± 83.2	(3.36 ± 3.05) × 10 ⁸	(2.24 ± 1.89) × 10 ⁹
2 Step Methane Oxidation				
GEM	6.35 ± 4.28	1.46 ± 1.02	(1.93 ± 5.03) × 10⁵	(1.67 ± 0.16) × 10⁴
SOSaG	9.6 ± 14.4	5.2 ± 12.0	(2.19 ± 6.83) × 10 ⁶	(0.70 ± 2.22) × 10 ⁸
ODE	(1.38 ± 1.75) × 10 ²	71.2 ± 46.8	(1.81 ± 2.63) × 10 ⁸	(2.54 ± 3.58) × 10 ⁹
DiffPDE	(1.22 ± 1.49) × 10 ²	52.3 ± 23.6	(1.02 ± 0.78) × 10 ⁸	(1.70 ± 1.92) × 10 ⁹
Water Gas Shift				
GEM	3.45 ± 1.40	1.54 ± 1.78	(7.05 ± 3.31) × 10 ³	(1.14 ± 0.19) × 10⁴
SOSaG	3.02 ± 1.11	1.84 ± 2.35	(6.07 ± 1.65) × 10³	(1.25 ± 0.18) × 10 ⁴
ODE	27.3 ± 7.9	20.0 ± 22.1	(8.57 ± 3.93) × 10 ⁸	(5.03 ± 4.15) × 10 ⁸
DiffPDE	27.9 ± 7.1	19.8 ± 16.2	(8.11 ± 6.54) × 10 ⁸	(4.33 ± 5.46) × 10 ⁸

E.2 Outlet Estimation Figures

Figures 3–7 give the estimated species concentration at the outlet from $\tau = 0.01 \rightarrow 30$ seconds for all methods for the same example run.

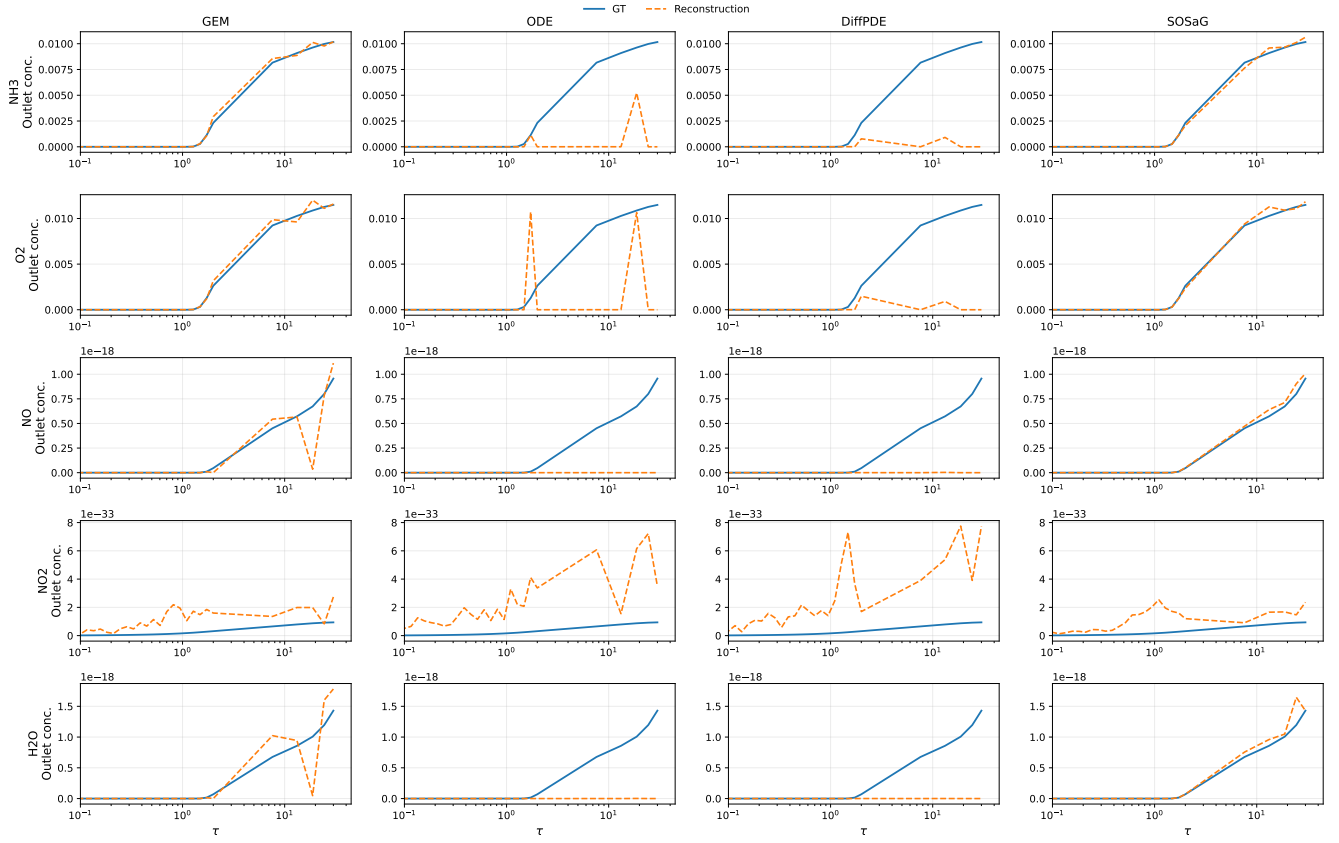


Figure 3: Species outlet comparison for Ammonia Oxidation.

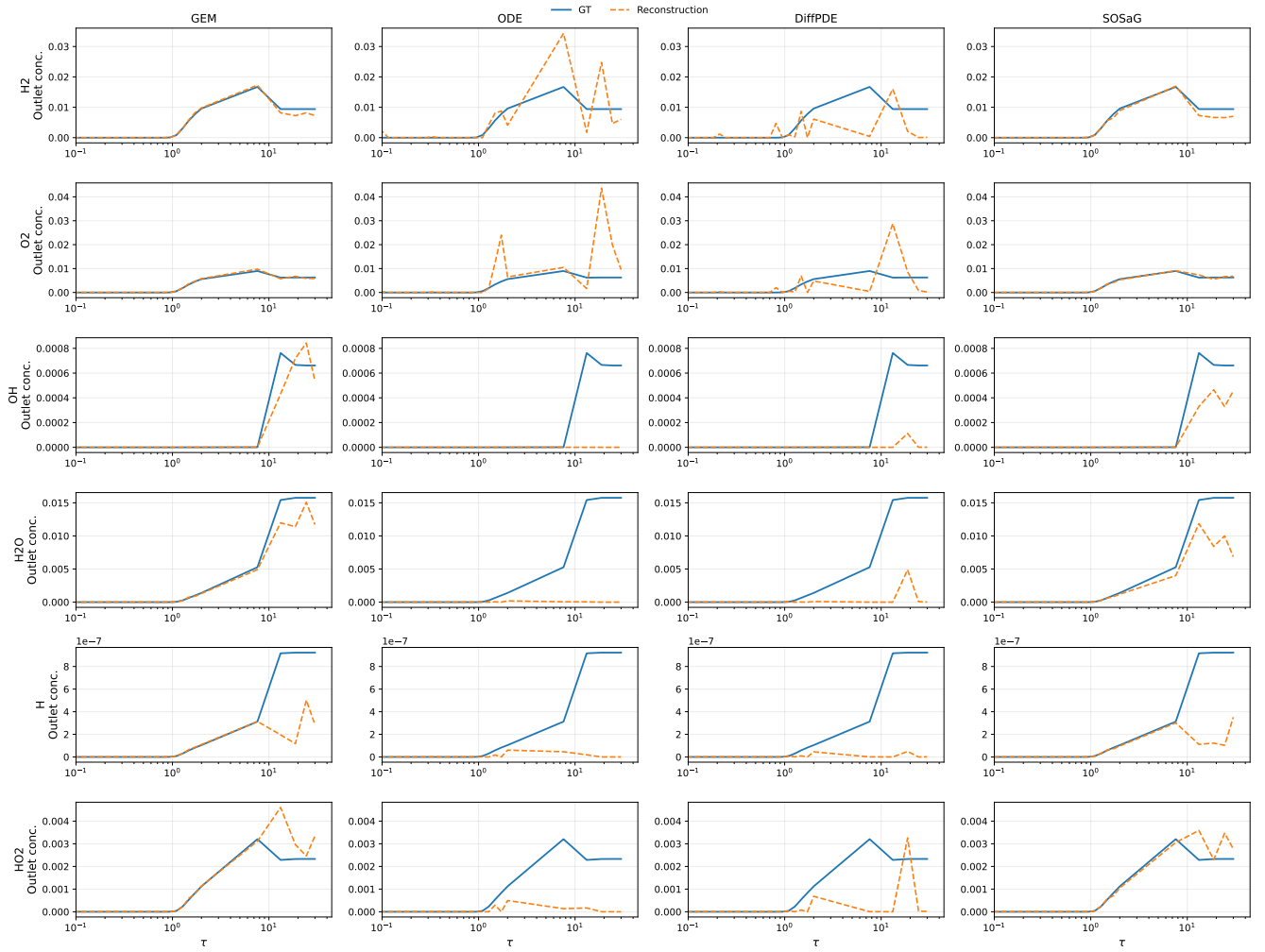


Figure 4: Species outlet comparison for Hydrogen Oxidation Subset.

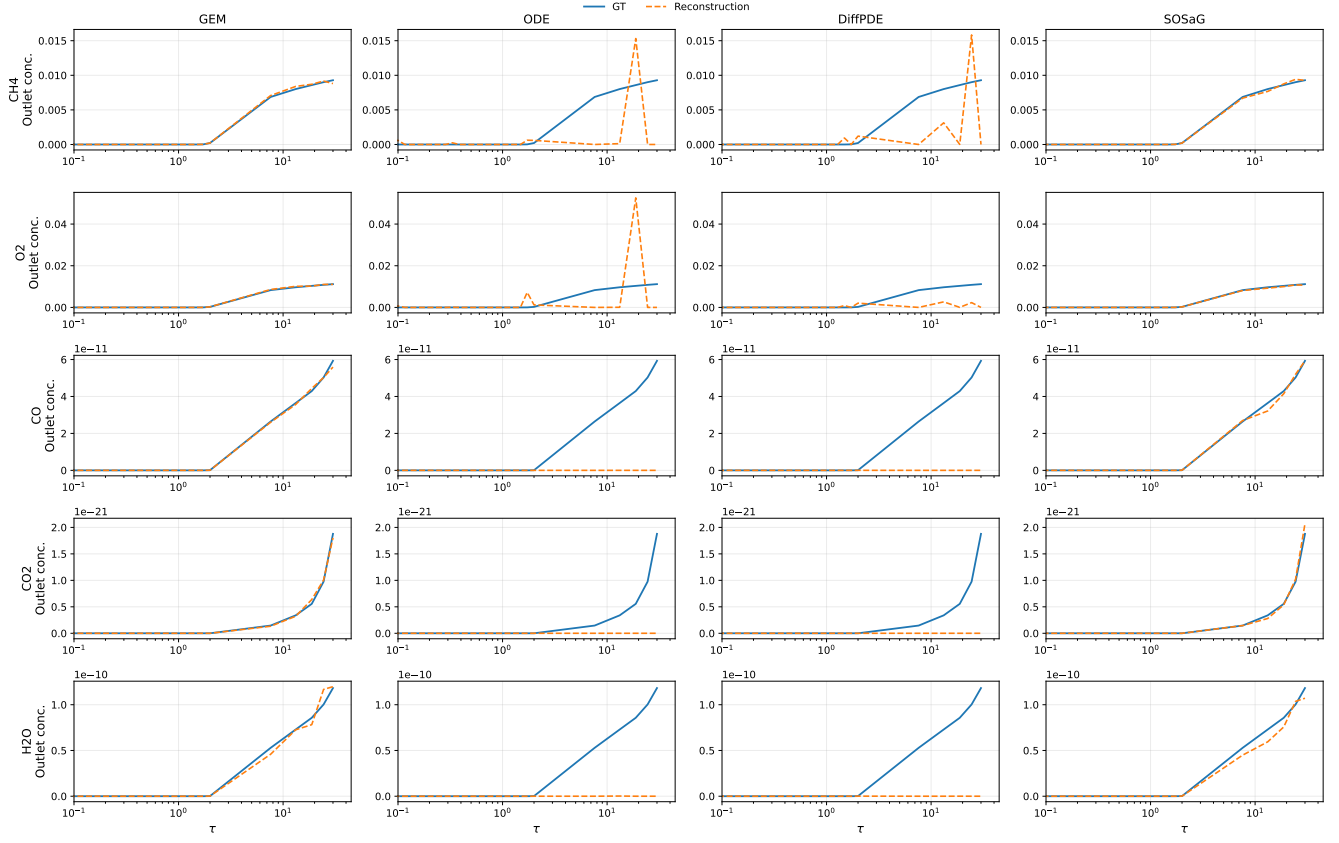


Figure 5: Species outlet comparison for Two-step Methane Oxidation.

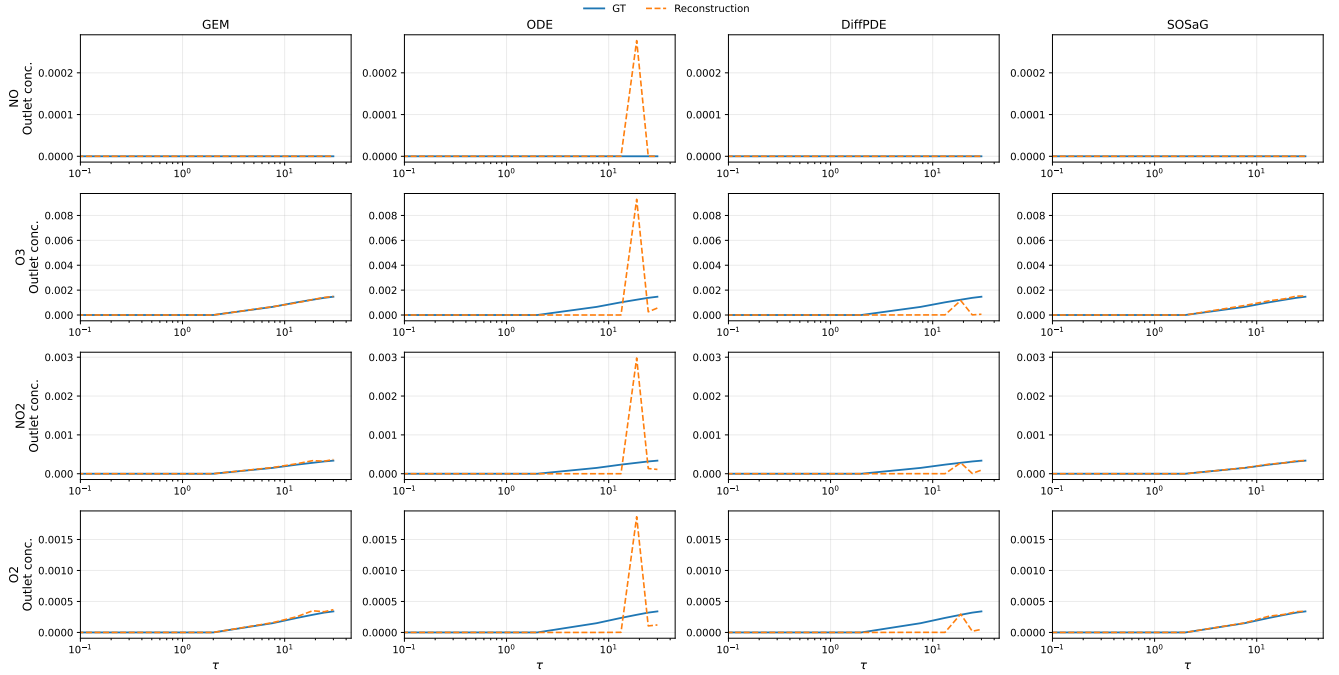


Figure 6: Species outlet comparison for $\text{NO} + \text{O}_3 \rightarrow \text{NO}_2$.

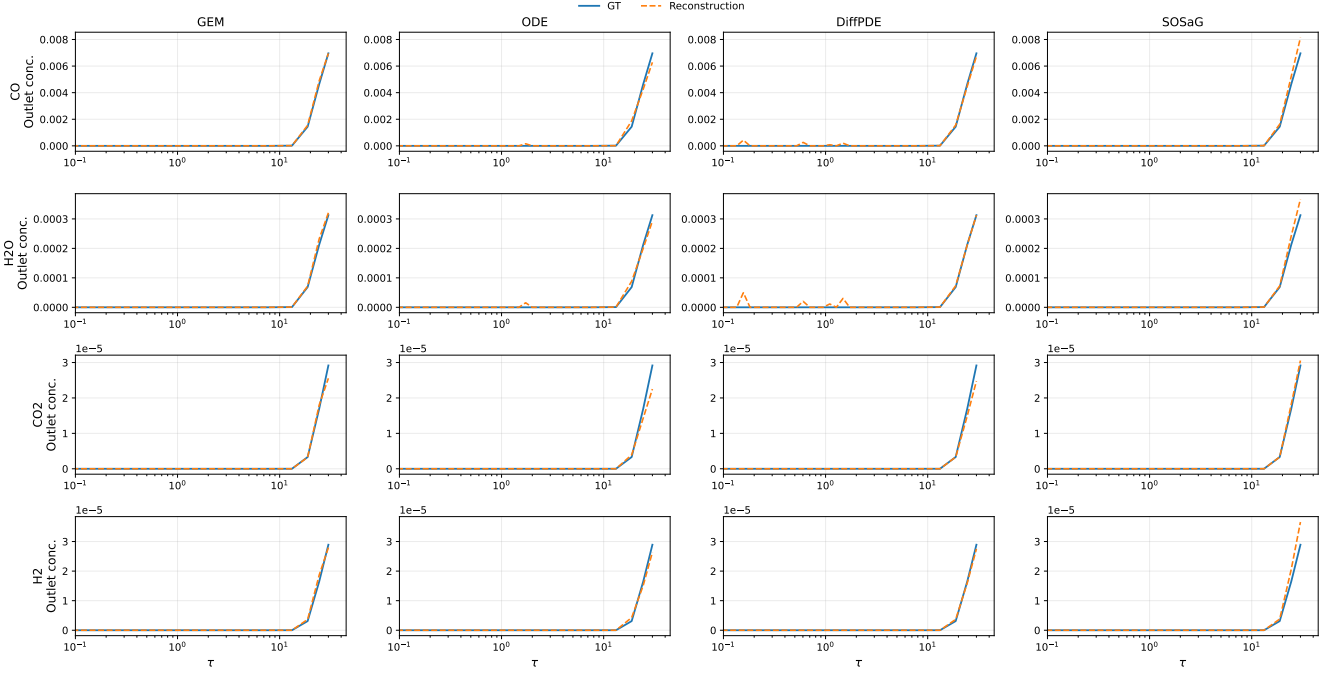


Figure 7: Species outlet comparison for Water Gas Shift.

E.3 Field Reconstruction Figures

Figures 8–31 show an example field reconstructions at selected time snapshots for each method. All methods use the same random seed for the example run.

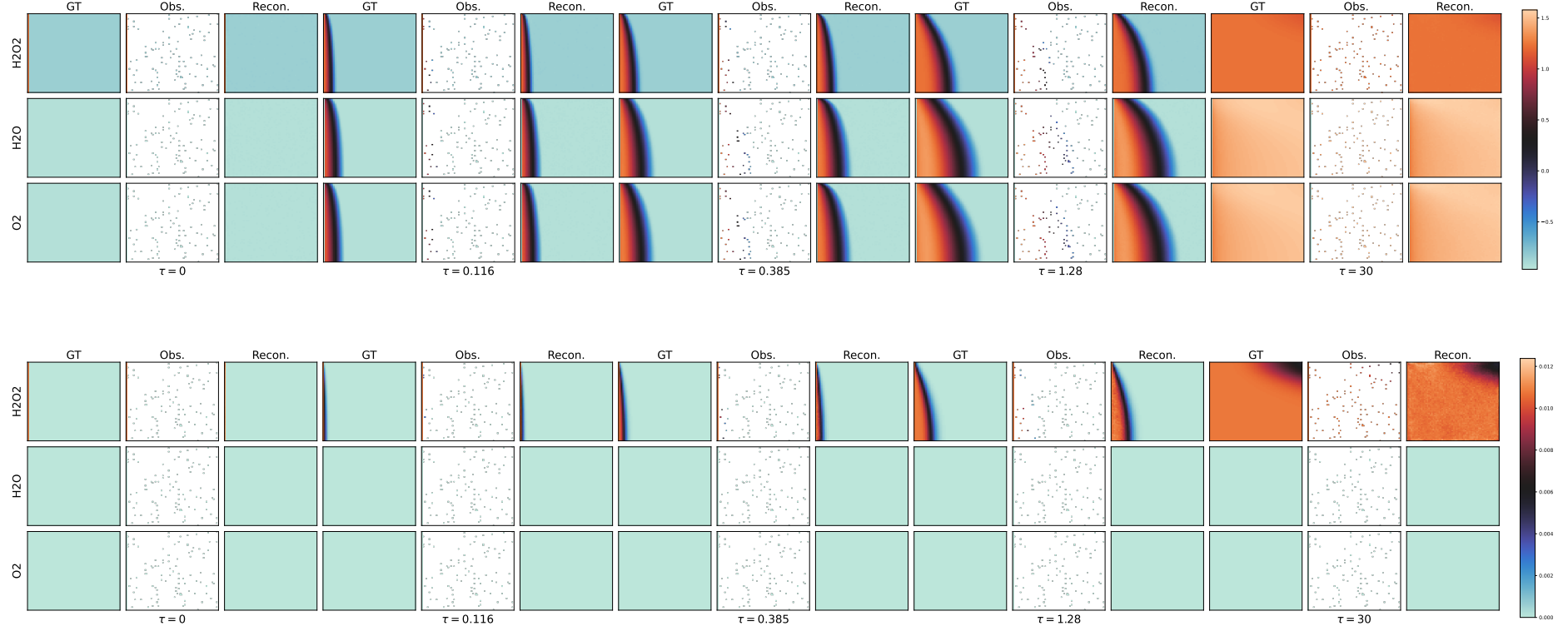


Figure 8: H₂O₂ Decomposition reconstruction panel for the GEM method for select PDE times. Top: Normalised. Bottom: Physical.

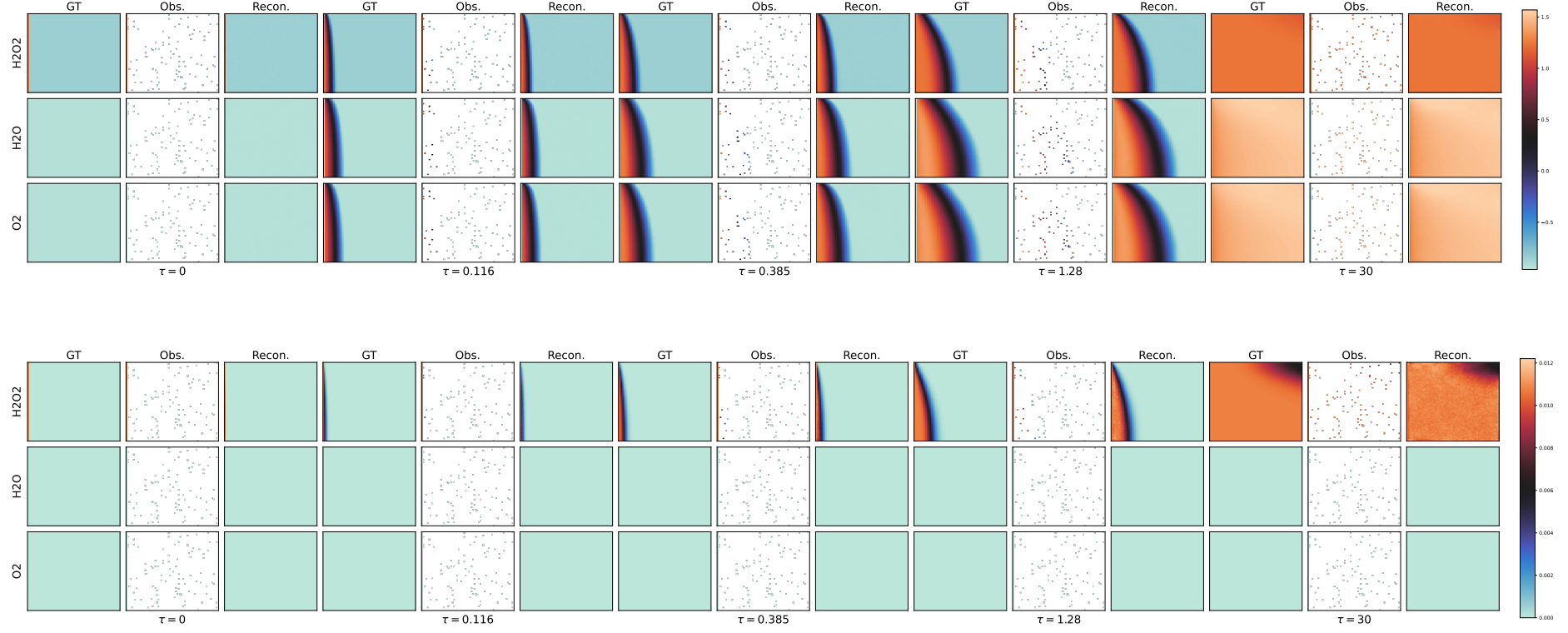


Figure 9: H₂O₂ Decomposition reconstruction panel for the SOSaG method for select PDE times. Top: Normalised. Bottom: Physical.

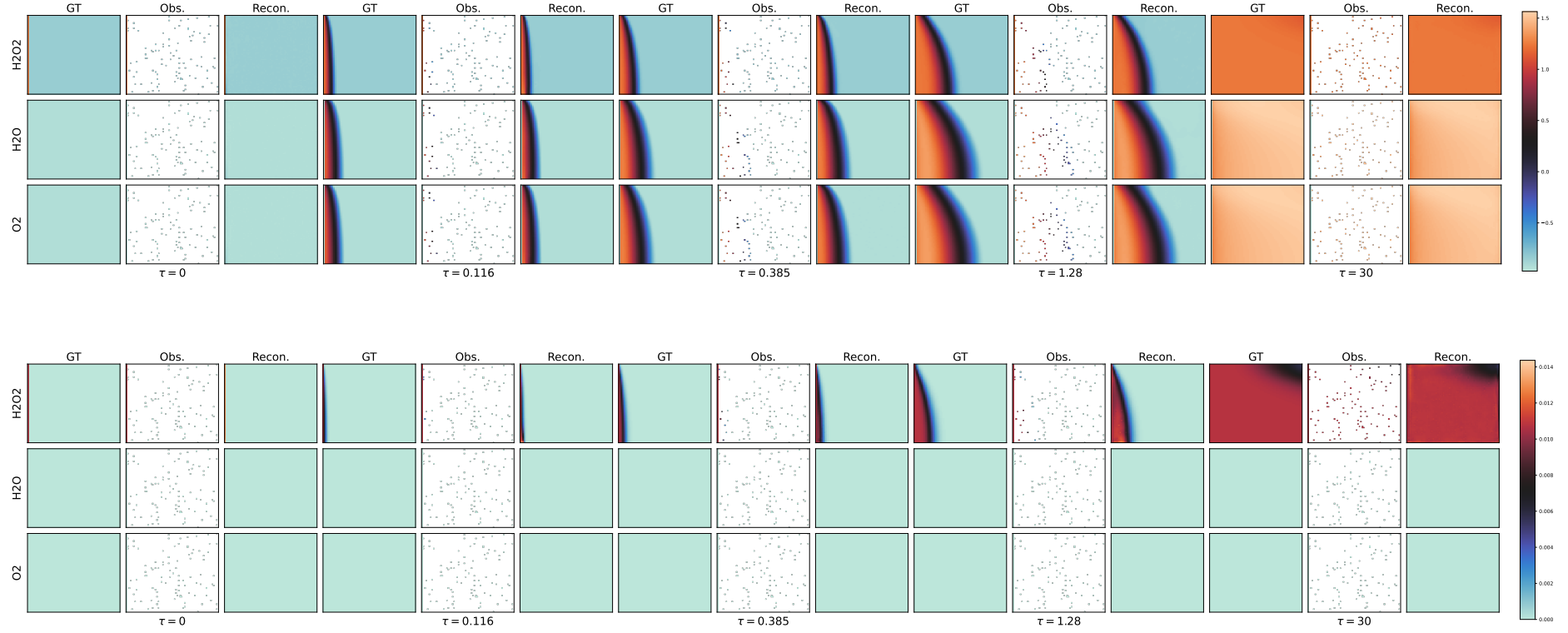


Figure 10: H₂O₂ Decomposition reconstruction panel for the ODE method for select PDE times. Top: Normalised. Bottom: Physical.

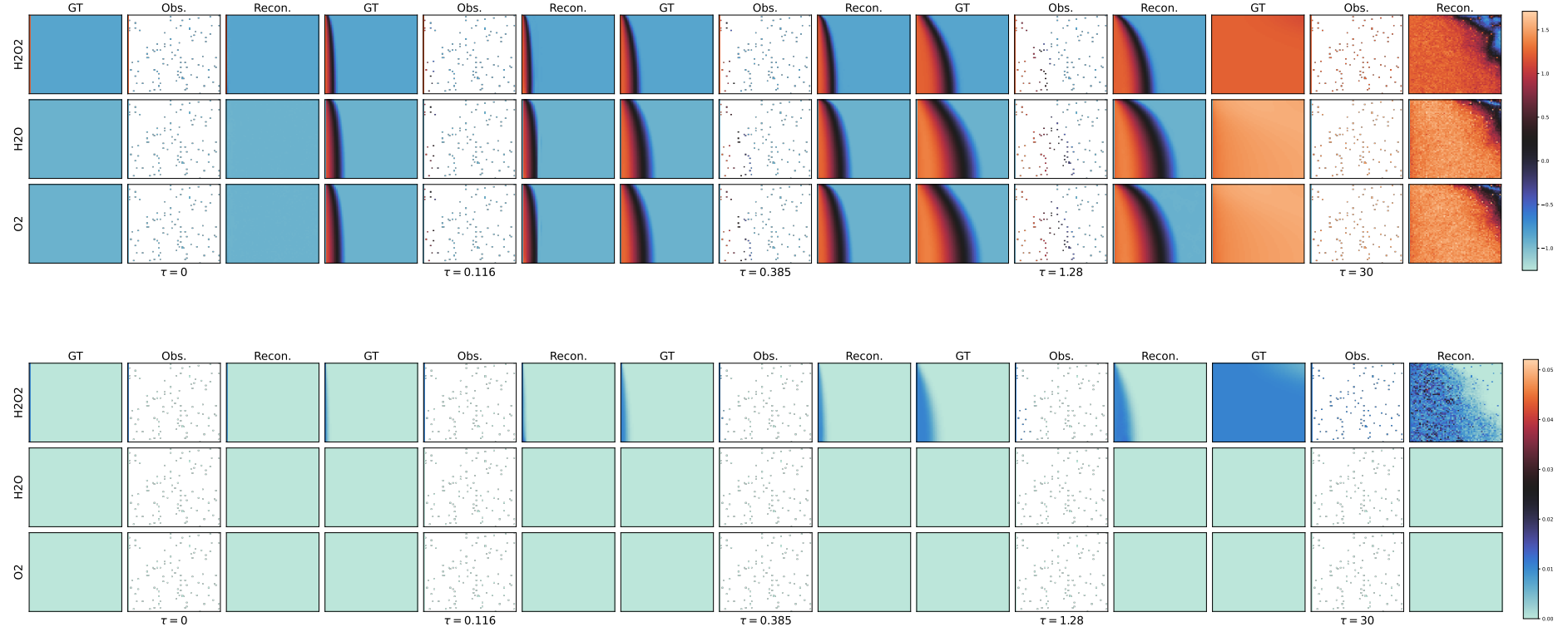


Figure 11: H₂O₂ Decomposition reconstruction panel for the DiffPDE method for select PDE times. Top: Normalised. Bottom: Physical.

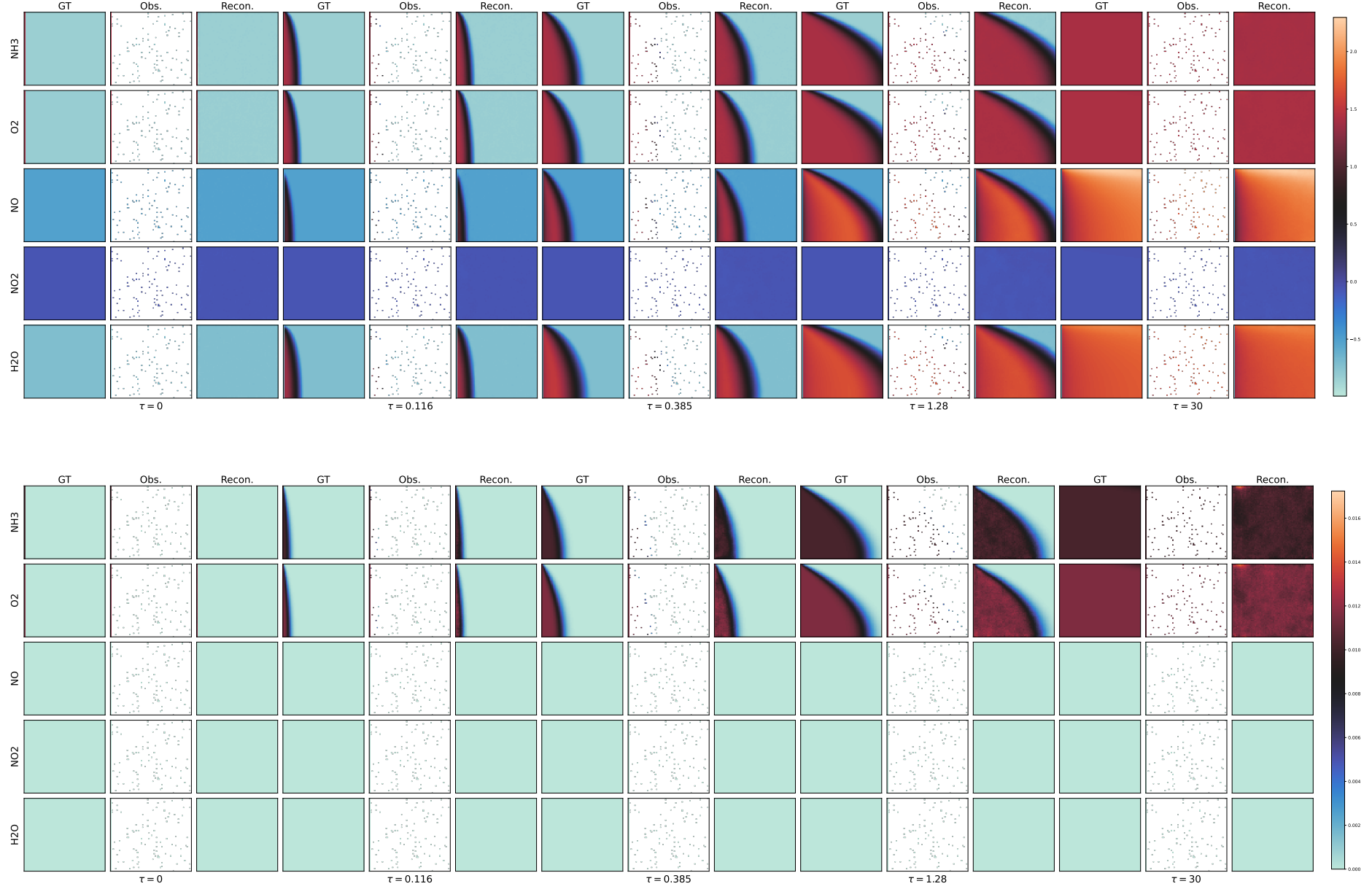


Figure 12: Ammonia Oxidation reconstruction panel for the GEM method for select PDE times. Top: Normalised. Bottom: Physical.

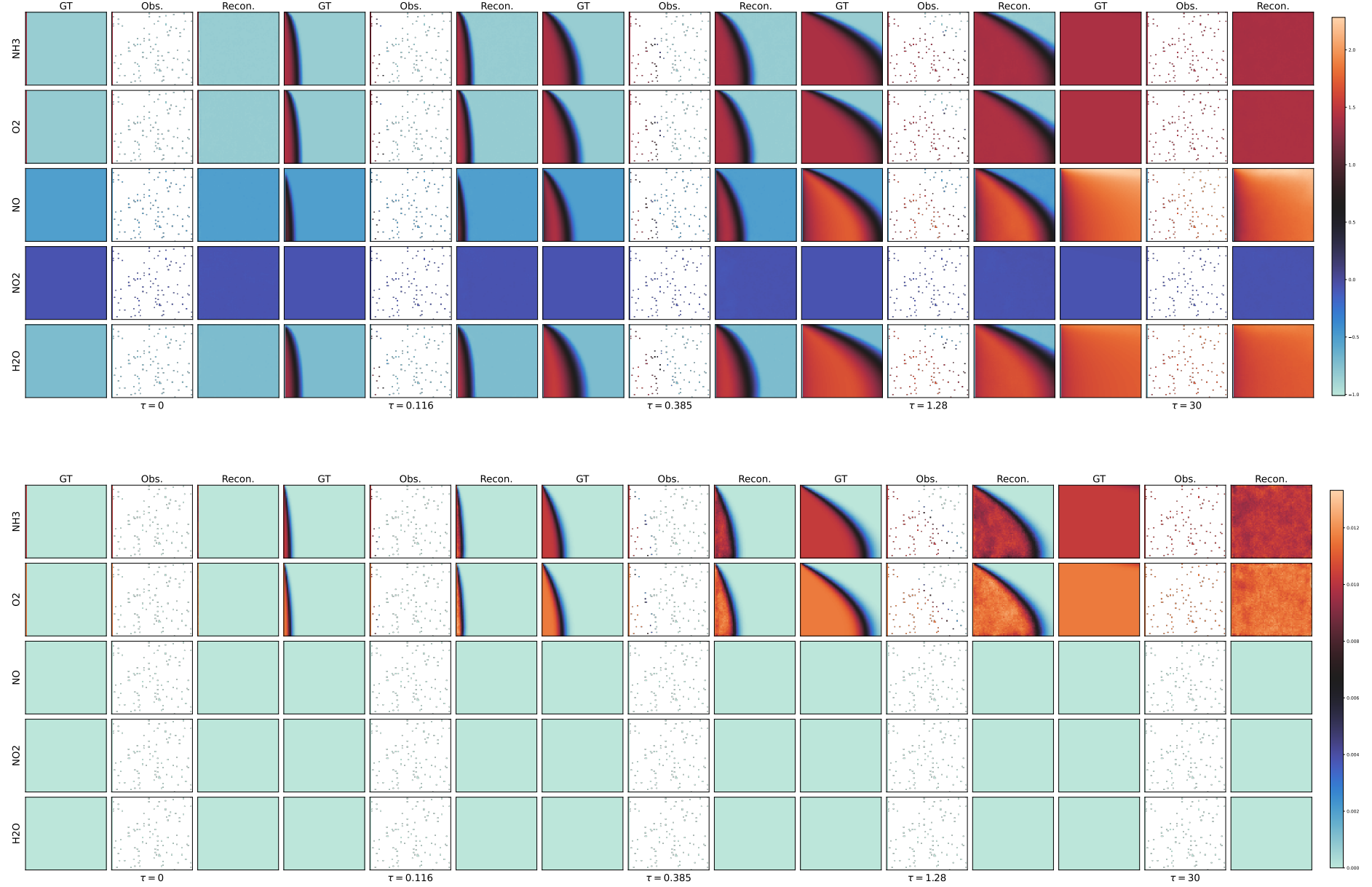


Figure 13: Ammonia Oxidation reconstruction panel for the SOSaG method for select PDE times. Top: Normalised. Bottom: Physical.

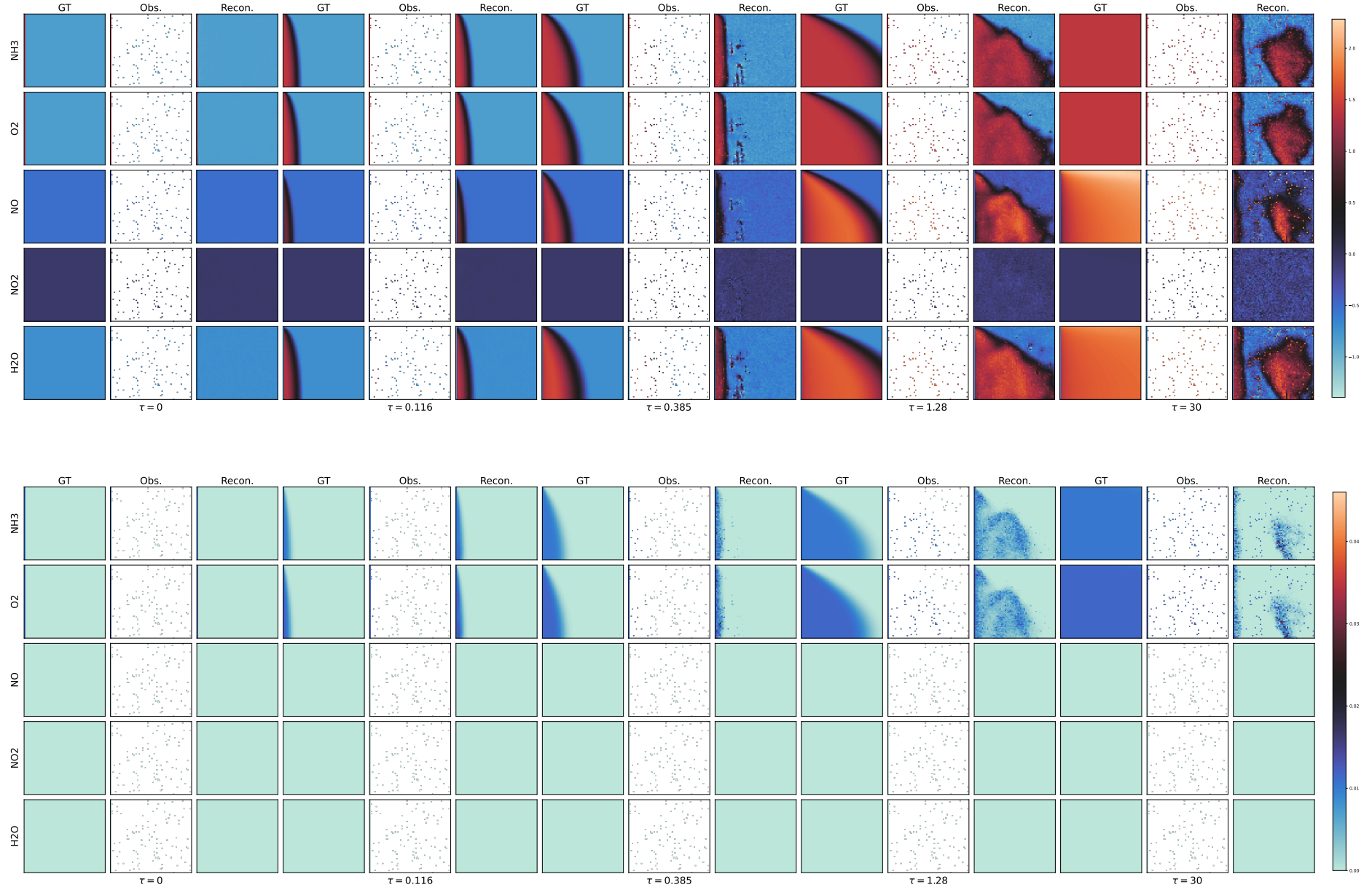


Figure 14: Ammonia Oxidation reconstruction panel for the ODE method for select PDE times. Top: Normalised. Bottom: Physical.

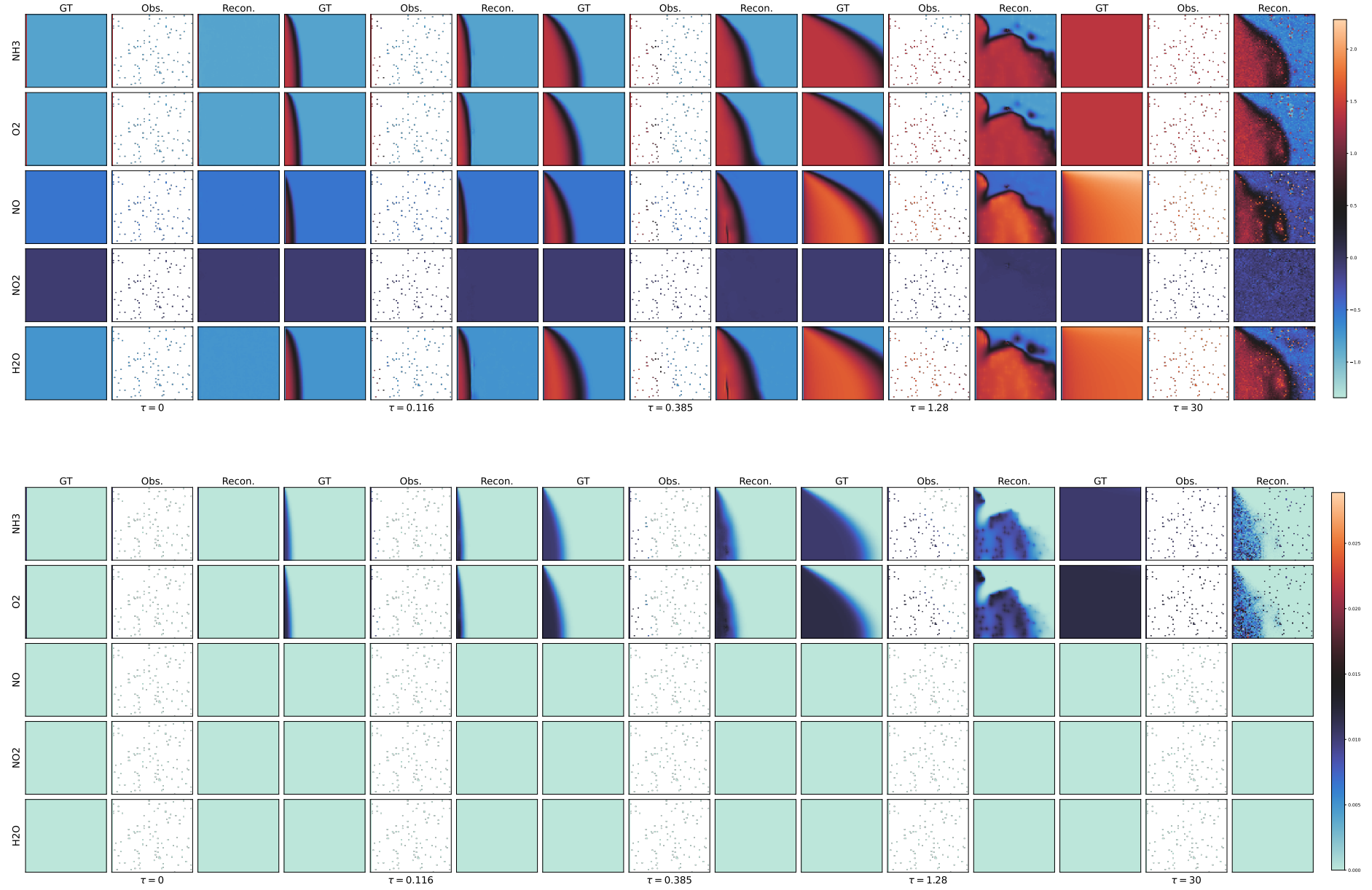


Figure 15: Ammonia Oxidation reconstruction panel for the DiffPDE method for select PDE times. Top: Normalised. Bottom: Physical.

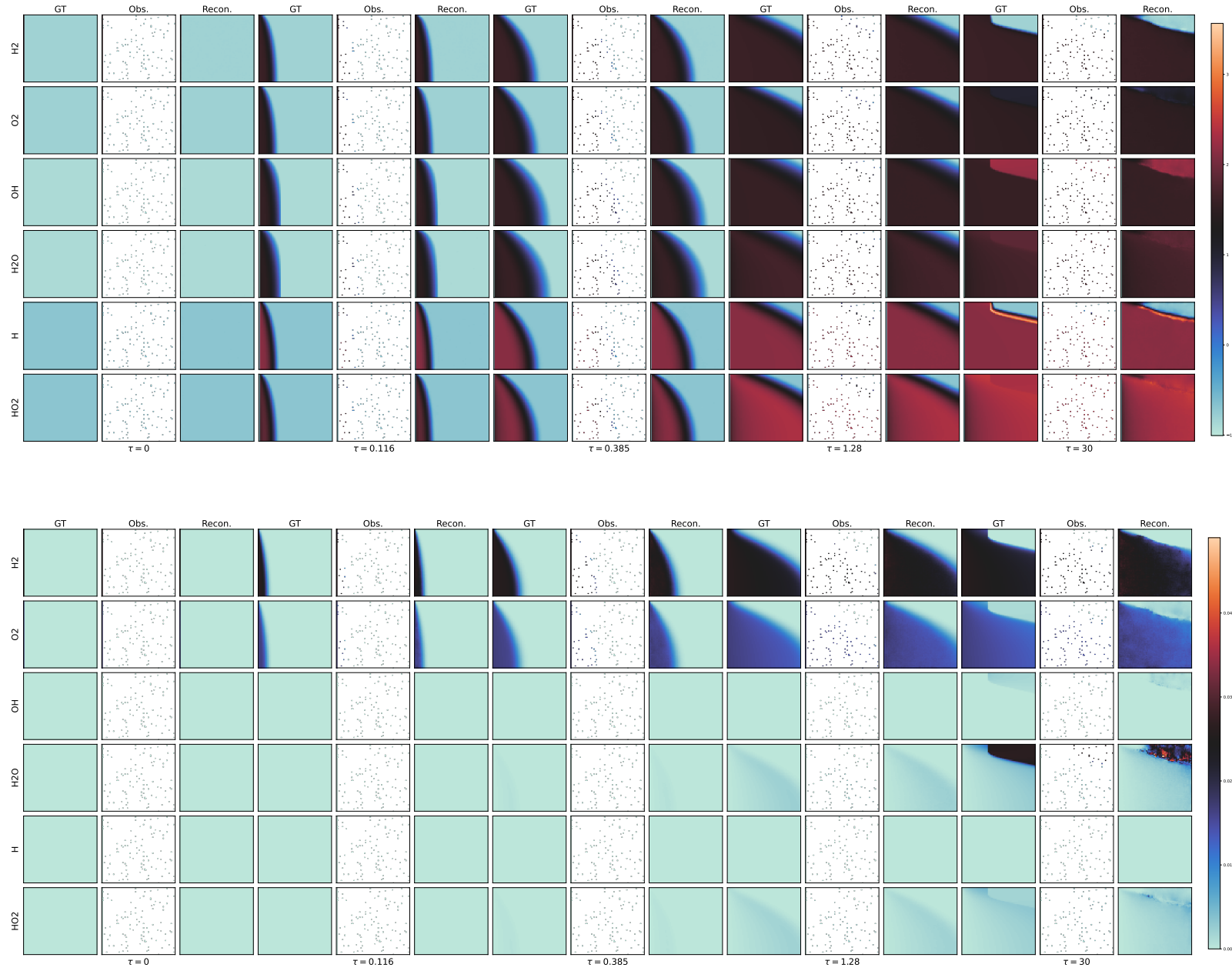


Figure 16: Hydrogen Oxidation Subset reconstruction panel for the GEM method for select PDE times. Top: Normalised. Bottom: Physical.

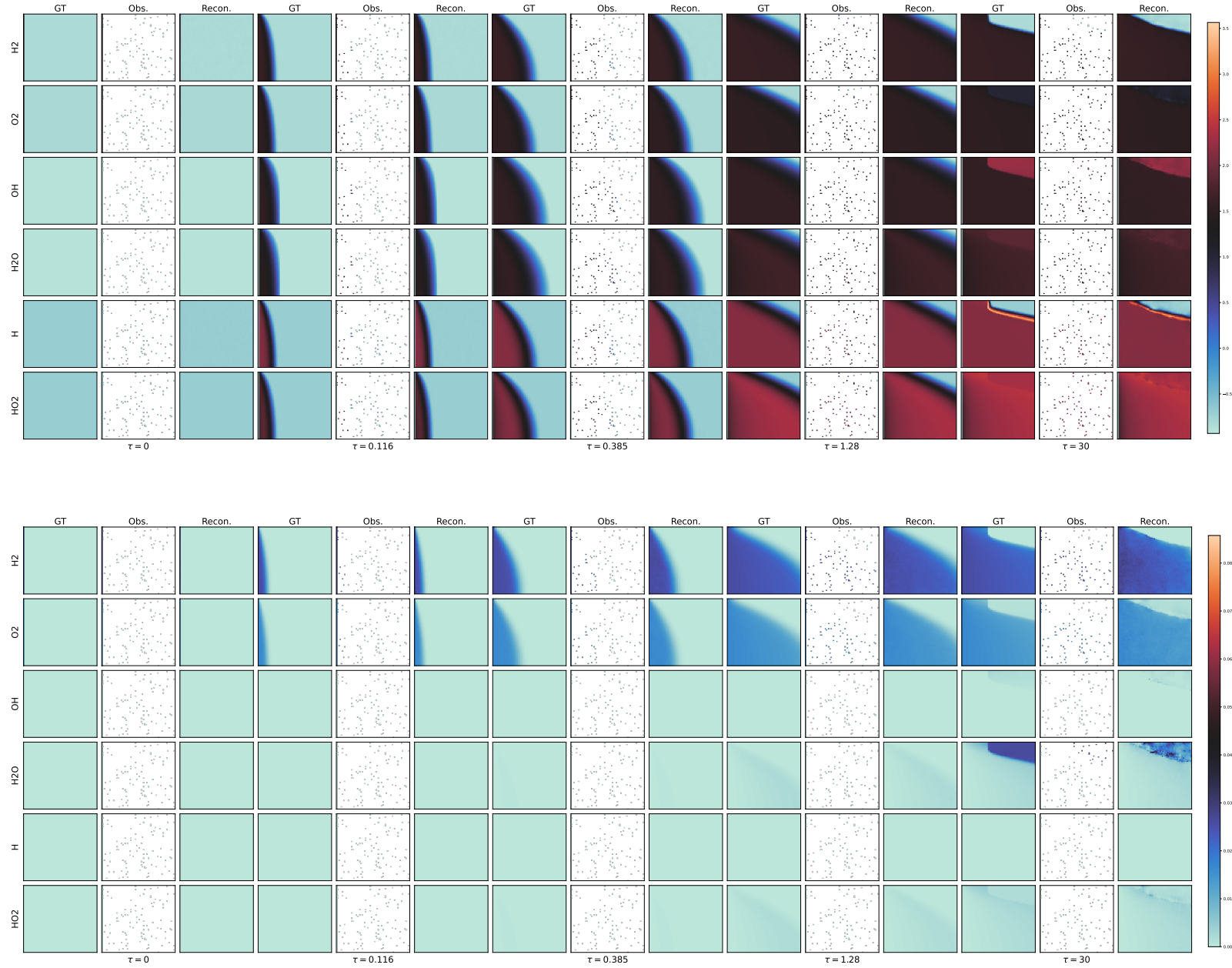


Figure 17: Hydrogen Oxidation Subset reconstruction panel for the SOSaG method for select PDE times. Top: Normalised. Bottom: Physical.

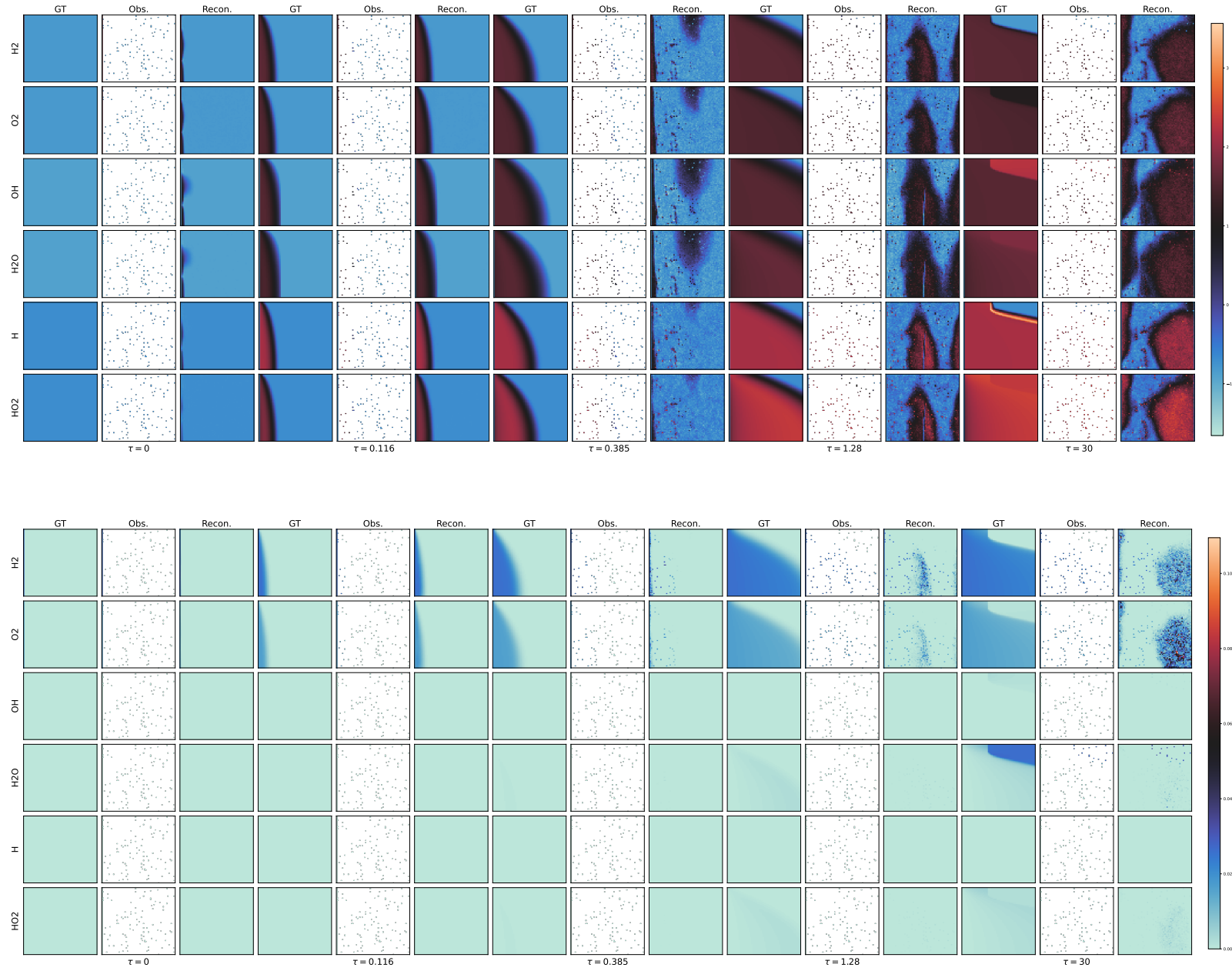


Figure 18: Hydrogen Oxidation Subset reconstruction panel for the ODE method for select PDE times. Top: Normalised. Bottom: Physical.

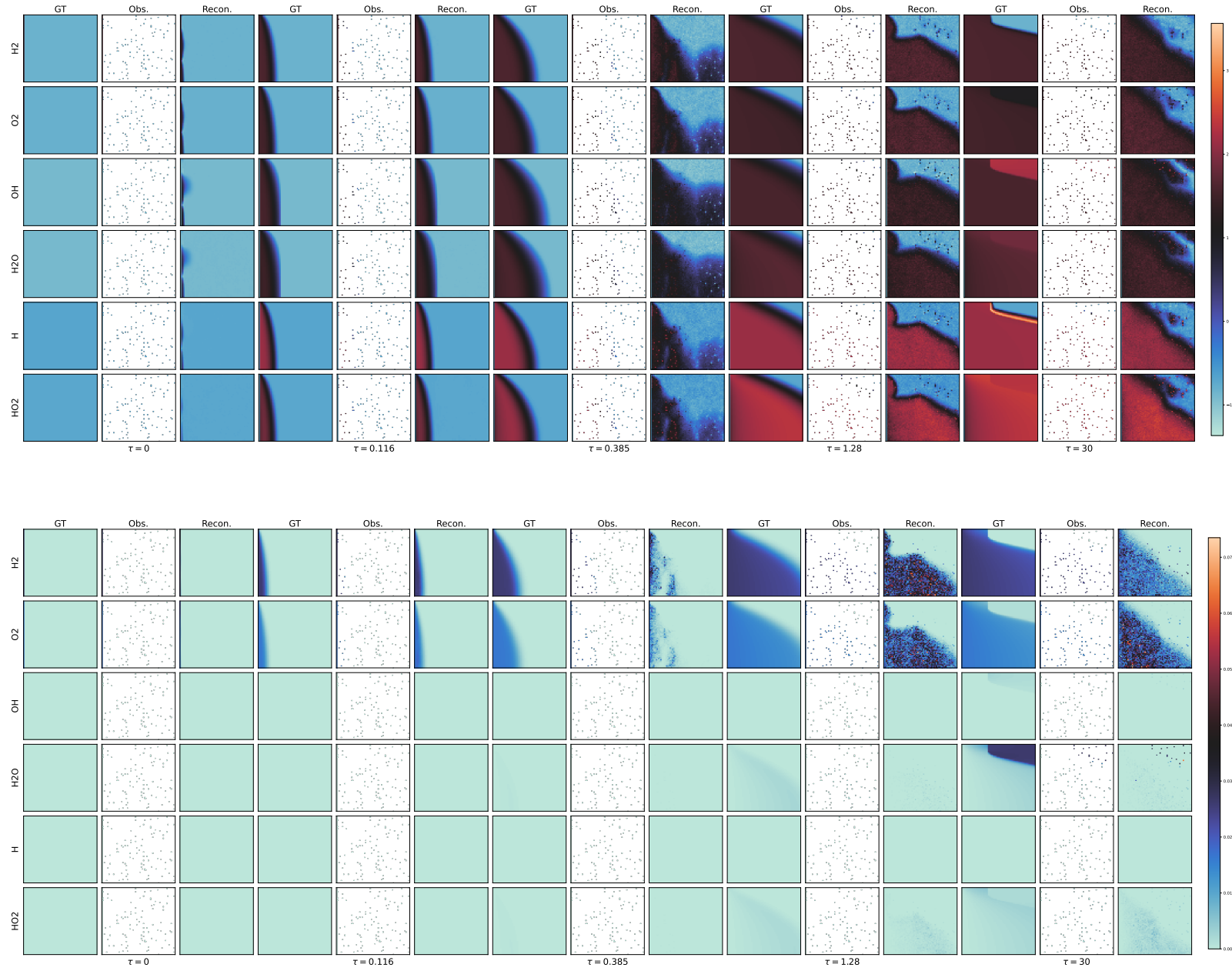


Figure 19: Hydrogen Oxidation Subset reconstruction panel for the DiffPDE method for select PDE times. Top: Normalised. Bottom: Physical.

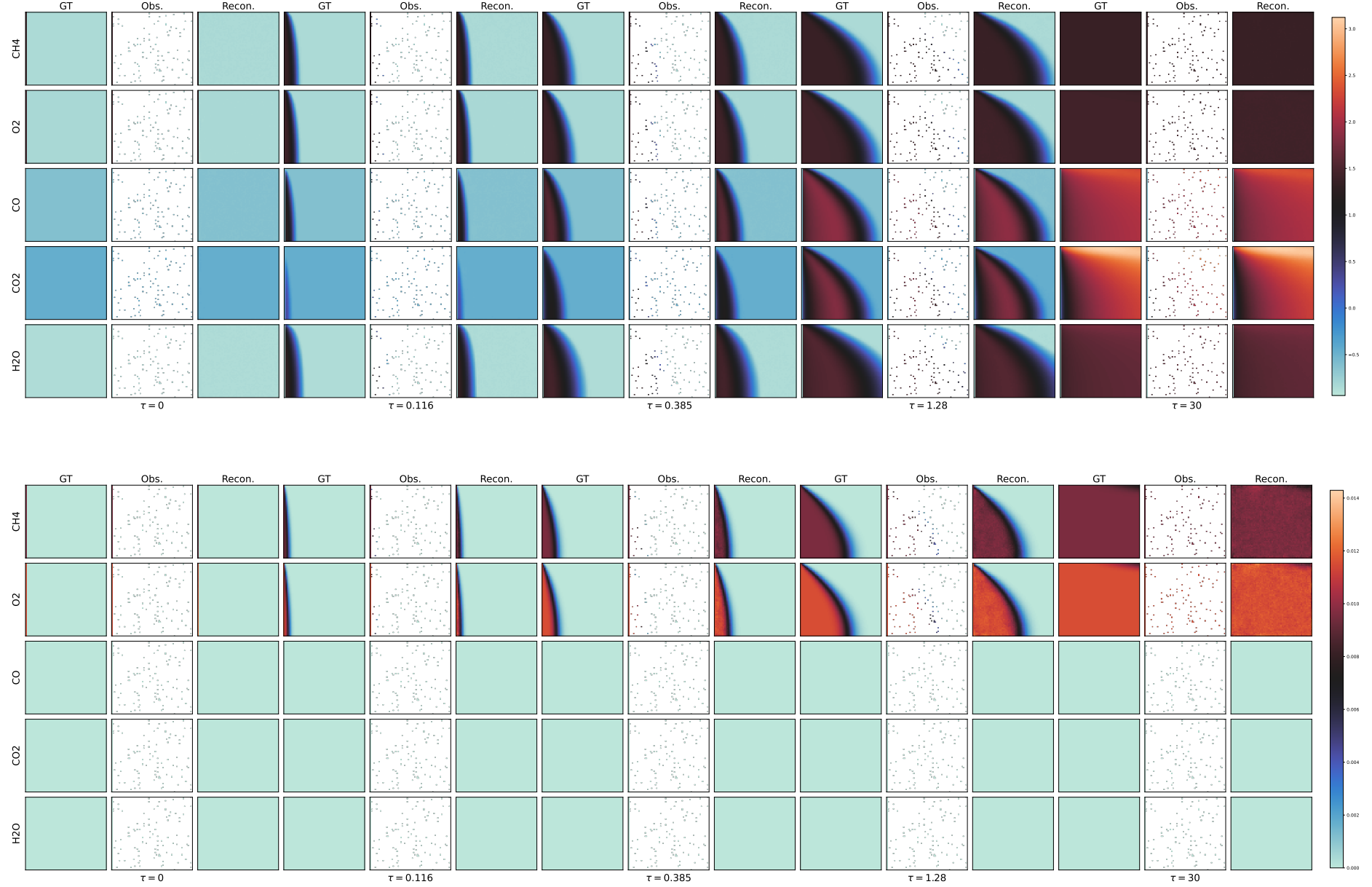


Figure 20: Two-step Methane Oxidation Subset reconstruction panel for the GEM method for select PDE times. Top: Normalised. Bottom: Physical.

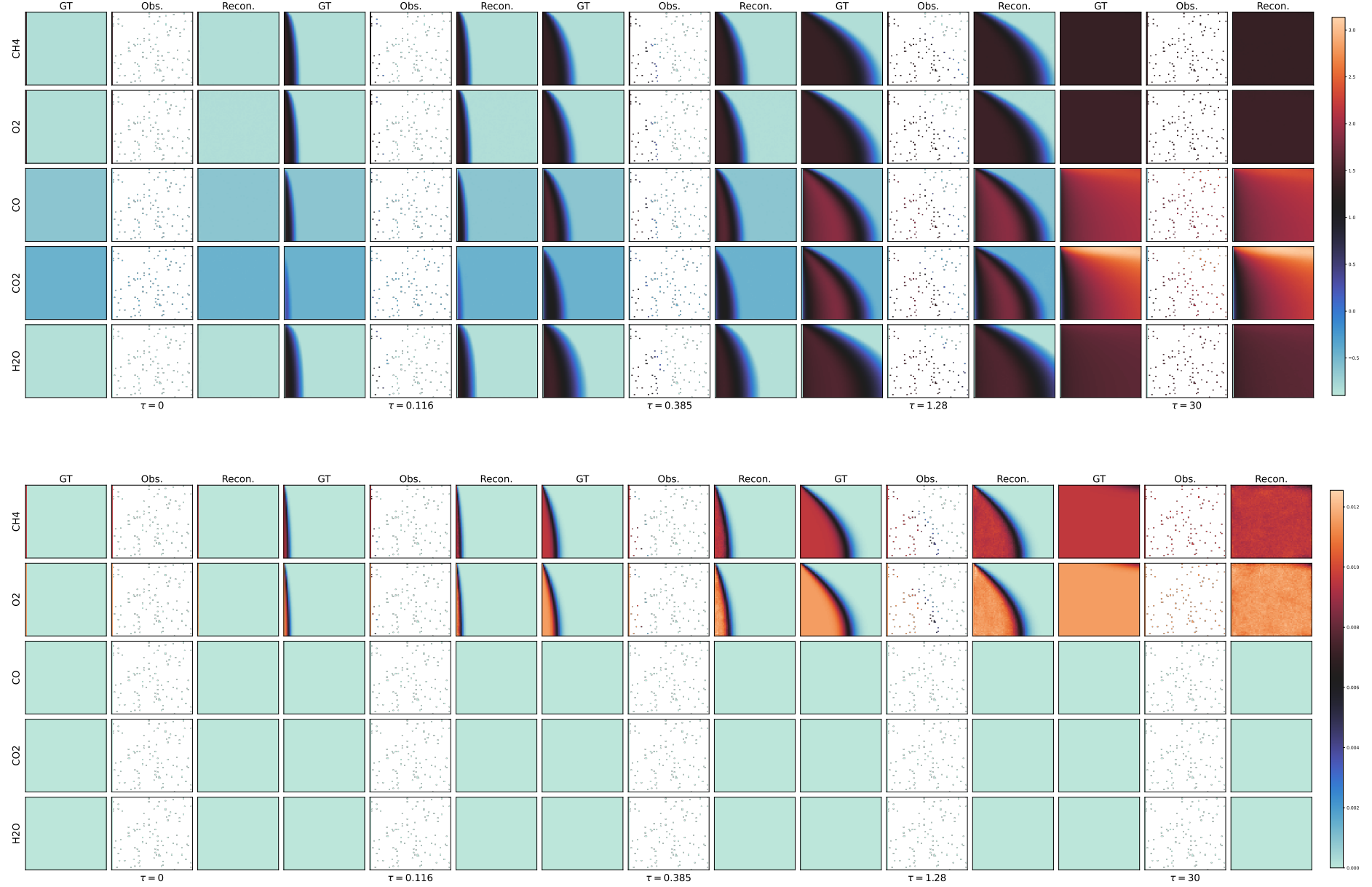


Figure 21: Two-step Methane Oxidation Subset reconstruction panel for the SOSaG method for select PDE times. Top: Normalised. Bottom: Physical.

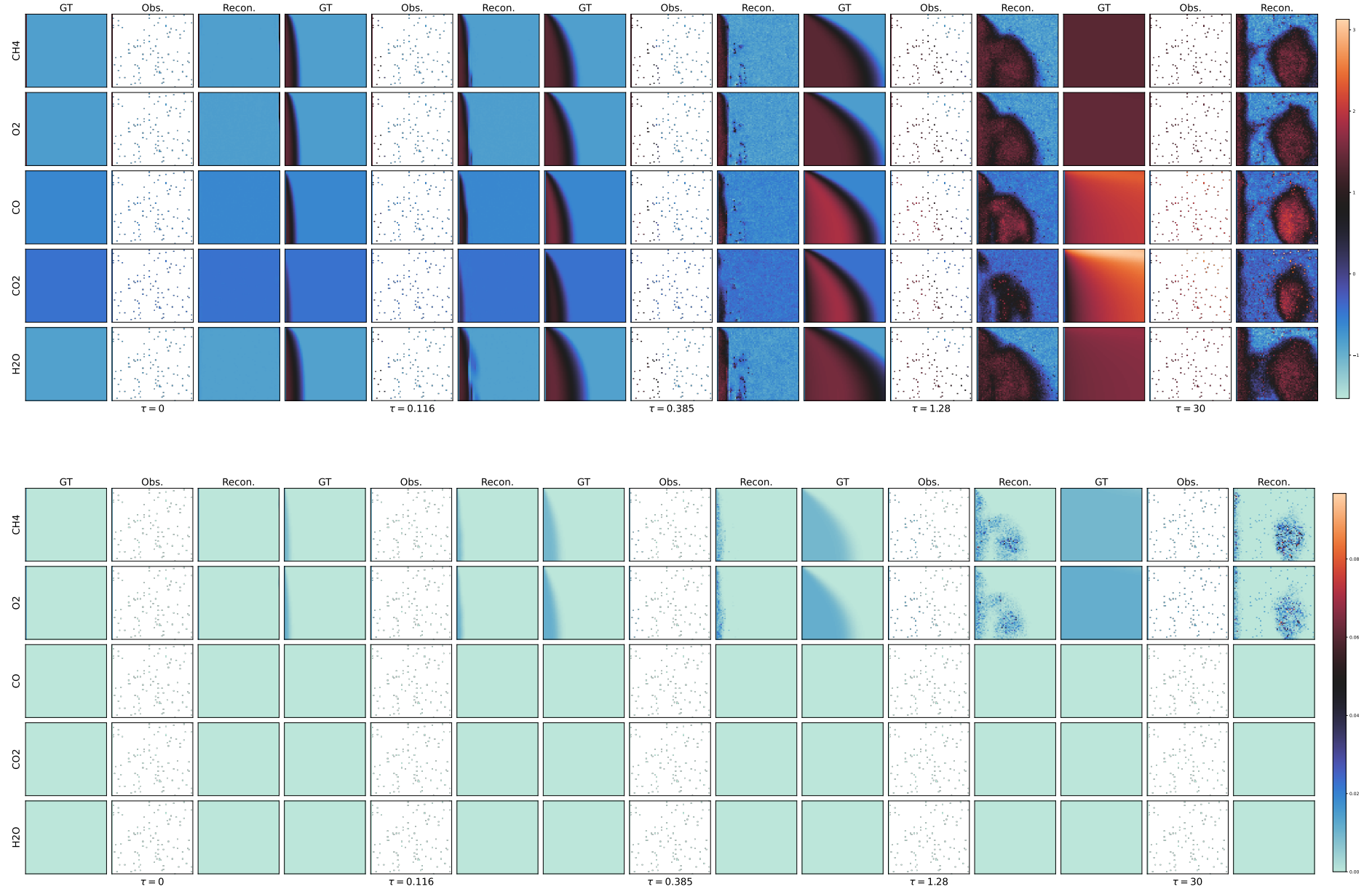


Figure 22: Two-step Methane Oxidation Subset reconstruction panel for the ODE method for select PDE times. Top: Normalised. Bottom: Physical.

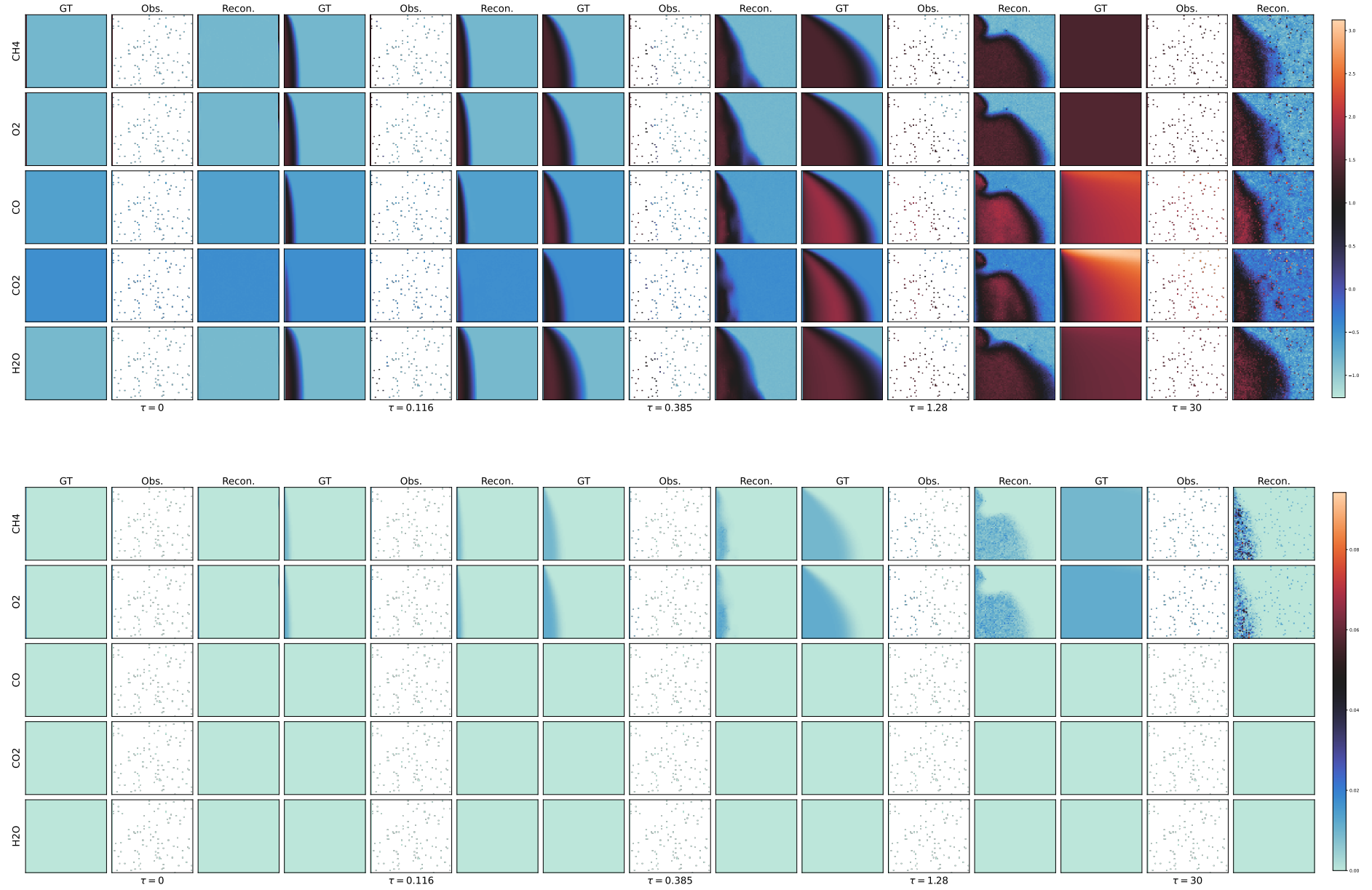


Figure 23: Two-step Methane Oxidation Subset reconstruction panel for the DiffPDE method for select PDE times. Top: Normalised. Bottom: Physical.

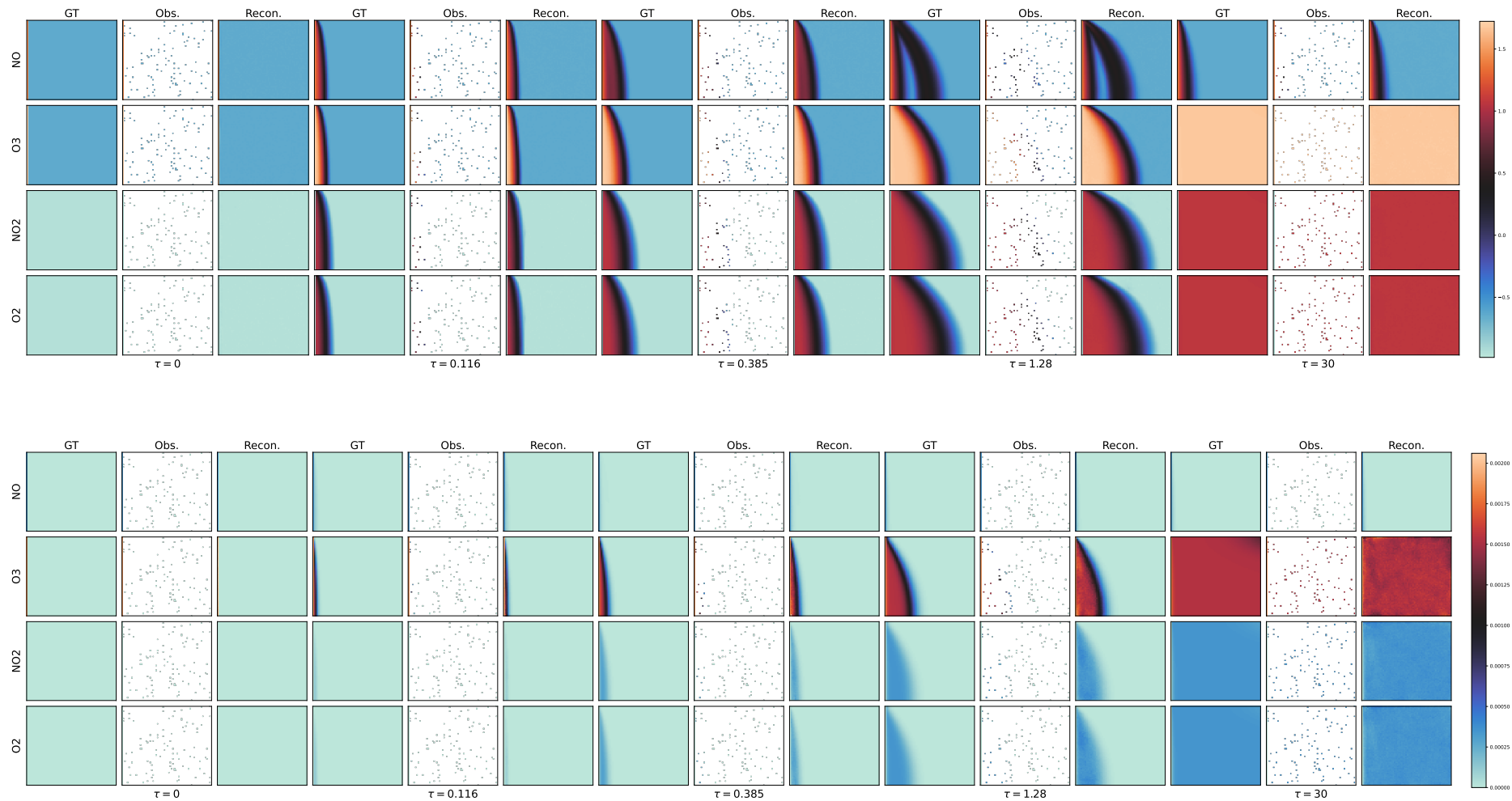


Figure 24: $\text{NO} + \text{O}_3 \rightarrow \text{NO}_2$ reconstruction panel for the GEM method for select PDE times. Top: Normalised. Bottom: Physical.

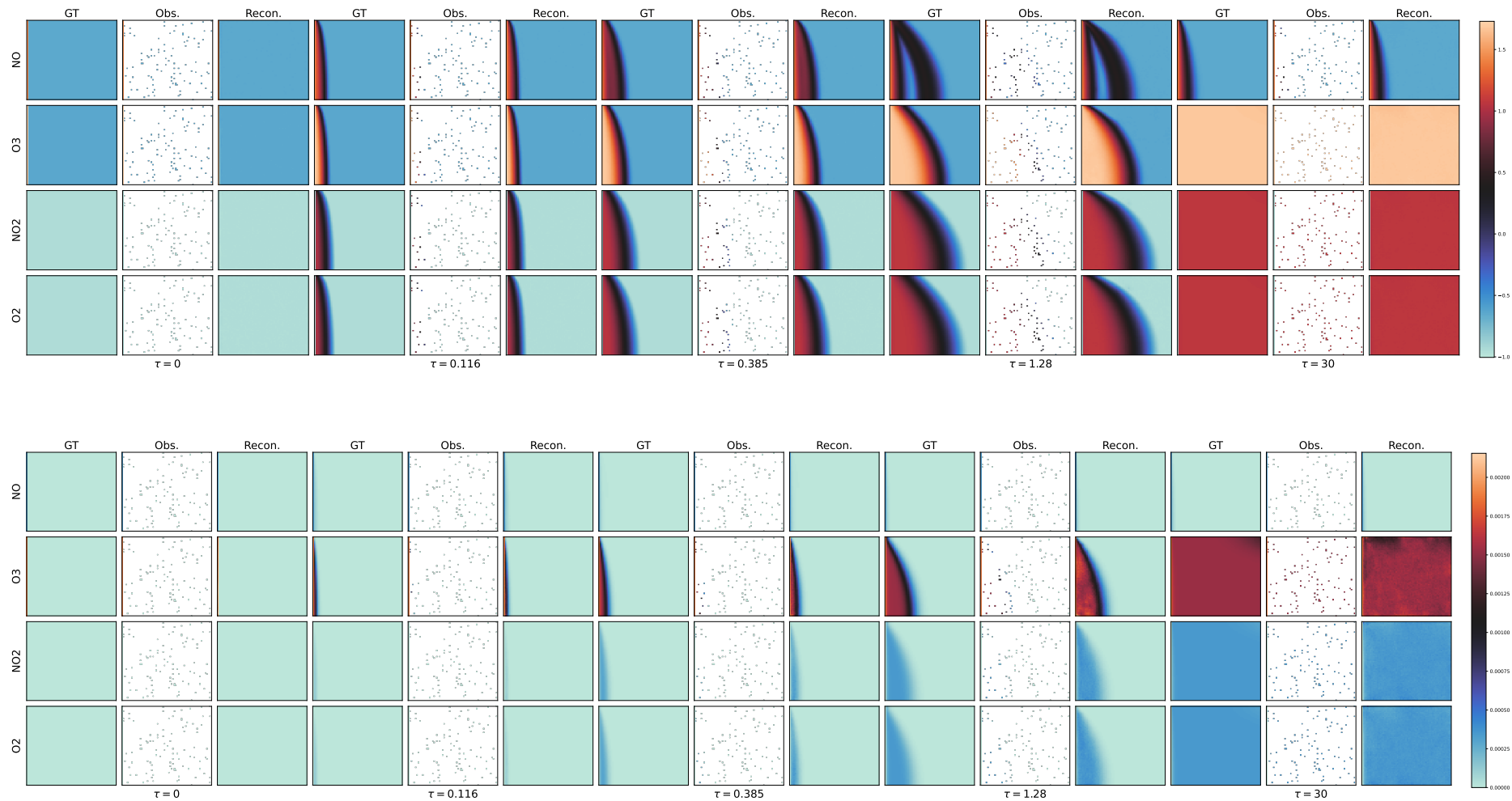


Figure 25: $\text{NO} + \text{O}_3 \rightarrow \text{NO}_2$ reconstruction panel for the SOSaG method for select PDE times. Top: Normalised. Bottom: Physical.

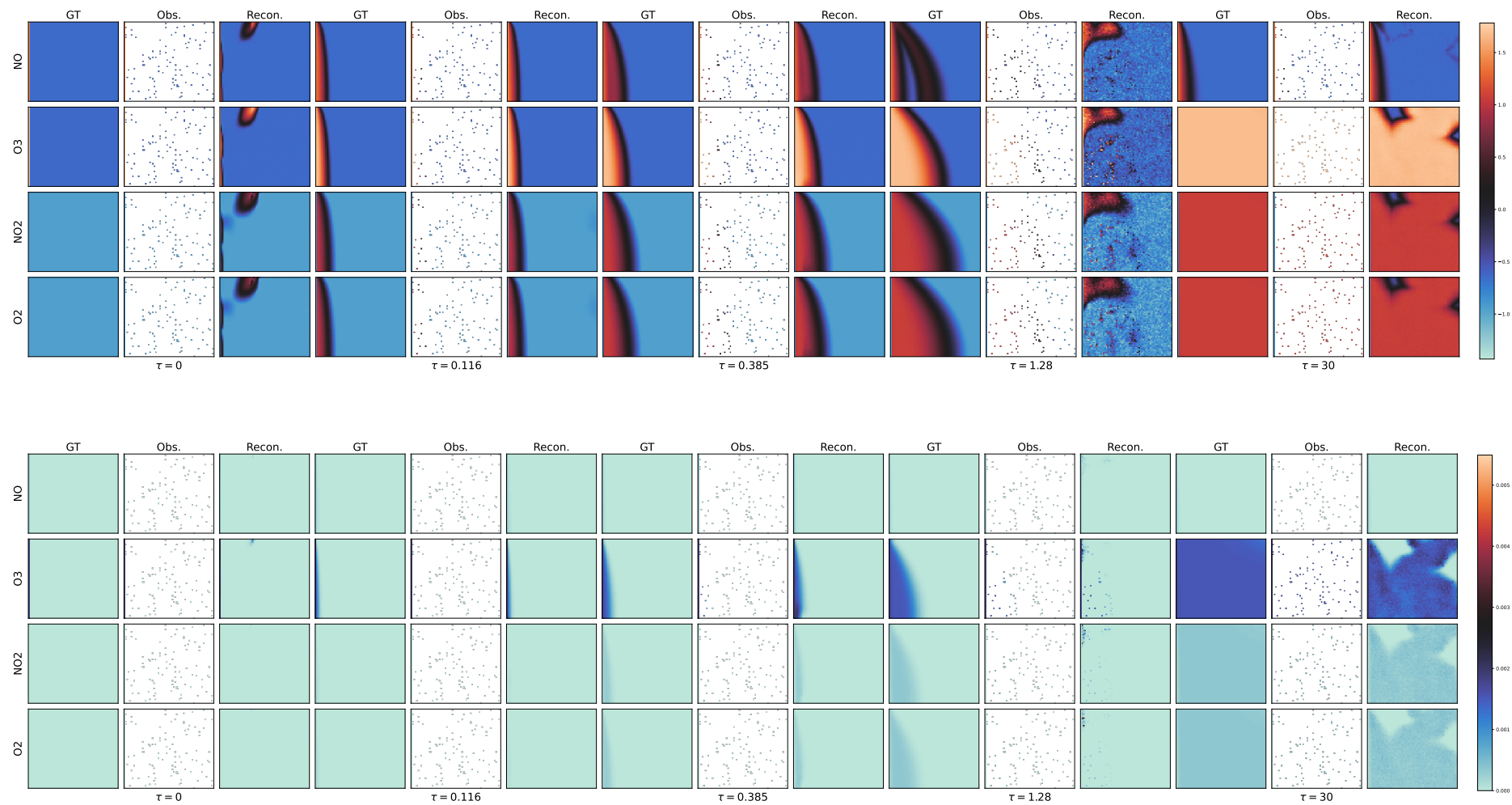


Figure 26: $\text{NO} + \text{O}_3 \rightarrow \text{NO}_2$ reconstruction panel for the ODE method for select PDE times. Top: Normalised. Bottom: Physical.

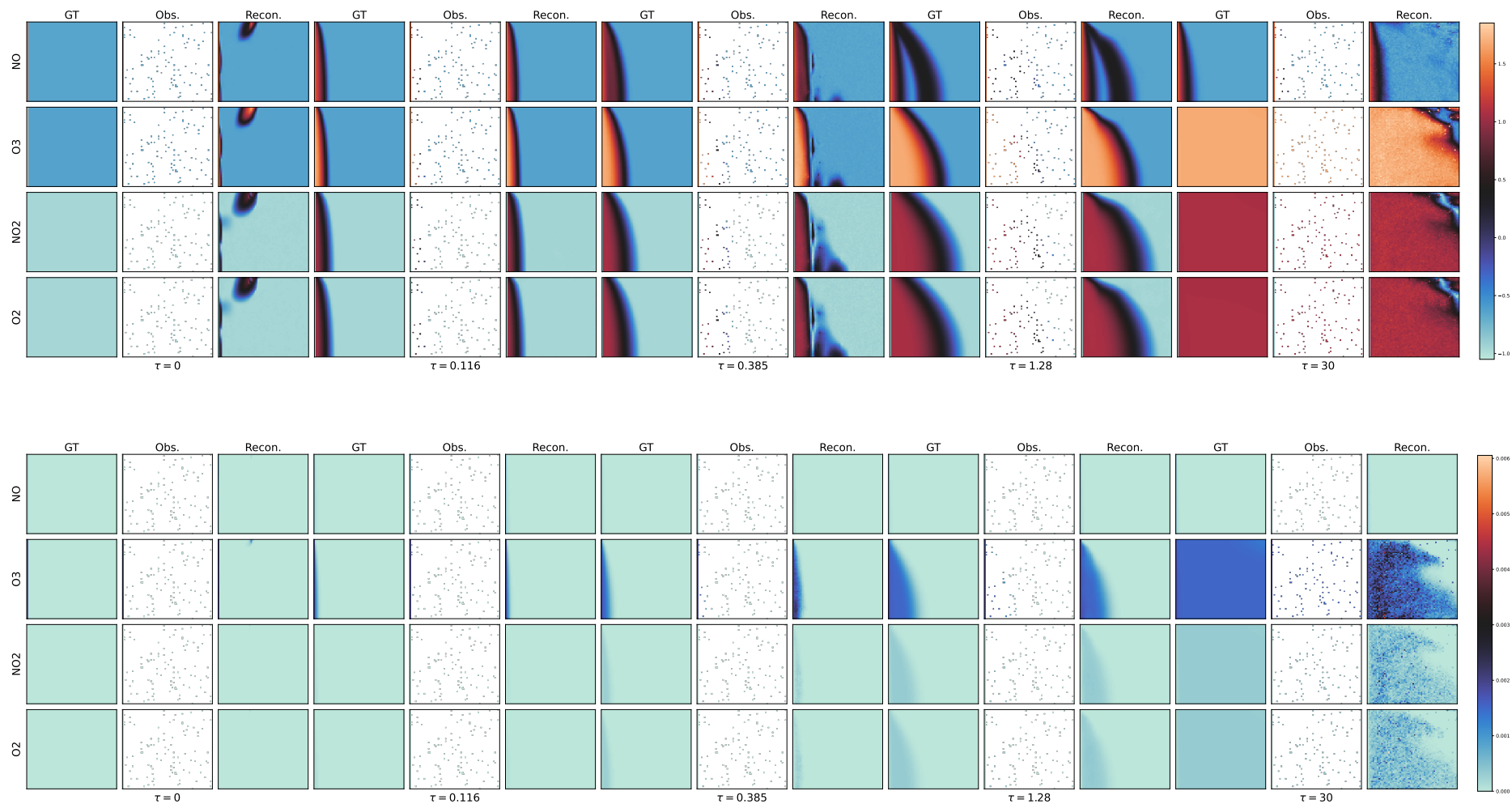


Figure 27: $\text{NO} + \text{O}_3 \rightarrow \text{NO}_2$ reconstruction panel for the DiffPDE method for select PDE times. Top: Normalised. Bottom: Physical.

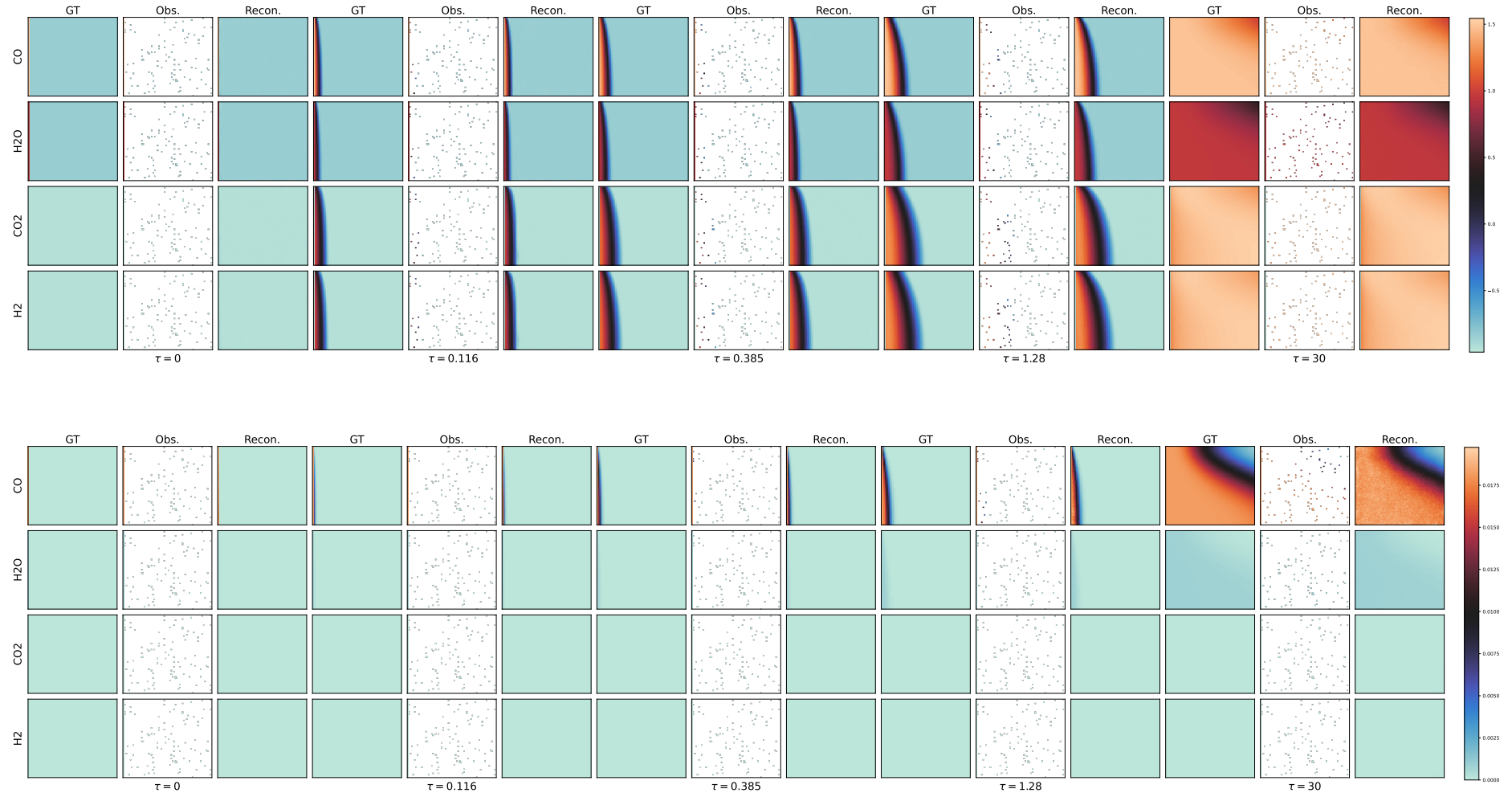


Figure 28: Water Gas Shift reconstruction panel for the GEM method for select PDE times. Top: Normalised. Bottom: Physical.

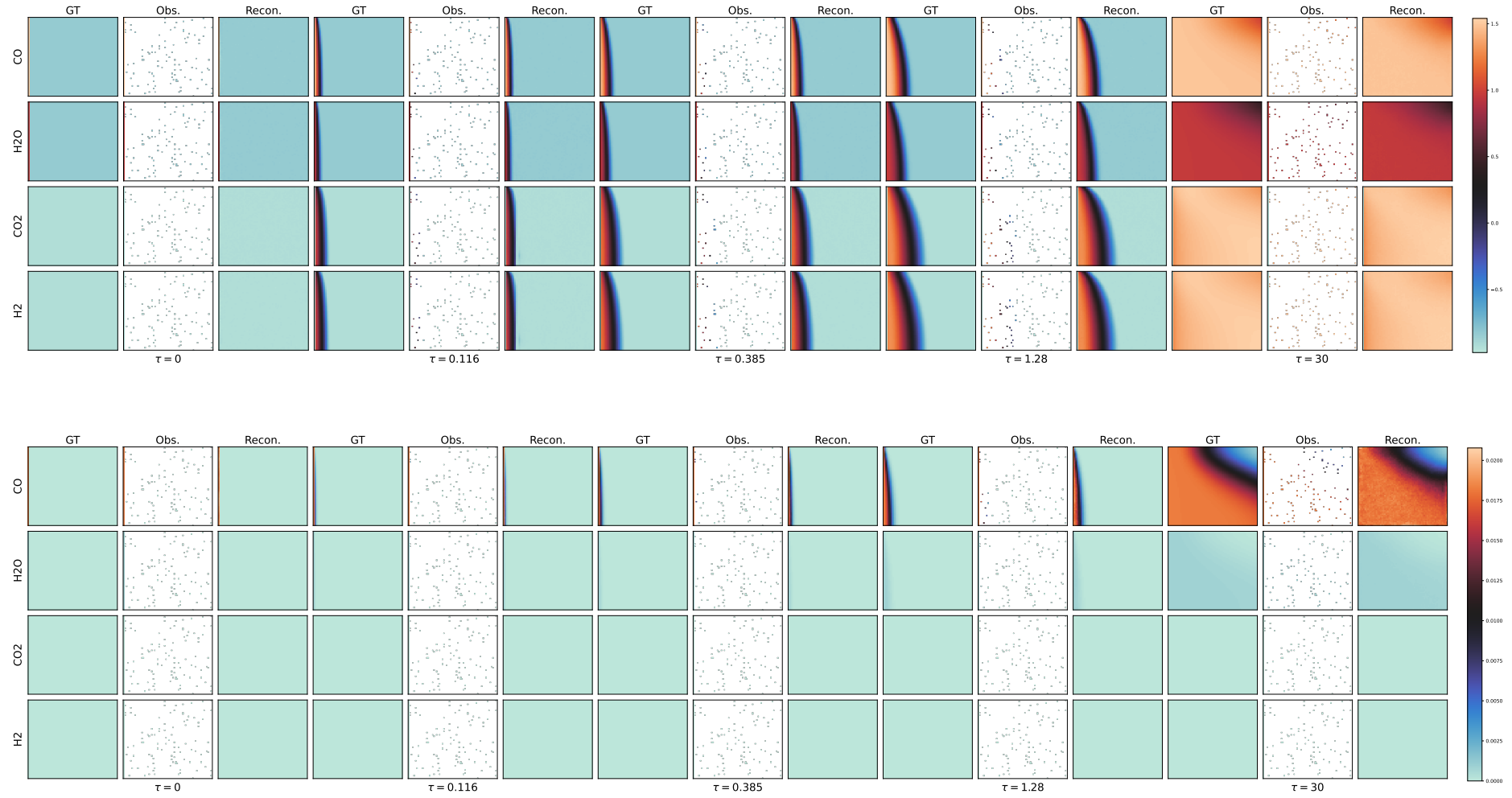


Figure 29: Water Gas Shift reconstruction panel for the SOSaG method for select PDE times. Top: Normalised. Bottom: Physical.

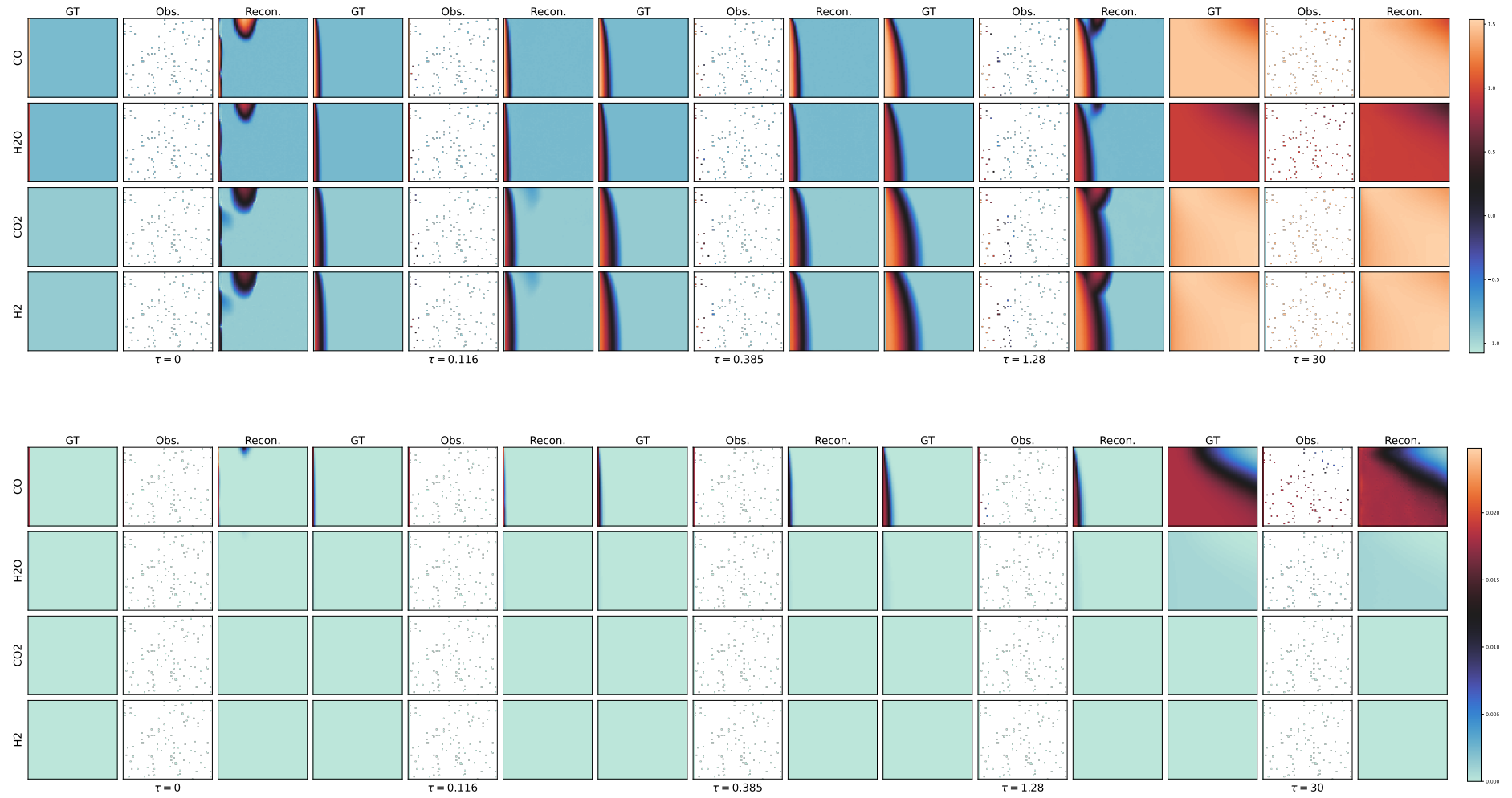


Figure 30: Water Gas Shift reconstruction panel for the ODE method for select PDE times. Top: Normalised. Bottom: Physical.

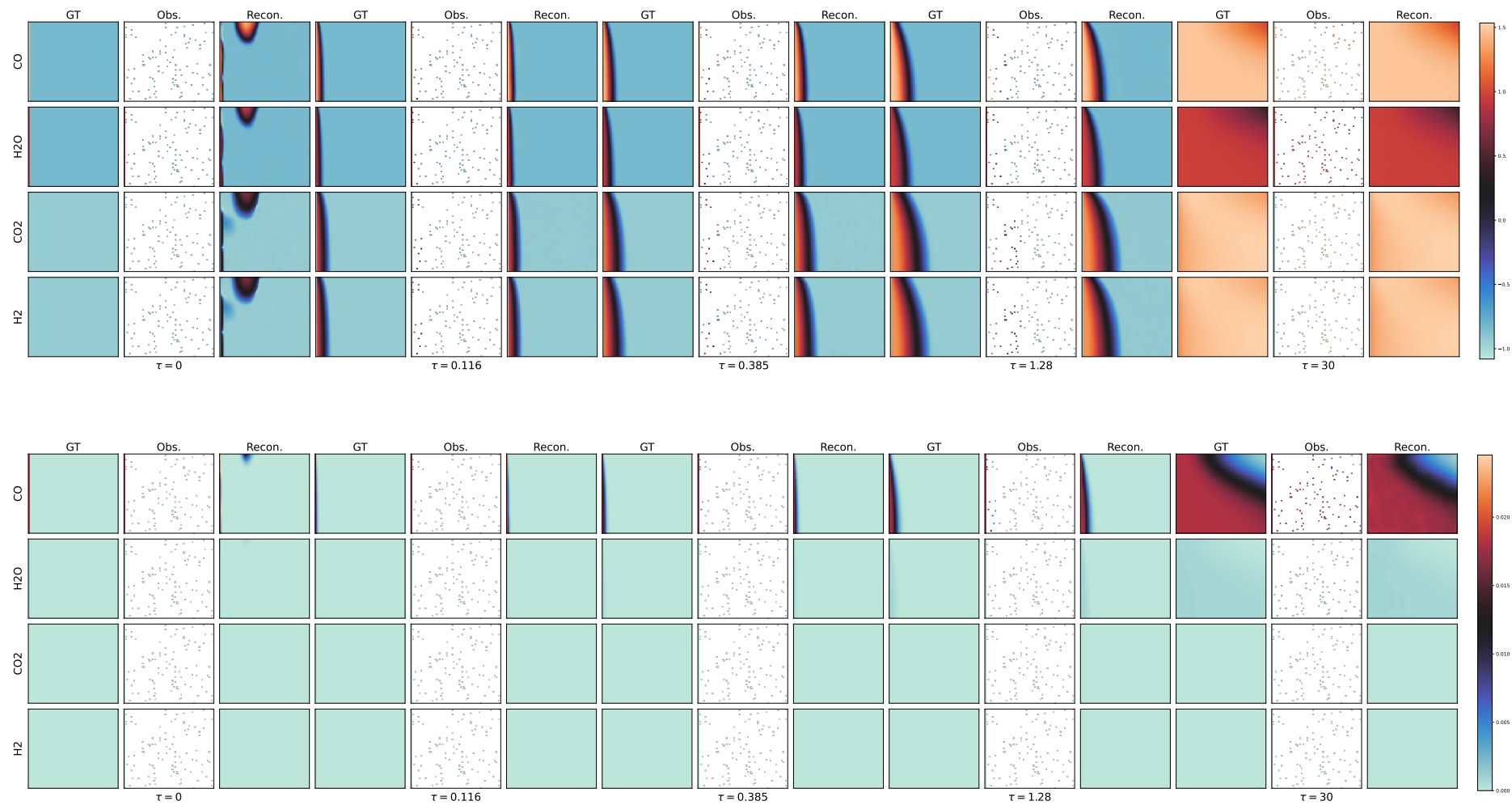


Figure 31: Water Gas Shift reconstruction panel for the DiffPDE method for select PDE times. Top: Normalised. Bottom: Physical.

F Error Maps

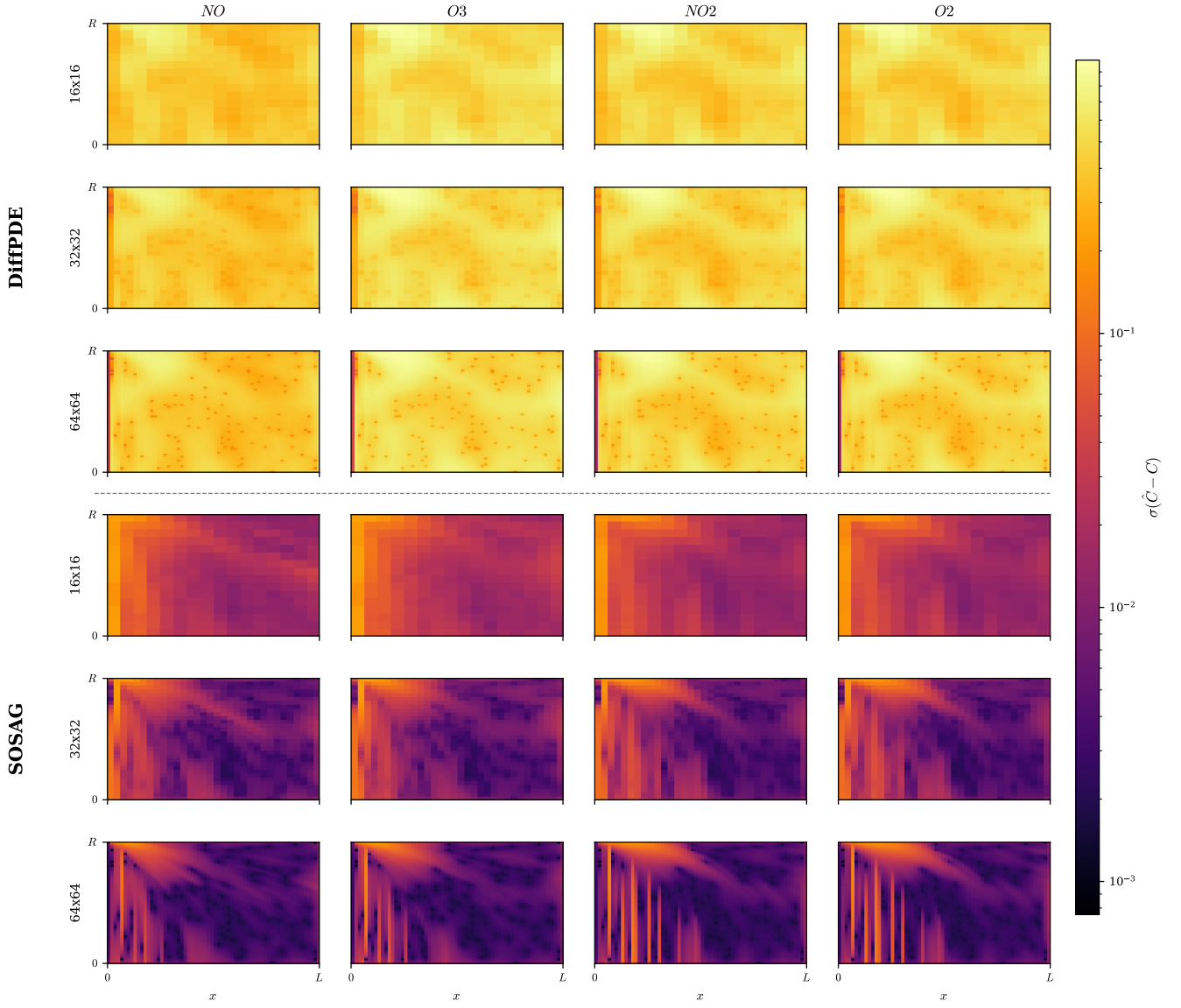


Figure 32: Map of the standard deviation of reconstruction errors for the DiffPDE and SOSaG methods on the NO-O₃ to NO₂ dataset for differing resolutions. The standard deviation is calculated per pixel over all timesteps across the 10 test set trajectories. SOSaG produces a lower error variability compared to DiffPDE.

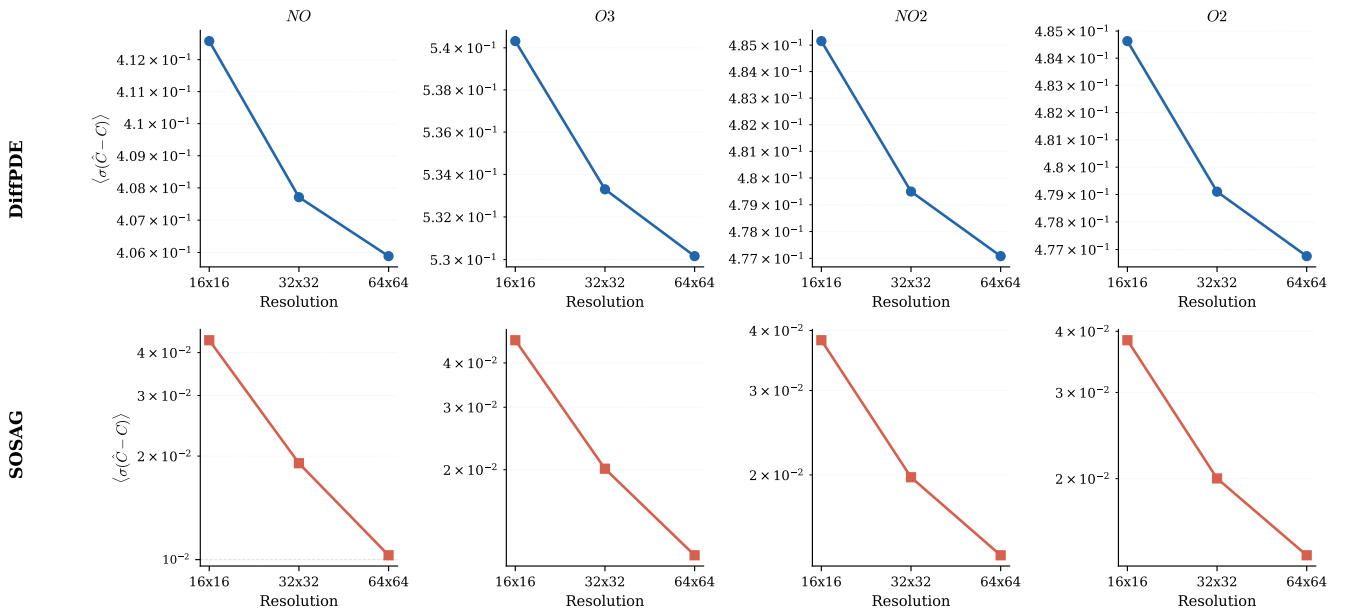


Figure 33: Line plots of the standard deviation of reconstruction errors for the DiffPDE and SOSaG methods on the NO-O₃ to NO₂ dataset for different resolutions. We notice that the error decreases as the resolution increases, an expected result. The standard deviation in the error drops at a faster rate when using SOSaG compared to DiffPDE.