`WaveStitch`: Flexible and Fast Conditional Time Series Generation With Diffusion Models (Appendix)

Algorithm 1 WaveStitch Inference with RePainting

1: **Input:** Windowed metadata ($\mathcal{A}_w$), masks ($\mathcal{M}_w$), and signals ($\mathcal{X}_w$); Timesteps M, window size w; Stride s; mini-batch size b; diffusion parameters $\{\alpha_t, \beta_t, \overline{\alpha}_t\}$; metadata $\mathcal{A} = \{\mathbf{a}^{(i)}\}_{i=1}^{M}$; Denoiser f_θ.

2: **Initialize outputs:**
$$\hat{\mathcal{X}}_w = \{\hat{\mathbf{x}}_{w,T}^{(j)} \sim \mathcal{N}(0, I)\}_{j=1}^{\lfloor (M-w)/s \rfloor}$$

3: **Divide** $\mathcal{X}_w, \mathcal{A}_w, \mathcal{M}_w$ **into** $(M-w)/(b \times s)$ **mini-batches**

4: **for** each mini-batch **do**

5:     **for** $(\mathbf{x}_w^{(j)}, \mathbf{a}_w^{(j)}, \mathbf{m}_w^{(j)})$ in mini-batch in **parallel**: **do**

6:         **for** step $t = T, T-1, \ldots, 1$ **do**

7:             **Conditional Forward noising**:
$$\hat{\mathbf{x}}_{w,t}^{(i)} = (1 - \mathbf{m}_w^{(i)}) \cdot (\sqrt{\overline{\alpha}_t} \cdot \mathbf{x}_w^{(i)} + \sqrt{1 - \overline{\alpha}_t} \cdot \epsilon^{(i)}) + \mathbf{m}_w^{(i)} \cdot \hat{\mathbf{x}}_{w,t}^{(i)}$$

8:             **One-step denoising**:
$$\hat{\mathbf{x}}_{w,t-1}^{(i)} = \frac{1}{\sqrt{\alpha_t}} \left(\hat{\mathbf{x}}_{w,t}^{(i)} - \frac{1 - \alpha_t}{\sqrt{1 - \overline{\alpha}_t}} \cdot f_\theta \left(\mathbf{a}_w^{(i)}, \hat{\mathbf{x}}_{w,t}^{(i)}, t \right) \right)$$

9:             **Re-introduce conditions:**
$$\hat{\mathbf{x}}_{w,t-1}^{(i)} = (1 - \mathbf{m}_w^{(i)}) \cdot \mathbf{x}_w^{(i)} + \mathbf{m}_w^{(i)} \cdot \hat{\mathbf{x}}_{w,t-1}^{(i)}$$

10:         **if** $s < w$ **then**

11:             **for** $i > 1$ in **parallel do**

12:                 **Stitch overlaps:**
$$\hat{\mathbf{x}}_w^{(i)(1:w-s)} = \hat{\mathbf{x}}_w^{(i-1)(1+s:w)}$$

13: **Merge windows:**
$$\hat{\mathcal{X}} = \hat{\mathbf{x}}_w^{(1)} \cup \left(\bigcup_{i \geq 2} \hat{\mathbf{x}}_w^{(i)(w-s+1:w)} \right)$$

14: **return** $\hat{\mathcal{X}}$

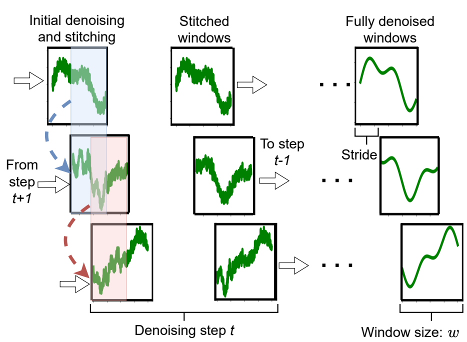

Figure 2: Parallel Denoising with Overlap Stitching

Table 1: Comparison of performance (MSE↓) of TSDiff under varying guidance strengths, for different datasets and tasks (*R*/*I*/*B*). Best score is in bold and second-best is underlined.

	0.0	**0.5**	**1.0**	**2.0**
(***R***)	$1.469_{.002}$	$1.223_{.007}$	$\underline{1.174_{.002}}$	$\mathbf{1.062_{.022}}$
AT (***I***)	$1.964_{.020}$	$\underline{1.480_{.003}}$	$\mathbf{1.395_{.020}}$	$1.932_{.002}$
(***B***)	$\mathbf{1.331_{.012}}$	$1.567_{.032}$	$\underline{1.446_{.007}}$	$1.609_{.004}$
(***R***)	$1.026_{.011}$	$\underline{0.964_{.009}}$	$\mathbf{0.880_{.014}}$	$1.017_{.007}$
MT (***I***)	$\mathbf{0.573_{.036}}$	$\underline{0.768_{.027}}$	$0.846_{.073}$	$0.809_{.041}$
(***B***)	$\underline{0.215_{.008}}$	$\mathbf{0.200_{.010}}$	$0.292_{.012}$	$0.267_{.006}$
(***R***)	$2.333_{.014}$	$\mathbf{1.763_{.017}}$	$2.444_{.010}$	$\underline{2.158_{.009}}$
BQ (***I***)	$\mathbf{1.342_{.011}}$	$\underline{1.422_{.025}}$	$2.780_{.016}$	$1.514_{.010}$
(***B***)	$0.244_{.022}$	$\mathbf{0.162_{.007}}$	$0.167_{.007}$	$\underline{0.164_{.012}}$
(***R***)	$0.785_{.008}$	$0.893_{.021}$	$\mathbf{0.688_{.007}}$	$\underline{0.760_{.011}}$
RS (***I***)	$0.971_{.037}$	$0.827_{.032}$	$\underline{0.723_{.007}}$	$\mathbf{0.637_{.025}}$
(***B***)	$\underline{0.298_{.006}}$	$0.458_{.011}$	$\mathbf{0.276_{.004}}$	$0.394_{.007}$
(***R***)	$\mathbf{1.405_{.011}}$	$1.667_{.005}$	$2.105_{.003}$	$\underline{1.410_{.008}}$
PE (***I***)	$2.171_{.092}$	$\mathbf{1.179_{.061}}$	$\underline{1.798_{.033}}$	$2.538_{.039}$
(***B***)	$\mathbf{0.369_{.001}}$	$0.469_{.002}$	$\underline{0.407_{.001}}$	$0.469_{.001}$
Avg.	$\underline{1.099}$	$\mathbf{1.002}$	1.161	1.116

A. RePaint-based WaveStitch

To condition on both metadata and available signal values during inference, we implement a RePainting-based version of `WaveStitch`, detailed in Algorithm 3. First, we apply a conditional mask to identify the known signal values, which is used to generate noised versions of the signal (step 1, Figure 1). The denoising process is then applied to the entire signal, with conditions re-introduced at each iteration to adjust the unconstrained parts of the signal in alignment with the known values.

We parallelize the generation of time series segments using overlapping windows. Each mini-batch of windows is denoised in parallel, conditioned on the historical signals and metadata. To enforce coherence across windows, the RePaint-based stitching mechanism simply **overwrites** overlapping regions in each window with the corresponding region from the preceding window, gradually aligning non-overlapping parts in subsequent iterations (see Figure 2, line 12). This approach, ensures that coherence emerges progressively through iterative refinement.

B. Impact of Guidance Strength (TSDiff)

For transparency on the tuning sensitivity of baselines, we included results analyzing the impact of guidance strength on `TSDiff` performance in Table 1 below, to contextualize performance variations. As seen in the table, the guidance strength of 0.5 gives the best performance on average, which we have consistently used for the experiments in the main text as well.

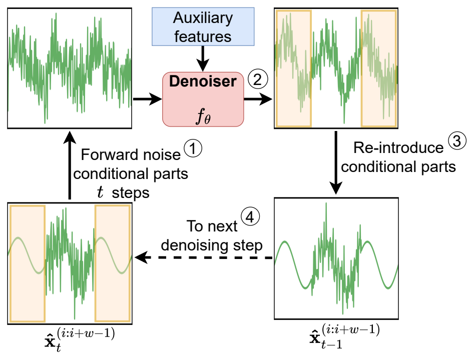

Figure 1: Conditional Denoising. Observed values in orange.