

# Supplementary Material to Nonlocal Neural Networks, Nonlocal Diffusion and Nonlocal Modeling

In the supplementary material, we aim to show some decay properties of the nonlocal equation

$$\begin{cases} z_t(\mathbf{x}, t) - \mathcal{L}z(\mathbf{x}) = 0, \\ z(\mathbf{x}, 0) = u(\mathbf{x}), \end{cases} \quad (1)$$

for  $\mathbf{x} \in \Omega \subset \mathbb{R}^d$  and  $t \geq 0$ . Here the nonlocal operator is defined as

$$\mathcal{L}z(\mathbf{x}) := \int_{\Omega} \rho(\mathbf{x}, \mathbf{y})(z(\mathbf{y}) - z(\mathbf{x}))d\mathbf{y}, \quad (2)$$

where the kernel

$$\rho(\mathbf{x}, \mathbf{y}) := \omega(u(\mathbf{x}), u(\mathbf{y})) \quad (3)$$

is symmetric and positivity-preserved. We should remark first that in all the proofs presented below, we always omit the time variable  $t$  if not referred to. For simplicity, we always assume that  $z(\mathbf{x})$  is a scalar function. We first derive some properties of the nonlocal operator  $\mathcal{L}$ .

**Proposition 1.** *The nonlocal operator  $\mathcal{L}$  defined in Eq. (2) admits the following properties:*

- (i) *If  $z(\mathbf{x})$  is a constant function for all  $\mathbf{x} \in \Omega$ , then  $\mathcal{L}z \equiv 0$ .*
- (ii)  *$\mathcal{L}z$  is mean zero for any function  $z$ , namely,*

$$\int_{\Omega} \mathcal{L}z(\mathbf{x})d\mathbf{x} = 0. \quad (4)$$

- (iii)  *$-\mathcal{L}$  is positive semi-definite, namely,*

$$-\int_{\Omega} \mathcal{L}z(\mathbf{x})z(\mathbf{x})d\mathbf{x} \geq 0, \quad (5)$$

for any function  $z$ .

*Proof.* Property (i) is straightforward. Property (ii) is true by the symmetry of the kernel function, namely,

$$\begin{aligned} \int_{\Omega} \mathcal{L}z(\mathbf{x})d\mathbf{x} &= \int_{\Omega} \int_{\Omega} \rho(\mathbf{x}, \mathbf{y})(z(\mathbf{y}) - z(\mathbf{x}))d\mathbf{y}d\mathbf{x} \\ &= \frac{1}{2} \int_{\Omega} \int_{\Omega} \rho(\mathbf{x}, \mathbf{y})(z(\mathbf{y}) - z(\mathbf{x}))d\mathbf{y}d\mathbf{x} + \frac{1}{2} \int_{\Omega} \int_{\Omega} \rho(\mathbf{y}, \mathbf{x})(z(\mathbf{x}) - z(\mathbf{y}))d\mathbf{y}d\mathbf{x} \\ &= 0. \end{aligned}$$

For Property (iii), using the symmetry of  $\rho$  again we can get

$$\begin{aligned} -\int_{\Omega} \mathcal{L}z(\mathbf{x})z(\mathbf{x})d\mathbf{x} &= -\int_{\Omega} \int_{\Omega} \rho(\mathbf{x}, \mathbf{y})(z(\mathbf{y}) - z(\mathbf{x}))z(\mathbf{x})d\mathbf{y}d\mathbf{x} \\ &= \frac{1}{2} \int_{\Omega} \int_{\Omega} \rho(\mathbf{x}, \mathbf{y})(z(\mathbf{y}) - z(\mathbf{x}))^2d\mathbf{y}d\mathbf{x}. \end{aligned}$$

Since  $\rho(\mathbf{x}, \mathbf{y}) \geq 0$  for all  $\mathbf{x}, \mathbf{y} \in \Omega$ , the above is always non-negative. □

It is easy to see that the discrete nonlocal operator  $\mathcal{L}^h$  defined in Section 4.2 of the main paper has the same properties in the discrete form. Now we can study the properties of the nonlocal integral-differential equation (1). The following gives a formal theorem on this observation.

**Theorem 2.** *The flow of Eq. (1) admits the following properties:*

(i) *The mean value is preserved, namely,*

$$\int_{\Omega} z(t, \mathbf{x}) d\mathbf{x} = \int_{\Omega} u(\mathbf{x}) d\mathbf{x}, \quad \text{for all } t \geq 0. \quad (6)$$

(ii) *As  $t \rightarrow \infty$ , the solution converges to a constant, which is the mean value of the initial condition  $u$ , namely,*

$$\lim_{t \rightarrow \infty} z(t, \mathbf{x}) = \frac{1}{|\Omega|} \int_{\Omega} u(\mathbf{x}) d\mathbf{x}, \quad (7)$$

where  $|\Omega|$  is the measure of the space domain.

*Proof.* For Property (i), by taking the time derivative of the integral of  $z$  over  $\Omega$ , we can get that

$$\frac{d}{dt} \int_{\Omega} z(t, \mathbf{x}) d\mathbf{x} = \int_{\Omega} z_t(t, \mathbf{x}) d\mathbf{x} = \int_{\Omega} \mathcal{L}z(\mathbf{x}) d\mathbf{x}.$$

From Property (ii) of Proposition 1, the above equations are equal to 0, which means that the mean value does not change over time.

For Property (ii), we first prove another property of energy decay, namely

$$\frac{d}{dt} \int_{\Omega} z(t, \mathbf{x})^2 d\mathbf{x} \leq 0.$$

Indeed, by Property (iii) of Proposition 1, we have

$$\begin{aligned} \frac{d}{dt} \int_{\Omega} z(t, \mathbf{x})^2 d\mathbf{x} &= 2 \int_{\Omega} z_t(t, \mathbf{x}) z(t, \mathbf{x}) d\mathbf{x} \\ &= 2 \int_{\Omega} \mathcal{L}z(\mathbf{x}) z(\mathbf{x}) d\mathbf{x} \\ &\leq 0. \end{aligned}$$

Moreover, in the proof of Property (iii) of Proposition 1, we can see that the above estimate is strictly negative unless  $z(\mathbf{x})$  is a constant for any  $\mathbf{x} \in \Omega$ . Then as  $t \rightarrow \infty$ ,  $z(t, \mathbf{x})$  must converge to a constant function, which has to be  $\frac{1}{|\Omega|} \int_{\Omega} u(\mathbf{x}) d\mathbf{x}$  by Property (i) of this theorem.  $\square$

Another observation is that the variance of  $z(t, \mathbf{x})$  is decreasing over iterations. The variance is defined as

$$\text{var}(z) := \frac{1}{|\Omega|} \int_{\Omega} \left( z(\mathbf{x}) - \frac{1}{|\Omega|} \int_{\Omega} z(\mathbf{y}) d\mathbf{y} \right)^2 d\mathbf{x}. \quad (8)$$

Then we can derive the following theorem.

**Theorem 3.** *Let  $z(t, \mathbf{x})$  be the solution to Eq. (1), then*

$$\frac{d}{dt} \text{var}(z) \leq 0. \quad (9)$$

*Proof.* Taking the time derivative of the variance, we can get

$$\begin{aligned} \frac{d}{dt} \text{var}(z) &= \frac{2}{|\Omega|} \int_{\Omega} \left( z(t, \mathbf{x}) - \frac{1}{|\Omega|} \int_{\Omega} u(\mathbf{y}) d\mathbf{y} \right) z_t(t, \mathbf{x}) d\mathbf{x} \\ &= \frac{2}{|\Omega|} \left( \int_{\Omega} z(\mathbf{x}) \mathcal{L}z(\mathbf{x}) d\mathbf{x} - \frac{1}{|\Omega|} \int_{\Omega} u(\mathbf{y}) d\mathbf{y} \int_{\Omega} \mathcal{L}z(\mathbf{x}) d\mathbf{x} \right) \\ &= \frac{2}{|\Omega|} \int_{\Omega} z(\mathbf{x}) \mathcal{L}z(\mathbf{x}) d\mathbf{x} \\ &\leq 0, \end{aligned}$$

where in the derivation we have used the fact that  $\int_{\Omega} \mathcal{L}z(\mathbf{x})d\mathbf{x} = 0$ .  $\square$

We are also interested in the convergence rate of  $z(t, \mathbf{x})$  to its mean value with respect to  $t$ . The following theorem tells us that the solution to Eq. (1) decays to its mean value with an exponential rate.

**Theorem 4.** *Suppose  $u \in L^2(\Omega)$ , i.e., square integrable. Then as  $t \rightarrow \infty$ , the solution  $z$  to Eq. (1) satisfies*

$$\|z(t, \mathbf{x}) - \bar{u}\|_{L^2(\Omega)} \leq Ce^{-\lambda t}, \quad (10)$$

with some positive constants  $C$  and  $\lambda$ , where

$$\bar{u} = \frac{1}{|\Omega|} \int_{\Omega} u(\mathbf{x})d\mathbf{x}. \quad (11)$$

Before we give the proof of Theorem 4, we establish a nonlocal-type Poincaré inequality as follows.

**Lemma 5 (Nonlocal Poincaré Inequality).** *Supposed that  $z(\mathbf{x}) \in L^2(\Omega)$  and  $\int_{\Omega} z(\mathbf{x})d\mathbf{x} = 0$ . Then there exists a constant  $C > 0$  such that*

$$\int_{\Omega} z(\mathbf{x})^2 d\mathbf{x} \leq C \int_{\Omega} \int_{\Omega} \rho(\mathbf{x}, \mathbf{y})(z(\mathbf{y}) - z(\mathbf{x}))^2 d\mathbf{y}d\mathbf{x}.$$

*Proof.* We show that

$$0 < m = \inf_{z(\mathbf{x}) \in L^2(\Omega), \|z\|_{L^2} = 1} \left( \int_{\Omega} \int_{\Omega} \rho(\mathbf{x}, \mathbf{y})(z(\mathbf{y}) - z(\mathbf{x}))^2 d\mathbf{y}d\mathbf{x} \right).$$

Clearly,  $m \geq 0$ . Suppose that  $m = 0$ , then there exists a sequence  $z_n(\mathbf{x}) \in L^2(\Omega)$  and  $\int_{\Omega} z_n(\mathbf{x})d\mathbf{x} = 0$  such that, for all  $n$ ,  $\|z_n\|_{L^2} = 1$  and

$$\lim_{n \rightarrow \infty} \int_{\Omega} \int_{\Omega} \rho(\mathbf{x}, \mathbf{y})(z_n(\mathbf{y}) - z_n(\mathbf{x}))^2 d\mathbf{y}d\mathbf{x} = 0.$$

Then  $\{z_n\}$  is precompact in  $L^2(\Omega)$ , which means that there exists a strong limit  $z_{\infty}$  of  $z_n$  in  $L^2(\Omega)$ . Then, on one hand, we have  $\|z_{\infty}\|_{L^2} = 1$  and  $\int_{\Omega} z_{\infty}(\mathbf{x})d\mathbf{x} = 0$ . On the other hand,

$$\begin{aligned} 0 &= \lim_{n \rightarrow \infty} \int_{\Omega} \int_{\Omega} \rho(\mathbf{x}, \mathbf{y})(z_n(\mathbf{y}) - z_n(\mathbf{x}))^2 d\mathbf{y}d\mathbf{x} \\ &= \lim_{n \rightarrow \infty} -2 \int_{\Omega} \mathcal{L}z_n(\mathbf{x})z_n(\mathbf{x})d\mathbf{x} \\ &= -2 \int_{\Omega} \mathcal{L}z_{\infty}(\mathbf{x})z_{\infty}(\mathbf{x})d\mathbf{x} \\ &= \int_{\Omega} \int_{\Omega} \rho(\mathbf{x}, \mathbf{y})(z_{\infty}(\mathbf{y}) - z_{\infty}(\mathbf{x}))^2 d\mathbf{y}d\mathbf{x}. \end{aligned}$$

The above integral is 0 if and only if  $z_{\infty}$  is a constant function. Then by  $\int_{\Omega} z_{\infty}(\mathbf{x})d\mathbf{x} = 0$ , we must have  $z_{\infty} \equiv 0$ , which contradicts to  $\|z_{\infty}\|_{L^2} = 1$ . This completes the proof.  $\square$

Now we can start to prove Theorem 4.

*Proof of Theorem 4.* Let  $\bar{z}(t, \mathbf{x}) = z(t, \mathbf{x}) - \bar{u}$ . Then we have

$$\int_{\Omega} \bar{z}(t, \mathbf{x})d\mathbf{x} = 0 \quad \text{for all } t \geq 0.$$

Since  $u \in L^2(\Omega)$ , it is also easy to get that  $\bar{z}(\cdot, \mathbf{x}) \in L^2(\Omega)$ . Moreover,  $\bar{z}$  is the solution to

$$\begin{cases} \bar{z}_t(t, \mathbf{x}) - \int_{\Omega} \rho(\mathbf{x}, \mathbf{y})(\bar{z}(t, \mathbf{y}) - \bar{z}(t, \mathbf{x}))d\mathbf{y} = 0, \\ \bar{z}(0, \mathbf{x}) = u(\mathbf{x}) - \bar{u}. \end{cases}$$

Then by the nonlocal Poincaré inequality, there exists a constant  $\lambda > 0$  such that

$$\begin{aligned} \frac{d}{dt} \int_{\Omega} \bar{z}(t, \mathbf{x})^2 d\mathbf{x} &= - \int_{\Omega} \rho(\mathbf{x}, \mathbf{y}) (\bar{z}(\mathbf{y}) - \bar{z}(\mathbf{x}))^2 d\mathbf{y} d\mathbf{x} \\ &\leq -2\lambda \int_{\Omega} \bar{z}(\mathbf{x})^2 d\mathbf{x}. \end{aligned}$$

Therefore, we can derive that

$$\frac{d}{dt} \left( e^{2\lambda t} \int_{\Omega} \bar{z}(t, \mathbf{x})^2 d\mathbf{x} \right) \leq 0.$$

Let  $C = \left( \int_{\Omega} \bar{z}(0, \mathbf{x})^2 d\mathbf{x} \right)^{1/2}$ . Then we have

$$\|z(t, \mathbf{x}) - \bar{u}\|_{L^2(\Omega)} \leq C e^{-\lambda t}.$$

□

To sum up, all the properties derived from the nonlocal equation are also true for local diffusions. Moreover, since the nonlocal equation (1) is a continuum generalization of the proposed nonlocal network which can be regarded as the discretization of Eq. (1), we can view the formulation of the proposed nonlocal blocks as an analogue of local diffusive terms.