

## *Inverse problems in neural field theory*

Article

Published Version

Potthast, R. ORCID: <https://orcid.org/0000-0001-6794-2500>  
and Beim Graben, P. (2009) Inverse problems in neural field  
theory. SIAM Journal on Applied Dynamical Systems, 8 (4).  
pp. 1405-1433. ISSN 1536-0040 doi: 10.1137/080731220  
Available at <https://centaur.reading.ac.uk/29359/>

It is advisable to refer to the publisher's version if you intend to cite from the  
work. See [Guidance on citing](#).

To link to this article DOI: <http://dx.doi.org/10.1137/080731220>

Publisher: Society for Industrial and Applied Mathematics

All outputs in CentAUR are protected by Intellectual Property Rights law,  
including copyright law. Copyright and IPR is retained by the creators or other  
copyright holders. Terms and conditions for use of this material are defined in  
the [End User Agreement](#).

[www.reading.ac.uk/centaur](http://www.reading.ac.uk/centaur)

**CentAUR**

Central Archive at the University of Reading

Reading's research outputs online

## Inverse Problems in Neural Field Theory\*

Roland Potthast<sup>†</sup> and Peter beim Graben<sup>‡</sup>

**Abstract.** We study inverse problems in neural field theory, i.e., the construction of synaptic weight kernels yielding a prescribed neural field dynamics. We address the issues of existence, uniqueness, and stability of solutions to the inverse problem for the Amari neural field equation as a special case, and prove that these problems are generally ill-posed. In order to construct solutions to the inverse problem, we first recast the Amari equation into a linear perceptron equation in an infinite-dimensional Banach or Hilbert space. In a second step, we construct sets of biorthogonal function systems allowing the approximation of synaptic weight kernels by a generalized Hebbian learning rule. Numerically, this construction is implemented by the Moore–Penrose pseudoinverse method. We demonstrate the instability of these solutions and use the Tikhonov regularization method for stabilization and to prevent numerical overfitting. We illustrate the stable construction of kernels by means of three instructive examples.

**Key words.** neural field theory, Amari equation, inverse problem, neural learning, Hebbian learning rule, Tikhonov regularization, kernel construction

**AMS subject classifications.** 92B20, 68T05, 82C32, 91E40, 45K05, 93B30, 93B05, 93B35, 93B40, 65R32, 65R30

**DOI.** 10.1137/080731220

**1. Introduction.** In this paper, we address the *inverse problem* for neural field theories. Neural fields are continuum limits of neural networks, which are generally described by integro-differential equations [1, 3, 9, 10, 12, 15, 16, 21, 22, 32, 33, 40, 42]. Performing this limit for a recurrent neural network of  $n$  leaky integrator units with activations  $u_i(t) \in \mathbb{R}$  [2, 4, 5, 19, 37, 41],

$$(1) \quad \tau \frac{\partial u_i}{\partial t}(t) = -u_i(t) + \sum_{j=1}^n w_{ij} f(u_j(t))$$

with positive time constant  $\tau > 0$ , synaptic weights  $w_{ij}$ , and the nonlinear squashing function  $f : \mathbb{R} \rightarrow [0, 1]$ , which is assumed to be smooth and monotonous, one obtains the  $m$ -dimensional *Amari equation* [1]

$$(2) \quad \tau \frac{\partial u}{\partial t}(x, t) = -u(x, t) + \int_D w(x, y) f(u(y, t)) dy, \quad x \in D, t > 0,$$

with *initial condition*

$$(3) \quad u(x, 0) = u_0(x), \quad x \in D.$$

\*Received by the editors July 25, 2008; accepted for publication (in revised form) by B. Ermentrout June 14, 2009; published electronically October 22, 2009.

<http://www.siam.org/journals/siads/8-4/73122.html>

<sup>†</sup>Department of Mathematics, University of Reading, Whiteknights, P.O. Box 220, Berkshire RG6 6AX, UK ([r.w.e.potthast@reading.ac.uk](mailto:r.w.e.potthast@reading.ac.uk)).

<sup>‡</sup>Department of Clinical Language Sciences, University of Reading, Harry Pitt Building, Whiteknights, P.O. Box 217, Reading RG6 6AH, UK ([p.r.beimgraben@reading.ac.uk](mailto:p.r.beimgraben@reading.ac.uk)).

In (2),  $w(x, y)$  denotes the *synaptic weight kernel* describing the connectivity between points  $y, x \in \mathbb{R}^m$  situated in a domain  $D \subset \mathbb{R}^m$  with  $m \in \mathbb{N}$ .

In applications of neural networks described by (1) in the engineering or cognitive sciences, one prescribes a certain dynamics  $v_i(t)$  of the nodes (or of a subset of nodes defining an output layer) and asks for a suitable weight matrix  $\mathbf{W} = (w_{ij})$  yielding the desired evolution. In other words, one solves the inverse problem estimating  $\mathbf{W}$  given  $v_i(t)$ . This is usually achieved by means of backpropagation algorithms [5, 7, 19, 25, 27, 28, 36, 38].

Related inverse problems in *neural field theories* comprise the estimation of the synaptic weight kernel  $w(x, y)$  from a prescribed dynamically evolving field  $v(x, t)$  within a time interval  $[0, T]$ , where  $T \in \mathbb{R}$  or  $T = \infty$  might be considered. Methods to solve this problem would be mandatory for modeling transient dynamic fields in cognitive neurodynamics [6, 14, 22, 35, 39].

In this paper, we approach the inverse problem for neural fields using techniques from functional analysis and numerical mathematics. Our approach relies upon recasting the Amari equation (2) into a simple linear perceptron equation in an infinite-dimensional Banach or Hilbert space. Such spaces are naturally given by  $BC(D)$  or  $L^2(D)$  for the neural domain  $D$  under consideration, and we need to investigate the infinite-dimensional operator-analogue  $W$  of the matrix  $\mathbf{W}$  above.

To our knowledge the field theoretical approach for neural inverse problems has not been widely studied yet. However, we believe that the reliable solution of neural field inverse problems on different levels of brain modeling can provide important new mathematical input into the extremely important field of cognitive neurodynamics. This work provides a first step in this direction by bringing together results from different communities and suggesting stable methods for solving important basic inverse problems for the Amari neural field equation.

Let us briefly review the well-known finite-dimensional case as an excursus. *Perceptrons* are one-layered neural networks with feed-forward architecture [19, 26, 34]. A linear perceptron of  $n$  input units with activations  $u_k$  ( $1 \leq k \leq n$ ) and  $m$  output units with activations  $v_i$  ( $1 \leq i \leq m$ ) is governed by the equation

$$(4) \quad v_i = \sum_{k=1}^n w_{ik} u_k$$

or, in matrix notation,

$$(5) \quad \mathbf{v} = \mathbf{W}\mathbf{u},$$

with  $\mathbf{u} \in \mathbb{R}^n$ ,  $\mathbf{v} \in \mathbb{R}^m$ , and  $\mathbf{W} \in \mathbb{R}^{m \times n}$ .

The inverse problem of determining the synaptic weights  $w_{ik}$  for such a *linear associator* can be straightforwardly solved if  $\ell$  distinguished orthogonal input vectors  $\mathbf{p}^{(j)}$  ( $1 \leq j \leq \ell$ ) should be associated with  $\ell$  output vectors  $\mathbf{q}^{(j)}$ . In this case, the weight matrix is simply given as

$$(6) \quad \mathbf{W} = \eta \sum_{j=1}^{\ell} \mathbf{q}^{(j)} \mathbf{p}^{(j)*},$$

where  $\eta = 1/\ell$  is the learning rate and “ $*$ ” denotes the matrix transpose. Equation (6) is called the *Hebb rule* [18, 19, 20]. If, on the other hand, the input vectors  $\mathbf{p}^{(j)}$  are not orthogonal

but still linearly independent, the solution reads

$$(7) \quad \mathbf{W} = \eta \sum_{j=1}^{\ell} \mathbf{q}^{(j)} \mathbf{M}^{-1} \mathbf{p}^{(j)*},$$

where  $\mathbf{M}$  is the overlap matrix of the input vectors,

$$(8) \quad (\mathbf{M})_{ik} = \eta \langle \mathbf{p}^{(i)}, \mathbf{p}^{(k)} \rangle.$$

This solution can also be expressed by means of the Moore–Penrose pseudoinverse method [19]. In general, this problem is ill-posed and nonunique. We will provide a detailed discussion of the phenomena for particular examples in section 6.2.

Here, we will study the kernel construction problem in an infinite-dimensional space  $X$ . We study *uniqueness* and *stability* of the problem by investigating different types of arguments, for example via smoothing properties of the operator  $W$  or by means of orthogonal systems in Hilbert spaces. In particular, we derive a solvability condition which corresponds to a *generalized Hebbian learning* rule. We will derive a continuous version of (7), where we work with *biorthogonal bases* in Hilbert spaces. In general, this inversion will be ill-posed and needs to be regularized, for which we will employ the Tikhonov regularization scheme.

The paper is structured as follows. In section 2 we collect basic definitions and formulate the full field neural inverse problem as an integro-differential equation. Then, that equation is reformulated as a kernel construction problem for a linear integral operator. Section 3 serves to derive results on existence, uniqueness, and stability for the kernel construction problem. In particular, here we observe the strong relationship between general results in Hilbert spaces and the convergence of the Hebbian learning rule for training of neural fields.

In section 4 we study the case of convolution kernels. We show that this case can be converted into an integral equation of the first kind with a compact operator. As a result the inverse problem is ill-posed; i.e., in general there is no solution to the problem, and even if we have a solution for given data, the solution does not depend in a stable way on the data. Further, we show that the ill-posedness of the particular problem yields the ill-posedness of the full field neural inverse problem.

Section 5 is dedicated to the explicit construction of approximate solutions by *biorthonormal basis functions*. When sequences of patterns obtained from sampling the prescribed solution with respect to the time variable build a Riesz basis in the underlying Hilbert space, we obtain feasibility of the construction. We use section 6 to describe a numerical realization of the above methods by means of the Moore–Penrose pseudoinverse or, more generally, the Tikhonov regularization method. We demonstrate the instability of the Moore–Penrose solution and how the difficulties can be remedied by regularization. Numerical examples implementing the theoretical results and showing the feasibility of the tools are provided. We conclude with three instructive examples for kernel construction problems for the Amari equation (2).

**2. The direct and the inverse problems.** We study neural fields  $u(x, t)$  depending on the space variable  $x \in D$  with some bounded domain  $D \subset \mathbb{R}^m$  and the time  $t \geq 0$  governed by the

Amari equation (2) with *initial condition* (3), with  $f : \mathbb{R} \rightarrow [0, 1]$  being some given smooth and monotonous function. We define the *synaptic weight operator*

$$(9) \quad (W\varphi)(x) := \int_D w(x, y)\varphi(y) dy, \quad x \in D,$$

which is acting on the space variable of the functions under consideration. Then, the neural field equation (2) is given by

$$(10) \quad \tau \frac{\partial u}{\partial s}(x, t) = -u(x, t) + \left\{ Wf(u(\cdot, t)) \right\}(x), \quad x \in D, t > 0.$$

We employ the functions

$$(11) \quad \psi(x, t) := \tau \frac{\partial u}{\partial t}(x, t) + u(x, t), \quad x \in D, t \geq 0,$$

and

$$(12) \quad \varphi(x, t) := f(u(x, t)), \quad x \in D, t \geq 0.$$

With  $\varphi$  and  $\psi$  given by (11) and (12), the neural field equation can be equivalently written as

$$(13) \quad \psi = W\varphi \quad \text{on } D \times \mathbb{R}^+.$$

This is a *generalized inverse operator problem*, where  $W$  is an integral operator with kernel  $w$ . The discretized version of the inverse problem (13) is the construction of a matrix  $\mathbf{W} \in \mathbb{R}^{n \times n}$  which satisfies

$$(14) \quad \psi^{(j)} = \mathbf{W}\varphi^{(j)}$$

for a family of elements  $\psi^{(j)}, \varphi^{(j)} \in \mathbb{R}^n$  for all  $j \in J$  with some index set  $J$ . This corresponds to the perceptron equation (5).

Here, we focus on the study of the continuous problem (13) and will employ a numerical quadrature scheme in the last section for numerical experiments. For basic results on existence of the solution to the direct problem [31] we use the notation

$$BC^{0,1}(D) \times L^1(D)$$

for the space of the kernel function  $w(x, y)$ . The notation means that  $w(x, \cdot)$  is in  $L^1(D)$  for each  $x \in D$  and that the element  $w(x, \cdot)$  depends Lipschitz continuously on the variable  $x \in D$  and is uniformly bounded on  $D \times D \subset \mathbb{R}^m \times \mathbb{R}^m$ .

Let  $X$  denote either the Banach space of bounded continuous functions  $BC(D \times D)$ , the space of square integrable functions  $L^2(D \times D)$ , or the space  $BC^{0,1}(D) \times L^1(D)$  from above. Then we can formulate the general inverse problem for the neural field equation as follows. Here, we consider either a finite time interval  $[0, T]$  with  $T \geq 0$  or infinite time intervals  $[0, \infty)$ , which we denote by  $T = \infty$ . We will assume that our functions are continuous on the compact interval  $[0, T]$ ; i.e., they are bounded for finite time.

**Definition 2.1** (full field neural inverse problem). *Given a function*

$$v \in BC(D) \times BC^1([0, T])$$

for some  $T \in \mathbb{R}^+ \cup \{\infty\}$  and with  $v(x, 0) = u_0(x)$ , the full field neural inverse problem is to construct a kernel  $w$  in  $X$  such that the solution of the neural field equation (2) is given by  $v$ .

By (13) or the upcoming version (19), we obtain a reformulation of the full field neural inverse problem into an special kind of a *linear integral equation*,

$$(15) \quad \psi(x, t) = \int_D w(x, y) \varphi(y, t) dy, \quad x \in D, t \in [0, T],$$

for  $w \in X$ . Here, we will use the first formulation (13) and note that the second formulation (19) is an equivalent problem in different spaces.

**Lemma 2.2.** *The full field neural inverse problem according to Definition 2.1 is equivalent to a linear kernel construction (15) given functions  $(\psi, \varphi)$  in the set*

$$U = \left\{ \left( \tau \frac{\partial v}{\partial t}(x, t) + v(x, t), f(v(\cdot, \cdot)) \right) : v \in BC(D) \times BC^1([0, T]) \right\}$$

for  $T \in \mathbb{R}^+ \cup \{\infty\}$ .

**Remarks.** 1. Instead of restricting our attention to the set  $U$ , we can study the more general problems where  $\varphi, \psi \in BC(D) \times BC([0, T])$ .

2. Note that time-behavior of the neural fields  $v$  under consideration can be highly *non-linear*. We will discuss this further in our numerical examples in section 6.

The study of the inverse problem given by Definition 2.1 or Lemma 2.2 includes an analysis of *uniqueness*, *existence*, and the *stability* or *ill-posedness* of the kernel construction. For practical applications a thorough understanding, in particular of uniqueness and stability issues, is of great importance.

**3. Existence, uniqueness, and smoothing.** As a first step in studying the solvability of the inverse problem we note some basic regularity properties of solutions to the neural field equation.

**Lemma 3.1.** *If  $f \in BC^n$ , then the solution to the neural field equation (2) is in  $C^n(\mathbb{R}^+)$  with respect to time; i.e.,  $t \mapsto u(\cdot, t)$  is  $n$  times continuously differentiable as a mapping from  $\mathbb{R}^+$  into  $BC(D)$ .*

**Proof.** A standard tool for analysis is the reformulation of the neural field equation in integral form. Integration over time from  $s = 0$  to  $s = t$  yields

$$(16) \quad \tau(u(x, t) - u(x, 0)) = - \int_0^t u(x, s) ds + \left\{ W \left( \int_0^t f(u(\cdot, s)) ds \right) \right\} (x)$$

with  $x \in D, t \geq 0$ . For later use for  $t \geq 0$  we define

$$(17) \quad \Psi(x, t) := \tau u(x, t) + \int_0^t u(x, s) ds, \quad x \in D,$$

and

$$(18) \quad \Phi(x, t) := \int_0^t f(u(\cdot, s)) ds, \quad x \in D.$$

Then, (10) can be rewritten in the form

$$(19) \quad \Psi(\cdot, t) = W\Phi(\cdot, t) \quad \text{on } D$$

for all  $t \geq 0$ . Now, we carry out an inductive proof. Clearly,  $u(x, t)$  is in  $C^1$  in dependence on the time variable as a function in  $BC(D)$ . Now, let  $u(x, t)$  be in  $C^k(\mathbb{R}^+)$  for some  $k \in \{1, 2, \dots, n-1\}$ . We consider the governing equation in the form (16). If the function  $u$  on the right-hand side is in  $C^k$ , then the field  $u(x, t)$  is in  $C^{k+1}(\mathbb{R}^+)$ . Inductively we now obtain that  $u$  is a  $C^n$  function on the real axis, and the proof is complete. ■

**Corollary 3.2 (conditions on existence).** *For the solvability of the full field neural inverse problem with  $f \in BC^n(\mathbb{R}^+)$  we need that the field  $v$  is in  $C^n(\mathbb{R}^+)$  with respect to the time variable. If  $v$  is not  $BC^n(\mathbb{R}^+)$ , then the inverse problem given  $v$  cannot have a solution.*

The following lemma prepares the subsequent arguments.

**Lemma 3.3.** *The norm of the operator  $W : BC(D) \rightarrow BC(D)$  defined in (9) is given by*

$$(20) \quad \|W\|_\infty = \sup_{x \in D} \int_D |w(x, y)| dy,$$

where  $BC(D)$  is equipped with its canonical norm  $\|\varphi\|_\infty := \sup_{x \in D} |\varphi(x)|$ .

*Proof.* We estimate

$$(21) \quad \begin{aligned} \|W\varphi\|_{BC(D)} &= \sup_{x \in D} \left| \int_D w(x, y) \varphi(y) dy \right| \\ &\leq \left( \sup_{x \in D} \int_D |w(x, y)| dy \right) \cdot \sup_{y \in D} |\varphi(y)|, \end{aligned}$$

which proves

$$(22) \quad \|W\|_\infty \leq \sup_{x, y \in D} \int_D |w(x, y)| dy.$$

To show equality we first consider the case where  $w(x, y) \neq 0$  on  $D \times D$ . We choose  $x_0$  such that

$$(23) \quad \int_D |w(x_0, y)| dy = \sup_{x \in D} \int_D |w(x, y)| dy.$$

Then, we define  $\varphi_0(y) := w(x_0, y)/|w(x_0, y)|$ , which is in  $BC(D)$  and has norm  $\|\varphi_0\|_\infty = 1$ . We calculate

$$(24) \quad |(W\varphi_0)(x_0)| = \left| \int_D w(x_0, y) \frac{w(x_0, y)}{|w(x_0, y)|} dy \right| = \sup_{x \in D} \int_D |w(x, y)| dy,$$

which proves equality in (22) under the assumptions that  $w$  has no zeros. In the case where  $w$  has zeros we define  $N := \{y \in D : w(x_0, y) = 0\}$ . To construct an appropriate sequence of continuous functions we define

$$(25) \quad \chi_n(y) := \begin{cases} 0, & d(y, N) \geq 1/n, \\ 1 - nd(y, N), & 0 \leq d(y, N) \leq 1/n, \end{cases}$$

for  $y \in D \setminus N$ , where  $d(y, N) := \inf_{z \in N} |y - z|$ . Since  $N$  is a closed bounded set we obtain

$$(26) \quad \int_{D \setminus N} \chi_n(y) dy \leq \int_{\{y: d(y, N) \leq 1/n\}} dy \rightarrow 0$$

for  $n \rightarrow \infty$ . Then, we define

$$(27) \quad \varphi_n(y) := \begin{cases} (1 - \chi_n(y)) \frac{w(x_0, y)}{|w(x_0, y)|} + \chi_n(y), & d(y, N) > 0, \\ 1, & \text{otherwise.} \end{cases}$$

By construction we have  $\varphi_n \in BC(D)$  and we observe  $\|\varphi_n\| = 1$ . We calculate

$$(28) \quad \begin{aligned} |(W\varphi_n)(x_0)| &= \left| \int_D w(x_0, y) \varphi_n(y) dy \right| \\ &= \left| \int_{D \setminus N} w(x_0, y) \left[ (1 - \chi_n) \frac{w(x_0, y)}{|w(x_0, y)|} + \chi_n(y) \right] dy \right| \\ &= \left| \int_{D \setminus N} \left[ (1 - \chi_n) w(x_0, y) \frac{w(x_0, y)}{|w(x_0, y)|} + \chi_n(y) w(x_0, y) \right] dy \right| \\ &= \left| \int_{D \setminus N} (1 - \chi_n) |w(x_0, y)| dy + \int_{D \setminus N} \chi_n w(x_0, y) dy \right| \\ &\rightarrow \int_D |w(x_0, y)| dy = \sup_{x \in D} \int_D |w(x, y)| dy, \quad n \rightarrow \infty, \end{aligned}$$

where we use (26) and  $w(x_0, y) = 0$  for  $y \in N$ . The limit (28) now proves equality in (22) in the general case. ■

To study the mapping  $w \mapsto W$  of the kernel  $w$  onto the operator  $W$  we equip the space  $BC(D \times D)$  with the norm

$$(29) \quad \|w\|_{BC(D \times D)} := \sup_{x \in D} \int_D |w(x, y)| dy.$$

**Theorem 3.4.** *For a bounded set  $D$  in  $\mathbb{R}^m$  the mapping*

$$(30) \quad K : BC(D \times D) \rightarrow BL(BC(D), BC(D)), \quad w \mapsto W,$$

*with  $W$  defined in (9) is a linear, bounded, and injective mapping but not surjective.*

*Proof.* Clearly, the mapping  $K$  is linear, since

$$(31) \quad \begin{aligned} K(w_1 + w_2)\varphi &= \int_D (w_1(\cdot, y) + w_2(\cdot, y)) \varphi(y) dy \\ &= \int_D w_1(\cdot, y) \varphi(y) dy + \int_D w_2(\cdot, y) \varphi(y) dy \\ &= Kw_1\varphi + Kw_2\varphi. \end{aligned}$$



We estimate

$$(32) \quad \begin{aligned} \|K(w)\varphi\|_{BC(D)} &= \sup_{x \in D} \left| \int_D w(x, y) \varphi(y) dy \right| \\ &\leq \|w\|_{BC(D \times D)} \cdot \|\varphi\|_\infty. \end{aligned}$$

This shows that  $K$  is bounded from  $BC(D \times D)$  into  $BL(BC(D), BC(D))$  with  $\|K\| = 1$ , where we obtain equality by looking at particular kernels  $w$  for which the supremum is obtained.

To show injectivity of  $K$  assume that  $w$  is a kernel such that  $W\varphi = 0$  for all  $\varphi \in BC(D)$ . We study a sequence of bounded continuous functions  $\delta_{z,n}(\cdot)$  approximating a delta function  $\delta_z$ . Then, we have

$$(33) \quad 0 = (W\delta_{z,n})(x) = \int_D w(x, y) \delta_{z,n}(y) dy \rightarrow w(x, z), \quad n \rightarrow \infty,$$

which shows  $w(x, z) = 0$  for all  $x, z \in D$ , and thus the operator  $K$  is injective.

The operator  $K$  is not surjective. This is shown by constructing examples  $V$  of bounded linear operators which cannot be represented by integral operators with continuous kernels. We note, for example, the *point evaluation operator*

$$(34) \quad (V\varphi)(x) := \varphi(x), \quad x \in D,$$

which clearly is bounded and linear but cannot be represented as an integral operator with a continuous kernel. Further examples for bounded linear operators which cannot be written as integral operators with continuous kernels are operators with singular kernels of various types. ■

In a second step we study bounded linear operators on the space  $L^2(D)$ . By the Cauchy–Schwarz inequality we obtain the norm estimate

$$(35) \quad \begin{aligned} \|W\varphi\|_{L^2(D)}^2 &\leq \int_D \left| \int_D w(x, y) \varphi(y) dy \right|^2 dx \\ &\leq \left( \int_{D \times D} |w(x, y)|^2 dy dx \right) \|\varphi\|_{L^2(D)}^2, \end{aligned}$$

which proves boundedness of the mapping

$$(36) \quad K : L^2(D \times D) \rightarrow BL(L^2(D), L^2(D)), \quad w \mapsto W.$$

However, here the estimate (35) is only an upper estimate, but the norm of  $W$  is given by the spectral radius  $\rho(W^*W)$ ; compare [24].

Consider any orthonormal basis  $\{\varphi_n, n \in \mathbb{N}\}$  in  $X$ . Then, every element  $\varphi \in X$  can be written as

$$(37) \quad \varphi = \sum_{n=1}^{\infty} \beta_n \varphi_n, \quad \beta = (\beta_1, \beta_2, \dots) \in \ell^2.$$

The image sequence  $\psi_n := W\varphi_n$  satisfies

$$(38) \quad \left\| \sum_{n=1}^{\infty} \beta_n \psi_n \right\| \leq \left\| W \sum_{n=1}^{\infty} \beta_n \varphi_n \right\| \leq \|W\| \cdot \|\beta\|_{\ell^2}.$$

By  $\beta_n = \langle \varphi_n, \varphi \rangle_{L^2(D)} = \int_D \varphi_n(y) \varphi(y) dy$  we can write  $W$  as

$$(39) \quad \begin{aligned} (W\varphi)(x) &= \sum_{n=1}^{\infty} \psi_n(x) \int_D \varphi_n(y) \varphi(y) dy \\ &= \int_D \left( \sum_{n=1}^{\infty} \psi_n(x) \varphi_n(y) \right) \varphi(y) dy, \quad x \in D. \end{aligned}$$

In general, we cannot expect the terms in round brackets to be a function in the space  $L^2(D \times D)$ . For example, when  $\psi_n = \varphi_n$  for all  $n$  we get a representation of the delta function  $\delta(x - y)$  with respect to the variable  $y$ ,

$$(40) \quad \begin{aligned} \int_D \varphi(z) \left( \sum_{n=1}^{\infty} \varphi_n(x) \varphi_n(z) \right) dz &= \sum_{n=1}^{\infty} \varphi_n(x) \int_D \varphi_n(z) \varphi(z) dz \\ &= \sum_{n=1}^{\infty} \varphi_n(x) \beta_n = \varphi(x), \end{aligned}$$

and therefore

$$(41) \quad \sum_{n=1}^{\infty} \varphi_n(x) \varphi_n(z) = \delta(x - z),$$

to be interpreted in  $L^2(D \times D)$  in the sense of (40), which is not an element of  $L^2(D \times D)$ .

When for  $x$  fixed the sequence  $(\psi_n(x))$  satisfies

$$(42) \quad \sum_{n=1}^{\infty} (1 + n^s)^2 |\psi_n(x)|^2 < \infty,$$

we obtain a kernel which is in the Sobolev space  $H^s(D)$  with respect to the variable  $y$ . Now, we formulate the following result, which can be seen as a special case of the theory of Hilbert–Schmidt integral operators.

**Theorem 3.5.** *A bounded linear operator  $W$  is an integral operator on  $L^2(D)$  if and only if the sum*

$$(43) \quad w(x, y) := \sum_{n=1}^{\infty} \psi_n(x) \varphi_n(y)$$

*is convergent in  $L^2(D \times D)$  for any orthonormal basis  $\{\varphi_n : n \in \mathbb{N}\}$  and  $\psi_n := W\varphi_n$ ,  $n \in \mathbb{N}$ .*

*Proof.* If the sum is convergent in  $L^2(D \times D)$ , then we can rewrite (39) as integral operator

$$\begin{aligned} (W\varphi)(x) &= \int_D \left( \sum_{n=1}^{\infty} \psi_n(x) \varphi_n(y) \right) \varphi(y) dy \\ (44) \qquad &= \int_D w(x, y) \varphi(y) dy, \quad x \in D. \end{aligned}$$

By (35) the operator is a linear bounded operator on  $L^2(D)$ . Now, assume that  $W$  is an integral operator with kernel  $w \in L^2(D \times D)$ . Then, for any orthonormal basis  $\{\varphi_n : n \in \mathbb{N}\}$  and  $\psi_n = W\varphi_n$  we obtain  $\psi_n \in L^2(D)$  and

$$\begin{aligned} \sum_{n=1}^N \psi_n(x) \varphi_n(y) &= \sum_{n=1}^N \left( \int_D w(x, z) \varphi_n(z) dz \right) \varphi_n(y) \\ &= \int_D w(x, z) \left( \sum_{n=1}^N \varphi_n(y) \varphi_n(z) \right) dz \\ (45) \qquad &\rightarrow w(x, y), \quad x, y \in D, \quad N \rightarrow \infty, \end{aligned}$$

which, according to (40), is satisfied in  $L^2(D \times D)$ . ■

Note that the kernel construction (43) corresponds to the Hebb rule (6) for a finite linear perceptron with orthogonal training vectors extended to infinite-dimensional function spaces.

One of the implications of the above theorem is that in general we cannot assume that we can control the neural field over an infinite time.

**Theorem 3.6.** *In general, the full field neural inverse problem with infinite time  $T = \infty$  is not solvable in  $BC(D \times D)$ ,  $BC^{0,\alpha}(D) \times L^1(D)$ , or  $L^2(D \times D)$ .*

*Proof.* We first treat the  $L^2$  case. Consider some orthonormal system  $\{\varphi_n, n \in \mathbb{N}\}$  in  $L^2(D)$  constructed out of Haar-type basis functions; i.e.,  $\varphi_n$  has values either zero or one on  $D$ . For simplicity, here we assume that  $f(0) = 0$ ; the general case can be treated analogously. In this case  $f(\varphi_n) = c \cdot \varphi_n$ , where  $c = f(1)$ . Consider  $t_n := n$  for  $n \in \mathbb{N}$  and let  $\{\chi_n : n \in \mathbb{N}\}$  be some  $C^\infty$ -smooth partition of unity,

$$(46) \qquad 1 = \sum_{n=1}^{\infty} \chi_n(t), \quad t \geq 0,$$

such that

$$(47) \qquad \chi_n(t) = \begin{cases} 1, & t \in B_\epsilon(t_n), \\ 0, & t \notin B_{1-\epsilon}(t_n), \end{cases}$$

with some  $1/4 > \epsilon > 0$ . Now, we define

$$(48) \qquad v(x, t) := \sum_{n=1}^{\infty} \chi_n(t) \varphi_n(x), \quad x \in D, \quad t > 0.$$

We obtain

- (a)  $\varphi(\cdot, t_n) = f(v(\cdot, t_n)) = c \cdot \varphi_n(\cdot)$  for  $n \in \mathbb{N}$ ,
- (b)  $\frac{dv}{dt}(\cdot, t_n) = 0$  for  $n \in \mathbb{N}$ , and thus
- (c)  $\psi(\cdot, t_n) = \varphi_n$  for  $n \in \mathbb{N}$ .

This means that to solve the full field neural inverse problem we need to construct an operator  $W$  mapping  $\varphi_n$  onto  $c^{-1} \cdot \varphi_n = W\varphi_n$  for  $n \in \mathbb{N}$ . However, as shown in (40) and Theorem 3.5, the kernel  $w(x, y)$  defined in (43) is not an element of  $L^2(D \times D)$ . This proves the statement of the theorem for  $L^2(D \times D)$ .

We construct a function in  $BC(D)$  which cannot be a solution to the full field neural inverse problem by

$$(49) \quad v(x, t) := t \cdot v_0(x), \quad t \geq 0.$$

Then we obtain  $\psi(\cdot, t) = v_0 \cdot (\tau + t) \sim t$ , but  $f(v(\cdot, t))$  will remain bounded for large  $t$  by  $|f(s)| \leq 1$  for  $s \in \mathbb{R}$ . ■

**Remarks.** 1. We remark that the function (49) is another counterexample for the case of square integrable functions above. However, examples of a different nature provide further insight into the overall nonexistence statement. Also, different counterexamples provide guidance if one seeks sufficient conditions to obtain solvability of the problem.

2. If we modify the points  $t_n$  and the intervals  $\epsilon$  to depend on  $n$ , we can make the same construction for a finite half-open interval  $[0, T)$ . Thus, in general full field neural inverse problems on half-open intervals are not solvable either.

**Compactness.** We conclude this section with basic compactness statements which imply instability of the kernel construction problem if we seek kernels in spaces of differentiable functions or in Sobolev spaces. We note that, by the compactness of the embedding

$$(50) \quad BC^n(D \times D) \rightarrow BC(D \times D),$$

the mapping  $K$  considered in the spaces

$$(51) \quad K : BC^n(D \times D) \rightarrow BL(BC(D), BC(D))$$

for  $n \geq 1$  is compact. Also, using  $H^s$ -smooth kernels in  $L^2$ , we obtain compactness of the mapping

$$(52) \quad K : H^s(D \times D) \rightarrow BL(L^2(D), L^2(D)).$$

We summarize the consequences of this compactness in the following theorem.

**Theorem 3.7.** *In the settings (51) and (52) the mapping  $K : w \mapsto W$  is compact. The mapping cannot have a bounded inverse, and thus the kernel  $w$  depends unstably on the right-hand side.*

**4. Special cases and consequences.** The goal of this section is to study the inversion of the neural field equation for special cases. In particular, we will investigate fields arising from the translation of an initial pulse. As a second result we will conclude the ill-posedness of the general problem from the ill-posedness in special cases.

*Translation of fields in time.* Consider an initial field  $u_0(x)$ , and define the traveling wave field

$$(53) \quad u(x, t) := u_0(x - vt), \quad x \in D, \quad t \geq 0,$$

with  $v \in \mathbb{R}^m$  [12, 15]. Then,  $u(x, t)$  describes a traveling wave in the direction  $v$ . We get

$$(54) \quad \varphi(x, t) = f(u(x, t)) = f(u_0(x - vt)) = \varphi_0(x - vt)$$

and

$$(55) \quad \begin{aligned} \psi(x, t) &= \tau \frac{\partial u_0(x - vt)}{\partial t} + u_0(x - vt) = \psi_0(x - vt) \\ &= -\tau v \cdot \nabla u_0(x - vt) + u_0(x - vt) \end{aligned}$$

for  $x \in D, t \geq 0$ . We remark that if  $\psi_0 = W\varphi_0$  and if the kernel  $w$  is a convolution kernel, i.e.,  $w(x, y) = \tilde{w}(x - y)$ , then we obtain

$$(56) \quad \begin{aligned} \psi(x, t) &= \psi_0(x - vt) \\ &= \int_D w(x - vt, y) \varphi_0(y) dy \\ &= \int_D \tilde{w}(x - vt - y) \varphi_0(y) dy \\ &= \int_D \tilde{w}(x - (y + vt)) \varphi_0(y) dy \\ &= \int_D \tilde{w}(x - \tilde{y}) \varphi_0(\tilde{y} - vt) d\tilde{y} \\ &= (W\varphi)(x, t), \end{aligned}$$

where we use  $\tilde{y} = y + vt$ . This means we need to study the equation

$$(57) \quad \begin{aligned} \psi_0(x) &= \int_D \tilde{w}(x - y) \varphi_0(y) dy \\ &= \int_D \varphi_0(x - \tilde{y}) \tilde{w}(\tilde{y}) d\tilde{y}, \quad x \in D, \end{aligned}$$

where we employed the substitution  $x - y = \tilde{y}$ . We summarize these transformations in the following theorem.

**Theorem 4.1.** *Given a field  $u$  as a traveling wave (53), there is a solution of the full field inverse neural problem with translation invariant kernel  $w(x, y) = \tilde{w}(x - y)$  if and only if the function  $\tilde{w}$  satisfies the integral equation (57).*

The integral equation (57) is an integral equation of the first kind with a continuous or  $L^2$  kernel. In both settings the operator is compact in  $BC(D)$  or  $L^2(D)$ , since it can be approximated in norm by a finite-dimensional operator via polynomial approximations of the kernel function  $\tilde{w}$ ; compare [23]. We now summarize this result and carry it over to the inverse neural field equation.

**Theorem 4.2.** *In general, (57) does not have a solution. The solutions depend unstably on the right-hand side in both  $BC(D)$  and  $L^2(D)$ . The full field inverse neural problem with traveling wave-type functions is ill-posed in the sense of Hadamard (compare [11]).*

*Proof.* We define the operator

$$(58) \quad (Ag)(x) := \int_D \varphi_0(x-y)g(y) dy, \quad x \in D.$$

Note that the kernel density needs to be defined on a larger area,

$$(59) \quad \tilde{D} := \{x-y : x, y \in D\}.$$

If  $\varphi_0 \in BC(\tilde{D})$ , then it is an integral operator with continuous kernel which is a compact operator on  $BC(D)$ . If  $\varphi_0 \in L^2(\tilde{D})$ , we first remark that in this case  $\varphi_0(x-\cdot) \in L^2(D)$  for every fixed  $x \in D$ , and the norm of this function is smaller than  $\|\varphi_0\|_{L^2(\tilde{D})}$ . Then we estimate

$$(60) \quad \begin{aligned} \|Ag\|_{L^2(D)}^2 &\leq \int_D \left| \int_D \varphi_0(x-y)g(y) dy \right|^2 dx \\ &= |D| \cdot \|\varphi_0\|_{L^2(\tilde{D})}^2 \|g\|_{L^2(D)}^2 \end{aligned}$$

such that  $A$  is bounded in  $L^2(D)$ . A polynomial approximation of the function  $\varphi_0$  on  $\tilde{D}$  in the norm  $L^2(\tilde{D})$  will yield a finite-dimensional operator approximation of the operator  $A$  by some operator sequence  $A_n$ . Thus,  $A$  is compact in  $L^2(D)$ .

Now, by standard arguments from functional analysis [23] we conclude that  $A$  cannot be surjective and that the inverse of  $A$  cannot be bounded; i.e., the solution of (57) depends unstably on the right-hand side. Thus, the problem is ill-posed in the sense of Hadamard [11]. This completes the proof. ■

We first formulate a helpful little lemma.

**Lemma 4.3.** *Let  $\tilde{X} \subset X$  be Banach spaces and  $A : X \rightarrow Y$  be a linear operator. We denote  $\tilde{A} := A|_{\tilde{X}}$  the restriction of  $A$  to  $\tilde{X}$  and equip  $\tilde{X}$  with the norm induced by the norm on  $X$ . If  $\tilde{A} : \tilde{X} \rightarrow Y$  does not have a bounded inverse on  $\tilde{A}(\tilde{X})$ , then the operator  $A$  cannot have a bounded inverse.*

*Proof.* If  $A$  has a bounded inverse on  $Y$ , then

$$(61) \quad A^{-1}|_{\tilde{A}(\tilde{X})}$$

is a bounded inverse of  $\tilde{A}$ . Thus  $A$  cannot have a bounded inverse. ■

The above lemma now has consequences for the full field neural inverse problem. We identify  $X$  with the space of kernels  $w \in L^2(D \times D)$  and  $\tilde{X}$  with the space of convolution kernels  $w(x, y) = \tilde{w}(x-y)$ , which is a linear subspace of  $X$ .

**Lemma 4.4.** *The embedding*

$$(62) \quad \mathcal{I} : L^2(\tilde{D}) \rightarrow L^2(D, D), \quad \tilde{w}(\cdot) \mapsto w(x, y) := \tilde{w}(x-y)$$

*is an injective and boundedly invertible linear mapping.*

*Proof.* The mapping (62) is clearly injective, since  $w(x, y) \equiv 0$  is obtained only with  $\tilde{w} \equiv 0$ . The boundedness of the mapping is obtained from

$$\begin{aligned} \|w\|_{L^2(D \times D)}^2 &= \int_D \int_D |w(x, y)|^2 dy dx \\ &= \int_D \left( \int_D |\tilde{w}(x - y)|^2 dy \right) dx \\ (63) \quad &\leq |D| \cdot \|\tilde{w}\|_{L^2(\tilde{D})}^2. \end{aligned}$$

Next, we show that the image space  $\mathcal{I}(L^2(\tilde{D}))$  is a closed subspace of  $L^2(D \times D)$ . To this end let  $\tilde{w}_n$  be a sequence of convolution kernels such that

$$(64) \quad w_n(x, y) = \tilde{w}_n(x - y) \rightarrow w_0(x, y), \quad n \rightarrow \infty,$$

in  $L^2(D, D)$ . Then by

$$(65) \quad w_0(y + z, y) = \lim_{n \rightarrow \infty} w_n(y + z, y) = \lim_{n \rightarrow \infty} \tilde{w}_n(z),$$

$w_0(x, y)$  will depend only on the difference  $x - y$ ; i.e.,  $\mathcal{I}(L^2(\tilde{D}))$  is closed. Now, the boundedness of the inverse is a consequence of the closed mapping theorem. ■

As a consequence of the above lemma and the isometric embedding of  $BC(\tilde{D})$  into  $BC(D \times D)$  by  $\tilde{w} \mapsto w(x, y) := \tilde{w}(x - y)$ , we obtain the instability of the kernel construction problem and thus of the full field neural inverse problem formulated in Definition 2.1.

**Corollary 4.5.** *The full field neural inverse problem is unstable and thus ill-posed in  $BC(D \times D)$  as well as in  $L^2(D \times D)$ .*

**5. Construction of solutions.** Solutions of operator equations in Banach spaces have been studied extensively in the literature. It is well known that for ill-posed equations standard inversion theory (see, for example, [8] for a summary of literature about the Moore–Penrose pseudoinverse) cannot be directly applied, but one needs to use appropriate regularization; compare, for example, [11, 13, 29, 30].

In this section we collect basic results about biorthogonal sets in Hilbert spaces, and we study the construction of a biorthogonal set in particular for a Riesz basis of elements. Biorthogonal sets are usually used for operator construction and are important ingredients for understanding the unstable behavior of the solutions of our dynamic kernel construction problem formulated in Definition 2.1.

In a Hilbert space  $X$  with scalar product  $\langle \cdot, \cdot \rangle$  two linearly independent sets of functions  $Q = \{\varphi_1, \varphi_2, \dots\}$  and  $R = \{\rho_1, \rho_2, \dots\}$  are called *biorthogonal* if

$$(66) \quad \langle \rho_i, \varphi_k \rangle = 0 \quad \text{for all } k \neq i, \quad \langle \rho_i, \varphi_i \rangle = c_i, \quad i \in \mathbb{N},$$

where  $c_i > 0$  for  $i \in \mathbb{N}$ . The construction of a biorthonormal set  $R$  to  $Q$  is usually carried out as follows. We define

$$(67) \quad V_k := \text{span}\{\varphi_1, \dots, \varphi_{k-1}, \varphi_{k+1}, \dots\}, \quad k \in \mathbb{N},$$

denote its orthogonal space by  $V_k^\perp$ , and remark that  $X = \overline{V_k} \oplus V_k^\perp$ . We conclude that  $\varphi_k \notin \overline{V_k}$ , since it is linearly independent of the other elements of  $V_k$ ; thus, its orthogonal projection  $\tilde{\rho}_k$  onto  $V_k^\perp$  cannot be zero. The biorthogonal elements are now given by

$$(68) \quad \rho_k := \frac{\tilde{\rho}_k}{\|\tilde{\rho}_k\|^2}.$$

Here the *ill-posedness* of the inverse problem is strongly related to the division by  $\|\tilde{\rho}_k\|$  in (68). We verify (68) by calculating  $\langle \rho_k, \varphi_i \rangle = \delta_{ki}$ ,  $k, i \in \mathbb{N}$ , where the result is clear by definition for  $i \neq k$  since  $\rho_k \in V_k^\perp$ , and for  $i = k$  we remark that  $\varphi_k = \tilde{\rho}_k + \tilde{\varphi}_k$ , where  $\tilde{\varphi}_k \in V_k$ , which yields

$$(69) \quad \langle \tilde{\rho}_k, \varphi_k \rangle = \langle \tilde{\rho}_k, \tilde{\rho}_k + \tilde{\varphi}_k \rangle = \langle \tilde{\rho}_k, \tilde{\rho}_k \rangle = \|\tilde{\rho}_k\|^2.$$

We have constructed the set  $R := \{\rho_1, \rho_2, \dots\}$ , which satisfies (66) with constants  $c_i = 1$ ,  $i \in \mathbb{N}$ . To show that the elements of  $R$  are linearly independent, assume that  $\sum_{j=1}^n \beta_j \rho_{k_j} = 0$  with constants  $\beta_j \in \mathbb{C}$  and  $k_j \in \mathbb{N}$  for  $j = 1, \dots, n$ . Then, by multiplication with  $\varphi_{k_i}$  for  $i = 1, \dots, n$  we derive

$$(70) \quad 0 = \left\langle \sum_{j=1}^n \beta_j \rho_{k_j}, \varphi_{k_i} \right\rangle = \sum_{j=1}^n \beta_j \langle \rho_{k_j}, \varphi_{k_i} \rangle = \beta_i, \quad i = 1, \dots, n,$$

and thus  $R$  is linearly independent. Recall that  $Q$  is called a Riesz basis in a Hilbert space  $H$  if there are constants  $c_1, c_2 > 0$  such that

$$(71) \quad c_1 \sum_{j=1}^{\infty} |\alpha_j|^2 \leq \left\| \sum_{j=1}^{\infty} \alpha_j \varphi_j \right\|^2 \leq c_2 \sum_{j=1}^{\infty} |\alpha_j|^2$$

for all  $\alpha = (\alpha_j)_{j \in \mathbb{N}} \in \ell^2$ . In this case the mapping

$$(72) \quad A : \ell^2 \rightarrow X, \quad \alpha \mapsto \sum_{j=1}^{\infty} \alpha_j \varphi_j$$

is a bounded and boundedly invertible mapping from  $\ell^2$  onto  $A(\ell^2) \subset X$ . A dual operator  $A' : X \rightarrow \ell^2$  with respect to the scalar products  $\langle \cdot, \cdot \rangle_{\ell^2}$  and  $\langle \cdot, \cdot \rangle_X$  is given by

$$(73) \quad A' : X \rightarrow \ell^2, \quad \psi \mapsto (\langle \varphi_j, \psi \rangle_X)_{j \in \mathbb{N}},$$

as can be obtained from

$$(74) \quad \langle A\alpha, \psi \rangle_X = \sum_{j=1}^{\infty} \alpha_j \langle \varphi_j, \psi \rangle_X = \langle \alpha, A'\psi \rangle_{\ell^2}.$$

We remark that  $(\langle \varphi_j, \psi \rangle_X)_{j \in \mathbb{N}} \in \ell^2$  since we have

$$(75) \quad \left| \sum_{j=1}^{\infty} \beta_j \langle \varphi_j, \psi \rangle_X \right| = \left| \left\langle \sum_{j=1}^{\infty} \beta_j \varphi_j, \psi \right\rangle_X \right| \leq \sqrt{c_2} \|\beta\|_{\ell^2} \|\psi\|_X$$



for all  $\beta \in \ell^2$  via (71). We estimate

$$(76) \quad \langle \alpha, A' A \alpha \rangle_{\ell^2} = \langle A \alpha, A \alpha \rangle_X \geq c_1 \|\alpha\|_{\ell^2}^2,$$

and thus according to the Lax–Milgram theorem the operator  $A' A$  is boundedly invertible in  $\ell^2$  with a bound given by  $1/c_1$ . Note that if  $c_1$  is small, the inverse can have a large norm and the equation is highly ill-conditioned. We calculate

$$(77) \quad A' A \alpha = \left( \langle \varphi_k, A \alpha \rangle_X \right)_{k \in \mathbb{N}} = \left( \sum_{j=1}^{\infty} \langle \varphi_k, \varphi_j \rangle_X \alpha_j \right)_{k \in \mathbb{N}},$$

and thus the operation of  $A' A$  on  $\ell^2$  can be expressed as a matrix multiplication with  $M$  where we employ the infinite *overlap matrix* (compare (8))

$$(78) \quad M := (\langle \varphi_k, \varphi_j \rangle)_{k,j \in \mathbb{N}},$$

which according to (76) is boundedly invertible. Then we define

$$(79) \quad \rho_k := \sum_{j=1}^{\infty} (M^{-1})_{k,j} \varphi_j, \quad k \in \mathbb{N},$$

and calculate

$$(80) \quad \langle \rho_k, \varphi_i \rangle_X = \left\langle \sum_{j=1}^{\infty} (M^{-1})_{k,j} \varphi_j, \varphi_i \right\rangle = \sum_{j=1}^{\infty} (M^{-1})_{k,j} \langle \varphi_j, \varphi_i \rangle = (M^{-1} M)_{ki} = \delta_{ki}$$

for  $k, i \in \mathbb{N}$ . Equation (79) provides a constructive method for calculating the biorthonormal basis functions.

Given two linearly independent sets of elements  $Q = \{\varphi_1, \varphi_2, \dots\} \subset X$  and  $S = \{\psi_1, \psi_2, \dots\} \subset X$ , we can now construct linear operators

$$(81) \quad V_n \varphi := \sum_{j=1}^n \psi_j \langle \rho_j, \varphi \rangle, \quad V \varphi := \sum_{j=1}^{\infty} \psi_j \langle \rho_j, \varphi \rangle,$$

where  $R = \{\rho_1, \rho_2, \dots\}$  is a biorthogonal basis for  $Q$ .

**Lemma 5.1.** *The operator  $V_n$  is linear and bounded on  $H$ . If both  $Q$  and  $S$  are Riesz bases of  $H$ , then  $V$  is linear and bounded on  $H$  as well. Further, we have*

$$(82) \quad \begin{aligned} V_n \varphi_i &= \begin{cases} \psi_i, & i = 1, \dots, n, \\ 0, & i > n. \end{cases} \\ V \varphi_i &= \psi_i, \quad i \in \mathbb{N}. \end{aligned}$$

**Proof.** We first note that  $(\langle \rho_j, \varphi \rangle)_{j \in \mathbb{N}}$  is in  $\ell^2$  with an argument analogous to (75). Then, the boundedness of  $V$  is a consequence of (71) for  $(\psi_j)_{j \in \mathbb{N}}$ .

**6. Numerical study of kernel constructions.** This final section serves to study explicit constructions of integration kernels. We will first briefly describe our numerical realization of the solutions. Then, we study three different basic tasks: (a) kernel construction to generate one-dimensional state dynamics in section 6.2, (b) kernel construction for generating two-dimensional pulses in section 6.3, and (c) kernel construction for constructing logical gates in section 6.4. The function  $f$  will be chosen as the logistic function

$$(83) \quad f(s) := \frac{1}{1 + e^{\beta(s-\eta)}}, \quad s \in \mathbb{R},$$

with parameters  $\beta = 10$ ,  $\eta = 0.5$  in all our examples.

**6.1. Explicit Euler scheme with numerical quadrature.** For the solution of the integro-differential equation (2) we employ an *explicit Euler method* with *numerical quadrature* given by the rectangular or trapezoidal rule.

In the two-dimensional case we choose  $D$  to be a rectangular region  $D = [a_1, b_1] \times [a_2, b_2] \subset \mathbb{R}^2$  and define a grid by

$$(84) \quad h_1 := \frac{b_1 - a_1}{n_1 - 1}, \quad h_2 := \frac{b_2 - a_2}{n_2 - 1}$$

and

$$(85) \quad y_{j,k} := (a_1 + j \cdot h_1, a_2 + k \cdot h_2),$$

for  $j = 0, \dots, n_1 - 1$ ,  $k = 0, \dots, n_2 - 1$ . The points  $y_{j,k}$  are reordered into a vector  $\mathbf{y} \in \mathbb{R}^N$  with  $N = n_1 \cdot n_2$  by

$$(86) \quad \mathbf{y}_\xi := y_{j,k}, \quad \xi = n_2 \cdot j + k, \quad j = 0, \dots, n_1 - 1, \quad k = 0, \dots, n_2 - 1.$$

Consider the space  $X := L^2(D) \cap BC(D)$  of continuous functions on  $D$  equipped with the maximum or the  $L^2$ -norm. A continuous function  $u \in X$  on  $D$  is discretized and represented by a vector  $\mathbf{u} \in \mathbb{R}^N$  defined by

$$(87) \quad \mathbf{u}_\xi := u(\mathbf{y}_\xi), \quad \xi = 1, \dots, N,$$

which defines a mapping  $Q_N : X \rightarrow \mathbb{R}^N$ . Let  $\chi_\xi$  for  $\xi = 0, \dots, N_1$  be standard piecewise linear hat functions on  $D$  such that

$$(88) \quad \chi_\xi(y) = \begin{cases} 1, & y = \mathbf{y}_\xi, \\ 0, & y = \mathbf{y}_\eta \text{ for } \eta \neq \xi. \end{cases}$$

Then the vectors  $\mathbf{u}$  can be mapped back into  $X$  by  $u := P_N \mathbf{u}$  with

$$(89) \quad (P_N \mathbf{u})(y) := \sum_{\xi=1}^{N-1} \mathbf{u}_\xi \chi_\xi(y), \quad y \in D.$$

With this mapping for continuous functions  $u$  we have the convergence  $P_N Q_N u \rightarrow u$  on  $D$  for  $N \rightarrow \infty$  pointwise in  $X$ , i.e., for every fixed  $u$ , both in the maximum and the  $L^2$ -norms.

For the operator  $W$  we employ a *collocation scheme*; i.e., the application of  $W$  to some vector  $\varphi$  is realized by a matrix multiplication  $\mathbf{W}\varphi$ , where the matrix  $\mathbf{W}$  incorporates an appropriate numerical quadrature [23] and evaluates  $W\varphi$  on the grid  $x_{j,k} := y_{j,k}$ ,  $j = 0, \dots, n_1$ ,  $k = 0, \dots, n_2$ , given by (85). In general this leads to

$$(90) \quad \mathbf{W}_{\xi,\eta} = w(\mathbf{x}_\xi, \mathbf{y}_\eta) \alpha_\eta, \quad \xi, \eta = 0, \dots, N-1,$$

with quadrature weights  $\alpha_\eta$ ,  $\eta = 0, \dots, N-1$ , where in the simple case of the rectangular rule we have  $\alpha_\eta = h_1 \cdot h_2$ .

The numerical discretization of (11) and (12) for fixed  $t \in \mathbb{R}$  according to (87) leads to vectors  $\varphi(t)$  and  $\psi(t)$  in  $\mathbb{R}^N$  representing  $\varphi(\cdot, t)$  and  $\psi(\cdot, t)$ . For  $\ell$  time discretization points

$$(91) \quad t_1 < t_2 < \dots < t_\ell$$

we abbreviate

$$(92) \quad \varphi^{(s)} := \varphi(t_s), \quad \psi^{(s)} := \psi(t_s), \quad s = 1, \dots, \ell.$$

As the next step we define the  $\mathbb{R}^{N \times \ell}$ -matrices

$$(93) \quad \mathbf{A} := (\varphi^{(1)}, \dots, \varphi^{(\ell)}) \quad \text{and} \quad \mathbf{B} := (\psi^{(1)}, \dots, \psi^{(\ell)}).$$

The discrete version of (14) or (15) is now given by the equation

$$(94) \quad \mathbf{B} = \mathbf{W}\mathbf{A}$$

for  $\mathbf{W}$ . We can try to calculate a solution by

$$(95) \quad \mathbf{W} = \mathbf{B}\mathbf{A}^\dagger,$$

with the Moore–Penrose pseudoinverse  $\mathbf{A}^\dagger := (\mathbf{A}^* \mathbf{A})^{-1} \mathbf{A}^*$  as a consequence of  $\mathbf{A}^\dagger \mathbf{A} = (\mathbf{A}^* \mathbf{A})^{-1} \mathbf{A}^* \mathbf{A} = I_{\mathbb{R}^{\ell \times \ell}}$ . However, in general  $\mathbf{A}^\dagger$  is ill-conditioned and the condition number quickly increases with  $\ell$ , which reflects the fact that the continuous operator does not have a bounded inverse when  $c_1$  defined in (71) tends to zero. For regularization of the numerical calculation of the operators  $V_n$  defined in (81) we use the Tikhonov inverse

$$(96) \quad \mathbf{R}_\alpha := (\alpha I + \mathbf{A}^* \mathbf{A})^{-1} \mathbf{A}^*$$

with  $\alpha > 0$ . For an existence proof for the inverse we refer to [23] or [13]. The regularized kernel is now calculated by

$$(97) \quad \mathbf{W}_\alpha = \mathbf{B}(\alpha I + \mathbf{A}^* \mathbf{A})^{-1} \mathbf{A}^*.$$

Our numerical examples are carried out by the following steps:

1. Given a field  $v$  according to Definition 2.1 or Lemma 2.2, we first calculate  $\varphi$  and  $\psi$  according to (11) and (12) at the points  $y_{j,k}$  defined in (85) at time steps as defined in (91).

2. This leads to vectors  $\varphi^{(s)}, \psi^{(s)} \in \mathbb{R}^N$  for  $s = 1, \dots, \ell$ . Then, we define the matrices  $\mathbf{A}$  and  $\mathbf{B}$  according to (93).
3. We calculate  $\mathbf{W}_\alpha$  according to (97). Here, we choose the regularization parameter  $\alpha > 0$  by visual inspection. Automatic methods for the determination of  $\alpha$  can be found, for example, in [13].
4. Finally, we test the neural field by solving the initial problem (2) with kernel given by inverse solution  $\mathbf{W}_\alpha$ . This leads to a solution  $v_\alpha(x, t)$  for a set of discrete points  $x \in D$  and  $t \in [0, T]$ .

It is important to note that in step 4 we need to use a different discretization for the time variable than in the steps 1 to 3. Using the same time and space discretization for the inversion and the tests is called *inverse crime*; compare [11]. If we employ the same discretization for simulation and inversion, we do not study the full inverse dynamical problem but a much more stable inversion of a finite-dimensional system. We will explicitly demonstrate this phenomenon in section 6.2. Here, we need to use a regularization parameter which leads to the desired system behavior for a wide range of time discretizations, not only for the particular time discretization used in the inversion.

*Multiple pulses and more complex processes.* Steps 1–3 can be carried out with the data for several pulses at the same time. In this case let  $\varphi^{(p,1)}, \dots, \varphi^{(p,\ell)}$  be the discretized vectors for the pulse with index  $p = 1, \dots, P$ . We change the definition (93) into

$$(98) \quad \mathbf{A} := (\varphi^{(1,1)}, \dots, \varphi^{(1,\ell)}, \varphi^{(2,1)}, \dots, \varphi^{(2,\ell)}, \varphi^{(3,1)}, \dots, \varphi^{(P,\ell)})$$

and analogously for  $\mathbf{B}$ . Then we continue as above. Numerical tests for this are carried out in section 6.4.

**6.2. Construction of one-dimensional order parameter dynamics.** As a first example we study a function  $v(x, t)$  defined for  $x \in [0, 2\pi]$  and  $t \in [0, T]$  by the continuous transition between a finite number of linearly independent states. Consider a set of linearly independent functions  $v_k \in L^2([0, 2\pi])$  for  $q = 1, \dots, Q$  and a set of order parameters (cf. [6, 17])  $\lambda_q : [0, T] \rightarrow \mathbb{R}$ ,  $q = 1, \dots, Q$ . Often, the order parameter function  $\lambda_q(t)$  is zero outside of some compactly supported set. Here, we choose a set of points  $t_q \in [0, T]$ ,  $q = 1, \dots, Q$ , such that  $0 = t_1 < t_2 < \dots < t_Q = T$ , and determine  $\lambda_q$  to be linear in each of the intervals  $[t_q, t_{q+1}]$  for  $q = 1, \dots, Q - 1$  with

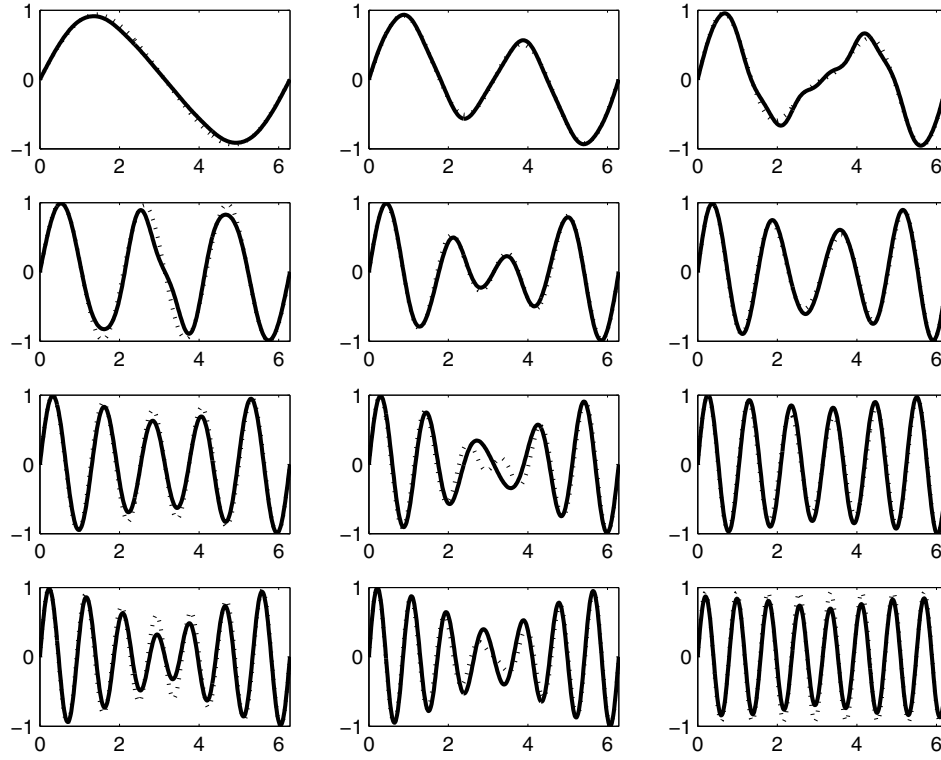
$$(99) \quad \lambda_q(t_p) = \begin{cases} 1, & p = q, \\ 0, & \text{otherwise;} \end{cases}$$

i.e., we use tent functions. We define our prescribed field by

$$(100) \quad v(x, t) = \sum_{q=1}^Q \lambda_q(t) v_q(x), \quad x \in [0, 2\pi], \quad t \in [0, T],$$

which is piecewise continuously differentiable with respect to time. Clearly, here we need to choose a one-dimensional version of the setting described in (84) to (97). For a numerical test we have chosen

$$(101) \quad v_q(x) := \sin(qx), \quad x \in [0, 2\pi],$$



**Figure 1.** As a numerical example for section 6.2 we test the solution  $v_\alpha(x, t)$  of the one-dimensional Amari equation for  $x \in [0, 2\pi]$  at 12 time slices  $t$ , where the kernel  $w_\alpha(x, y)$ ,  $x, y \in [0, 2\pi]$ , is reconstructed as described in steps 1–3 in section 6.1 with  $\ell = 100$  training patterns. The true dynamics is shown as a dotted line, and the constructed dynamics is solid, where the  $x$ -axis is the space variable and the  $y$ -axis the size of the neural fields under consideration studied without units. Here, we have used  $\alpha = 1$ , and the time discretization for the simulation used half of the grid size of the time discretization for the training patterns; i.e., for simulation we employed  $\ell = 200$  in (102). As a result, the basic features of the prescribed dynamics are successfully reproduced by the simulation with the constructed neural kernel, though some smaller features in time slice 8 are not fully captured.

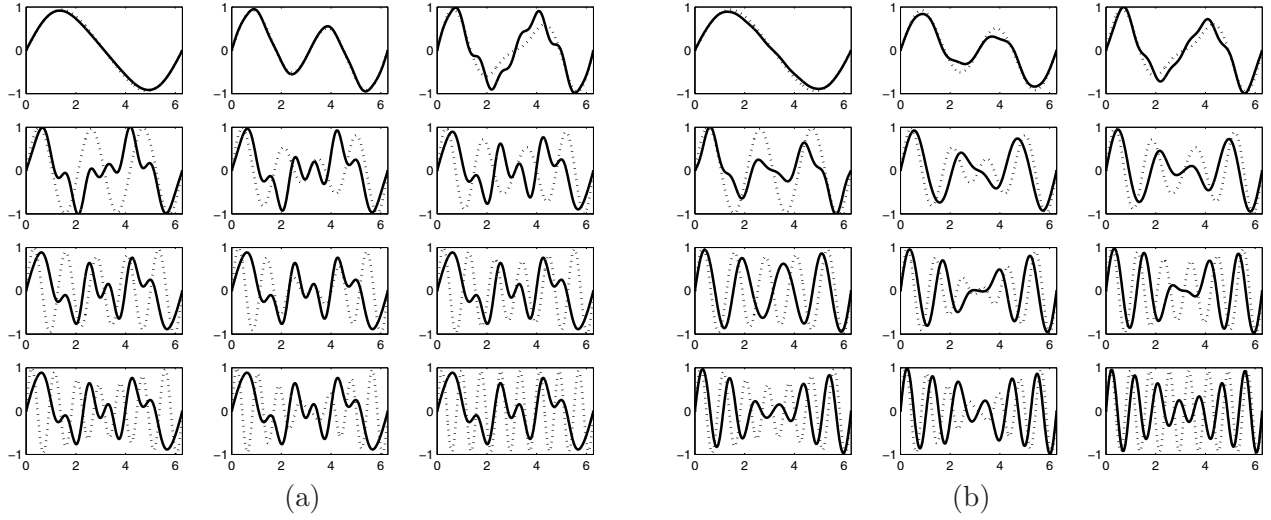
for  $q = 1, \dots, Q = 4$  or  $Q = 8$ . We employed  $N = 320$  space points and  $\ell = 100$  time steps for training, i.e., a time discretization

$$(102) \quad t_s = \frac{s}{\ell}T, \quad s = 1, \dots, \ell.$$

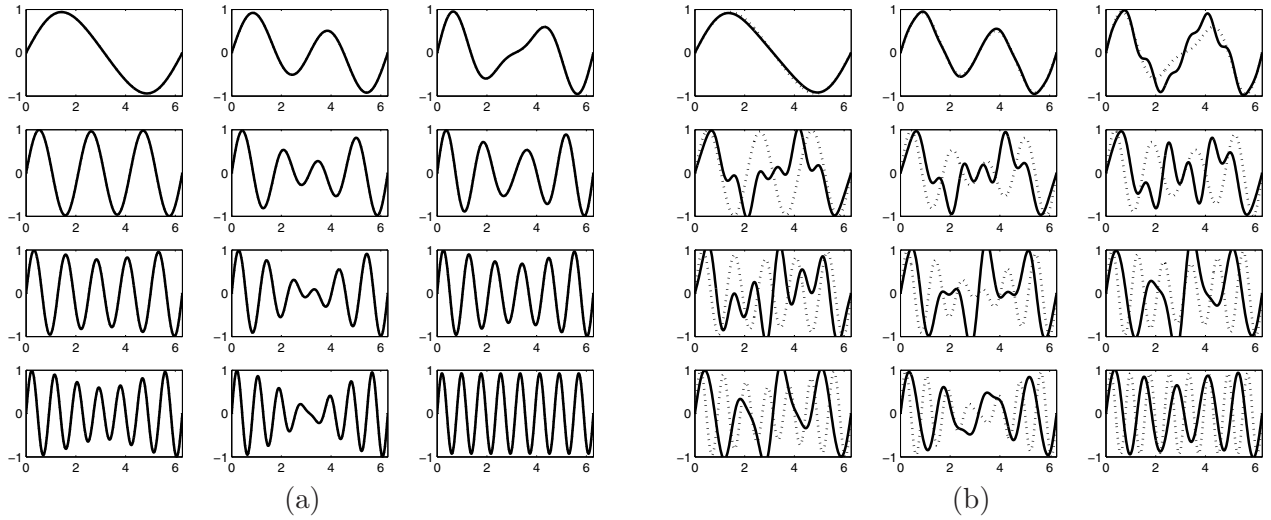
Further parameter choices have been  $\eta = 0.3$ ,  $\tau = 2$ , and  $T = 7$ . In the figures we display the function  $v(x, t)$  or  $v_\alpha(x, t)$  on a selection of *time slices* given by (102) with  $s = 1, 10, 19, 28, \dots, 91, 100$  and  $\ell = 100$ .

We will use this example to explain the choice of the parameters in more detail and to illustrate important features of the dynamical inverse problem under consideration in Figures 1, 2, 3, and 4.

*Choice of the number of discretization points  $N$  and  $\ell$ .* We need to discuss the choice of the space discretization  $N$  and number of time points  $\ell$ . Recall that we want to construct a kernel  $w$  such that the dynamics of the system defined by  $w$  via (2) reproduces a prescribed

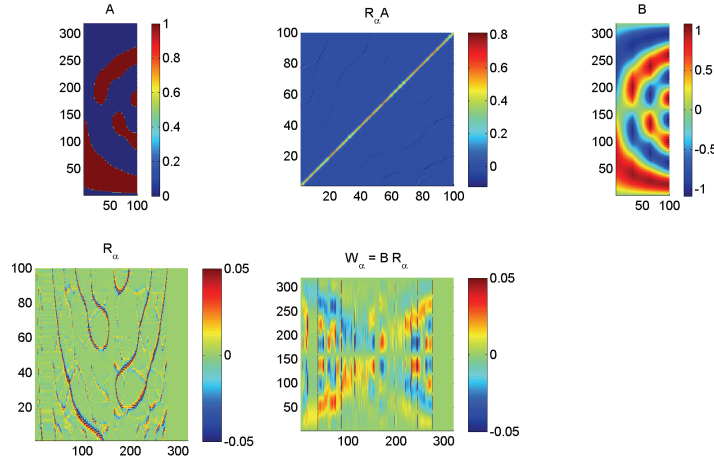


**Figure 2.** The figure demonstrates the effect of an inappropriate choice of the regularization parameter  $\alpha$ . We employ the same basic setup as in Figure 1. Here, we have used  $\alpha = 0.1$  (a) or  $\alpha = 30$  (b). The time discretization for the simulation used half of the grid size used in the time discretization for the training patterns; i.e., for simulation we employed  $\ell = 200$  in (102). Here, the regularization parameters  $\alpha = 0.1, 30$  are either too low or too large to reproduce the desired dynamics.



**Figure 3.** We illustrate an inverse crime as described in section 6.2. We employ the same basic setup as in Figure 1. Here, we have used  $\alpha = 0.01$  for both images. The time discretization for the simulation uses the same grid for training and simulation (a) which gives very good results. However, if we change the time discretization  $\ell$  in (102) for simulation (b) to a smaller grid constant, we see that the kernel which is reconstructed no longer provides the correct dynamics.

dynamics. Here we demand that the training patterns  $\varphi^{(s)}$ ,  $s = 1, \dots, \ell$ , for different time steps be linearly independent such that the matrix  $\mathbf{A}$  defined in (93) is invertible. In fact, this is not a necessary condition, but having linearly dependent training patterns generates several



**Figure 4.** We study the behavior of the matrices  $\mathbf{A}$ ,  $\mathbf{B}$ ,  $\mathbf{R}_\alpha$ ,  $\mathbf{R}_\alpha \mathbf{A}$ , and  $\mathbf{W}_\alpha$ , which are calculated in the kernel construction process for a special example defined in (100). A detailed discussion is given in section 6.2.

further questions which we avoid discussing in this work. The vectors  $\varphi^{(1)}, \dots, \varphi^{(\ell)}$ , which are elements of  $\mathbb{R}^N$  as defined in (87), can be linearly independent only if  $N \geq \ell$ . But in general we need  $N$  to be much larger than  $\ell$ , since  $\varphi^{(s)}$  is a discretized version of the function  $v(x, t_s)$ ,  $x \in [0, 2\pi]$ . If the functions  $v(\cdot, t_s)$  for  $s = 1, \dots, \ell$  are linearly independent over  $[0, 2\pi]$ , then for sufficiently large  $N$  the vectors  $\varphi^{(s)}$ ,  $s = 1, \dots, \ell$ , will be linearly independent. Here, for the particular given functions (100), given  $\ell$ , we determined  $N$  sufficiently large such that this is the case. For  $\ell = 100$  we needed to choose  $N \geq 320$ .

*Uniqueness of the kernel  $\mathbf{W}$ .* In general, the reconstructed kernel  $\mathbf{W}_\alpha$  is not uniquely determined. Only if the set of functions  $\varphi^{(1)}, \dots, \varphi^{(\ell)}$  spans the space  $\mathbb{R}^N$  is  $\mathbf{W}$  determined by the given dynamics, since in this case  $\mathbf{A}$  is bijective. Since mostly  $N \gg \ell$ , this condition will not be satisfied. Also, consider training patterns which are zero on some part  $V \subset D$  of the underlying set  $D$ . Then any kernel  $\tilde{w}$  which has nonzero values  $\tilde{w}(x, y)$  for  $x, y \in V$  only will not change the dynamics for these training patterns; i.e., it is in the nullspace of our inverse problem. We can add  $\tilde{w}$  to any solution and obtain an equivalent solution. Further, functions of the type  $\tilde{w}$  cannot be reconstructed from the given data.

*Inverse crimes and the time discretization for simulation.* An *inverse crime* is a well-known phenomenon in the theory of inverse problems; compare [11]. It arises if you solve an inverse problem using a particular choice of discretization parameters and then test it by solving the forward problem with the same choice of discretization parameters or vice versa. In this case the match is usually very good, since one basically inverts a finite-dimensional linear system. Even if the system is ill-conditioned in this case, the error is small. However, if one uses a different choice of discretization parameters for reconstruction and testing, the errors in the reconstruction play a much stronger role and the match is usually much worse.

For dynamical inverse problems, inverse crimes arise when the time discretization for training the kernels and the time discretization for simulation are identical. In this case one obtains a nearly perfect fit of the prescribed and the simulated curves. We illustrate

this in Figure 3(a), where we carry out the simulation with  $\ell = 100$ . We then change the time discretization for the simulation (keeping  $\ell = 100$  training patterns) to  $\ell = 200$  for the simulation time steps. The result as displayed in Figure 3(b) shows that there is no match at all—the instability of the problem fully destroys the dynamics. Here we worked with a regularization parameter  $\alpha = 0.01$ , which is too small to produce stable solutions. Stable solutions are obtained with  $\alpha = 1$ , which is demonstrated in Figure 1.

*Choice and influence of the regularization parameter  $\alpha$ .* We also illustrate different choices of the regularization parameter  $\alpha$  in Figures 1, 2, and 3. There are various methods for the automatic choice of  $\alpha$  in the theory of inverse problems, but it is also common to start with manual studies of the behavior of reconstructions when new problems are addressed. Here, we show the behavior of reconstructions for different  $\alpha$  in Figures 1, 2, and 3. The reconstructions with  $Q = 8$ ,  $N = 320$ , and  $\ell = 100$  training patterns are stable for values around  $\alpha = 1$ . For values smaller than  $\alpha = 0.5$  the instabilities of the reconstruction impede the desired dynamical behavior. For values larger than  $\alpha = 10$  the damping is too large, and we do not get the desired results. Usually, the practicable values of  $\alpha$  differ strongly depending on the choice of parameters and the particular problem under consideration.

*Insight into the behavior of  $\mathbf{R}_\alpha$  and  $\mathbf{W}_\alpha$ .* Finally, we illustrate the particular behavior of the matrices  $\mathbf{R}_\alpha$  and  $\mathbf{W}_\alpha$  for this example in Figure 4. For better visibility here we have used  $Q = 4$  functions only. The matrices  $\mathbf{A}$  and  $\mathbf{B}$  show the time behavior of the functions under consideration, where the time is given by the  $x$ -axis and the  $y$ -axis shows the space dimension. The product  $\mathbf{R}_\alpha \mathbf{A}$  is shown as a test; it is a damped version of the identity matrix. The most interesting numerical result here is the particular form of the matrix  $\mathbf{R}_\alpha$ , shown in the bottom left of Figure 4. We observe that the algorithm generates a kind of *edge detection*. If we multiply a row  $\mathbf{a}$  of  $\mathbf{A}$  with  $\mathbf{R}_\alpha$ , it will basically generate nonzero contributions where we have a change of values between neighboring entries of  $\mathbf{a}$  at places where you see the red-blue lines in the graphical display of  $\mathbf{R}_\alpha$ . This acts as a space-sensitive edge detection algorithm to distinguish between the different input functions at different time steps. The matrix  $\mathbf{W}_\alpha$  is shown for the sake of completeness; we observe that it does not show particular interesting features.

**6.3. Two-dimensional pulse reconstruction.** Next, we demonstrate inversion by constructing kernels for two-dimensional pulses. First, we have chosen a particular kernel

$$(103) \quad w(x, y) := e^{-R|x-y|^2} \cdot \chi_{x \geq y}(x, y) - \chi_{x_1 < y_1}(x, y) - \chi_{x_2 < y_2}(x, y), \quad x, y \in D,$$

where we use

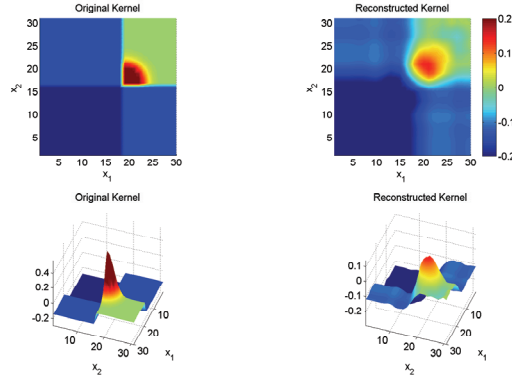
$$(104) \quad \chi_{\text{condition}}(x, y) = \begin{cases} 1, & \text{condition is satisfied,} \\ 0, & \text{otherwise,} \end{cases}$$

and  $x \geq y$  if and only if  $x_1 \geq y_1$  and  $x_2 \geq y_2$ . The kernel times the grid constants  $h_1, h_2$  for fixed  $y$  dependent on  $x$  is shown in Figure 5. A full plot of the kernel depending on  $\mathbf{x}$  and  $\mathbf{y}$  is shown in Figure 6(right).

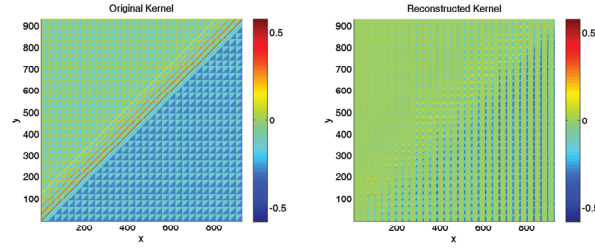
We construct  $v$  by starting with some initial values given by

$$(105) \quad v(x, 0) = \begin{cases} 1, & |x - x_0| \leq c, \\ 0, & \text{otherwise,} \end{cases}$$





**Figure 5.** Synaptic weight kernel reconstruction. Left column: a particular kernel generating a moving pulse is prescribed. Right column: kernel reconstructed by Tikhonov regularization. Both kernels are plotted for one fixed  $y$  dependent on  $x$ .



**Figure 6.** Original and reconstructed full weight kernel  $w(x,y)$  as defined in (95).

with some constant  $c > 0$  and initial pulse center  $x_0 \in D$ . Then, the time dependent pulse is calculated by solving the forward problem (2). As training patterns we use

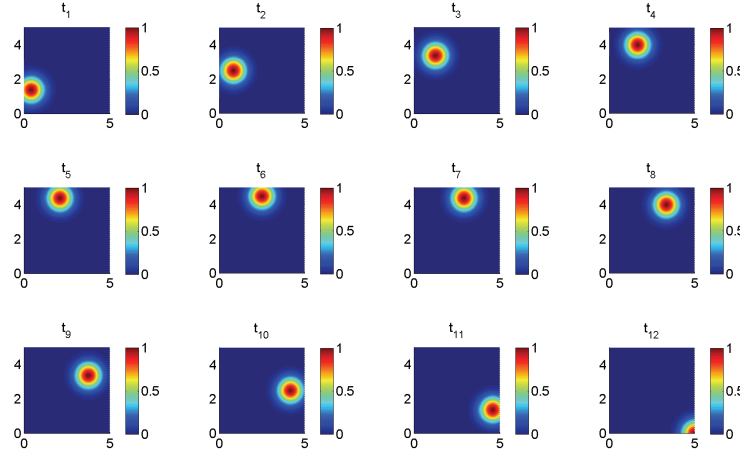
$$(106) \quad \varphi_1 := v(\cdot, t_1), \dots, \varphi_\ell := v(\cdot, t_\ell)$$

with  $t_1 \leq t_2 \leq t_3 \leq \dots \leq t_\ell$ . The reconstructed kernel  $\mathbf{W}_\alpha(\cdot, y)$  is zero if  $y$  is not in the path of the pulse  $v$ , which is expected since we calculate a minimum norm solution to match the given data. If  $y$  is in the path of  $v(\cdot, t)$ , the reconstruction has been a smoothed version of the original kernel, as shown in Figure 5(b). This effect is well known from various inverse problems [13, 11, 29], where regularization leads to a smoothing such that reconstructed functions are a smoothed version of the true solutions. If  $y$  is not in the path of the prescribed pulse, we expect the reconstructed kernel  $w(\cdot, y)$  to vanish, since the Tikhonov regularization calculates a solution with minimal norm. This is confirmed by our results.

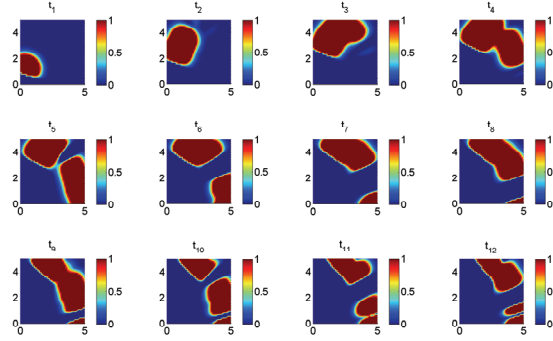
Second, we have chosen a *Gaussian pulse* by

$$(107) \quad v(x, t) := e^{-R|x-x_0(t)|^2}, \quad t \in [0, T],$$

following a parabolic path  $x_0(t) := (q_1 t, q_2 t - q_3 t^2)$  with constants  $R$ ,  $q_1$ ,  $q_2$ , and  $q_3$ . We have then calculated  $\varphi$  and  $\psi$  from (12) and (11) on the grid (85). As training patterns we



**Figure 7.** Original traveling neural pulse for kernel construction problem.

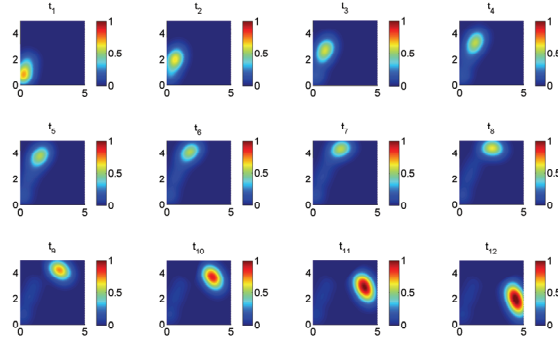


**Figure 8.** Traveling neural pulse for kernel constructed by Moore–Penrose pseudoinverse (without regularization) as given by (95).

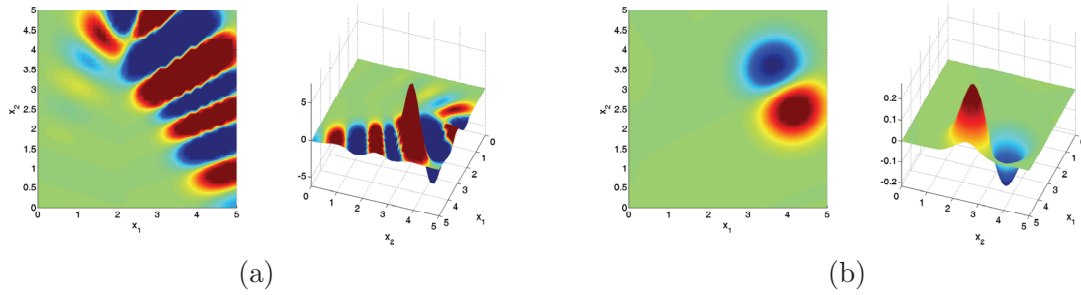
use (106) at the times  $t_1, \dots, t_\ell$ . Then, we construct the kernel  $\mathbf{W}$  using (97). The original pulse at sample times  $t_1, \dots, t_{12}$  is shown in Figure 7. We show a reconstructed pulse without regularization in Figure 8 and a regularized construction in Figure 9. The corresponding synaptic weight kernels are shown in Figure 10.

Clearly, even without regularization, if we keep the grid in space and time fixed, we are able to calculate an exact match for the pulse at the times  $t_1, \dots, t_\ell$ . This is a particular form of an *inverse crime* [11], where we consider original and reconstruction in the same fixed discretization. Here, we change the time discretization for calculating the pulses after reconstruction. In general the ill-conditioning of the matrix  $\mathbf{W}$  leads to the high oscillations in the reconstructed kernel as shown in Figure 8 or 10(a). However, with appropriate regularization (where we use  $\alpha = 10^3$  in this section), we obtain nice smooth kernels as shown in Figure 10(b) and a pulse which follows the predefined path as shown in Figure 9.

**6.4. Construction of logical gates.** Our final example shows the construction of a logical gate with XOR functionality; i.e., we construct a classical *nonlinear input-output relation* such



**Figure 9.** Traveling neural pulse for kernel constructed by Tikhonov regularization (97) with  $\alpha = 10^3$ .

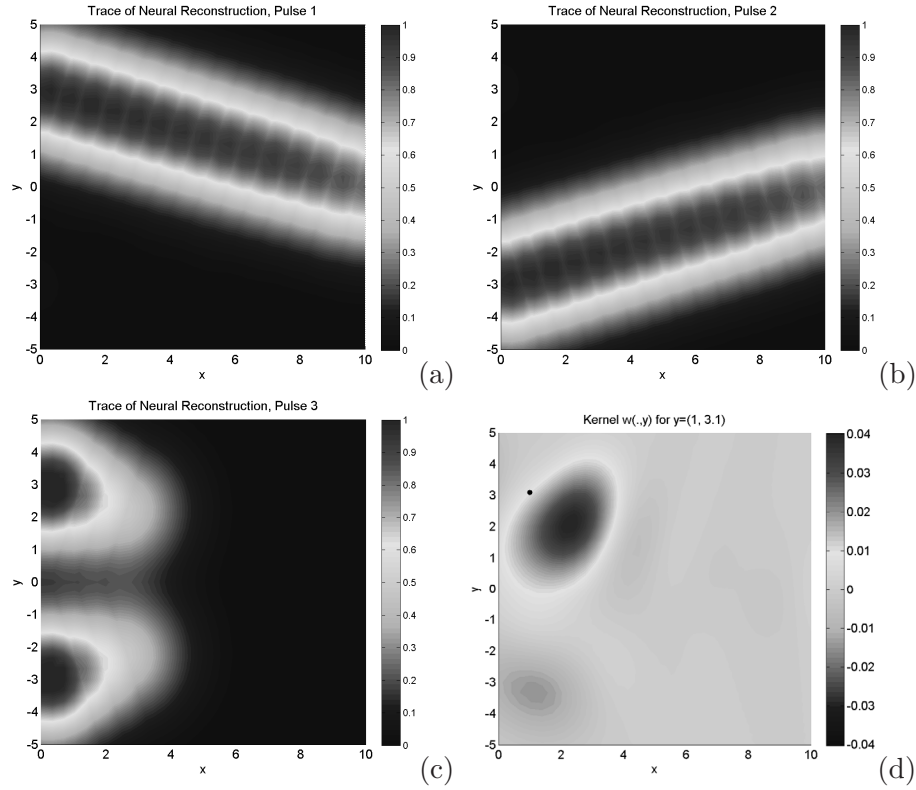


**Figure 10.** Unregularized (a) and regularized (b) synaptic weight kernels for the traveling neural pulse inverse problem.

that  $(1,0)$  is mapped onto 1,  $(0,1)$  is mapped onto 1, and both  $(0,0)$  and  $(1,1)$  are mapped onto 0. The input is a pulse into the domain  $D$  either around the point  $(0,3)$  or around  $(0,-3)$ . Output is collected at the point  $(10,0)$ .

For solving the inverse problem we employed the setting as described in (84) to (97) in the version for multiple pulses (98). Here, we have used a discretization of  $30 \times 31$  points in the domain  $D$  and 80 time discretization points for the inverse problem. The pulse had the form given in (107). We used regularization with  $\alpha = 1$ . The tests in Figure 11 have been carried out with 160 time discretization points to avoid an *inverse crime*. Here, a trace of the pulse is shown; i.e., we plotted all the time slices in the same figure. The result shows that the construction of a logical gate is possible with the linear techniques described in the above sections.

**7. Conclusion.** In this paper, we discussed inverse problems for a broad class of neural field theories described by the Amari equation (2), where synaptic weight kernels are to be constructed from prescribed solutions of the forward problem. By recasting the Amari equation into a linear perceptron equation in function space, we were able (a) to analyze the ill-posedness of the inverse problem, (b) to study generalized Hebbian learning as an appropriate linear training algorithm, and (c) to employ Tikhonov regularization for stable numerical implementations. We demonstrated the technique by means of three instructive



**Figure 11.** We show a trace of the constructed potential for three different neural pulses establishing a logical gate with the **XOR** functionality; i.e., either pulse entering at  $(0, 3)$  or  $(0, -3)$  is transmitted through the neural tissue layer to the point  $(10, 0)$ , but both pulses inserted simultaneously lead to extinction; i.e., they are not transmitted. The kernel  $w(\cdot, y)$  for  $y = (1, 3.1)$  is shown in (d)—here an active potential in the point  $(1, 3.1)$  inhibits the potential at points around  $(1, -3)$ .

examples: (i) temporal transitions between static spatial activation patterns governed by an order parameter dynamics, (ii) kernel reconstruction for a traveling Gaussian pulse, and (iii) solution of the logical XOR problem by linear means.

To our knowledge inverse problems in neural field theory have not been widely studied yet. Most work has been done in studying the forward problem under particular assumptions such as homogeneous and isotropic convolution kernels [1, 9, 10, 12, 15, 16, 21, 22, 33, 40, 42]. On the other hand, highly nonlinear and computationally expensive gradient inversion methods, such as backpropagation, are used in connectionist applications of neural networks, e.g., in order to solve the XOR problem [5, 7, 19, 25, 27, 28, 36, 38]. In contrast, our study shows that such persistent problems can be straightforwardly tackled by linear methods within the domain of neural field theory.

Therefore, we think that reliable solutions of neural field inverse problems on different levels of brain modeling can provide important new mathematical input into the extremely important field of cognitive neurodynamics. Our work provides a first step in this direction by bringing together results from different communities and suggesting stable methods for solving important basic inverse problems for neural field equations.

## REFERENCES

- [1] S.-I. AMARI, *Dynamics of pattern formation in lateral-inhibition type neural fields*, Biolog. Cybernet., 27 (1977), pp. 77–87.
- [2] J. A. ANDERSON AND E. ROSENFELD, EDS., *Neurocomputing. Foundations of Research*, Vol. 1, MIT Press, Cambridge, MA, 1988.
- [3] P. BEIM GRABEN, *Foundations of neurophysics*, in Lectures in Supercomputational Neuroscience: Dynamics in Complex Brain Networks, P. beim Graben, C. Zhou, M. Thiel, and J. Kurths, eds., Springer, Berlin, 2008, pp. 3–48.
- [4] P. BEIM GRABEN AND J. KURTHS, *Simulating global properties of electroencephalograms with minimal random neural networks*, Neurocomputing, 71 (2008), pp. 999–1007.
- [5] P. BEIM GRABEN, T. LIEBSCHER, AND J. KURTHS, *Neural and cognitive modeling with networks of leaky integrator units*, in Lectures in Supercomputational Neuroscience: Dynamics in Complex Brain Networks, P. beim Graben, C. Zhou, M. Thiel, and J. Kurths, eds., Springer, Berlin, 2008, pp. 195–223.
- [6] P. BEIM GRABEN, D. PINOTSIS, D. SADDY, AND R. POTTHAST, *Language processing with dynamic fields*, Cognitive Neurodynamics, 2 (2008), pp. 79–88.
- [7] H. BERSINI, M. SAERENS, AND L. G. SOTELINO, *Hopfield net generation, encoding and classification of temporal trajectories*, IEEE Trans. Neural Networks, 5 (1994), pp. 945–953.
- [8] E. BOASSO, *On the Moore-Penrose inverse, EP Banach space operators, and EP Banach algebra elements*, J. Math. Anal. Appl., 339 (2008), pp. 1003–1014.
- [9] M. BREAKSPEAR, J. A. ROBERTS, J. R. TERRY, S. RODRIGUES, N. MAHANT, AND P. A. ROBINSON, *A unifying explanation of primary generalized seizures through nonlinear brain modeling and bifurcation analysis*, Cerebral Cortex, 16 (2006), pp. 1296–1313.
- [10] P. C. BRESSLOFF, *Bloch waves, periodic feature maps, and cortical pattern formation*, Phys. Rev. Lett., 89 (2002), paper 088101.
- [11] D. COLTON AND R. KRESS, *Inverse Acoustic and Electromagnetic Scattering Theory*, Springer-Verlag, New York, Berlin, 1998.
- [12] S. COOMBES, G. J. LORD, AND M. R. OWEN, *Waves and bumps in neuronal networks with axo-dendritic synaptic interactions*, Phys. D, 178 (2003), pp. 219–241.
- [13] H. W. ENGL, M. HANKE, AND A. NEUBAUER, *Regularization of Inverse Problems*, Springer, Berlin, 1996.
- [14] W. ERLHAGEN AND G. SCHÖNER, *Dynamic field theory of movement preparation*, Psych. Rev., 109 (2002), pp. 545–572.
- [15] G. B. ERMENTROUT AND J. B. MCLEOD, *Existence and uniqueness of travelling waves for a neural network*, Proc. Roy. Soc. Edinburgh Sect. A, 123 (1993), pp. 461–478.
- [16] J. S. GRIFFITH, *A field theory of neural nets: I. Derivation of field equations*, Bull. Math. Biophys., 25 (1963), pp. 111–120.
- [17] H. HAKEN, *Synergetics. An Introduction*, Springer, New York, 1983.
- [18] D. O. HEBB, *The Organization of Behavior*, Wiley, New York, 1949; partly reprinted in [2].
- [19] J. HERTZ, A. KROGH, AND R. G. PALMER, *Introduction to the Theory of Neural Computation*, Perseus Books, Cambridge, MA, 1991.
- [20] J. J. HOPFIELD, *Neural networks and physical systems with emergent collective computational abilities*, Proc. Natl. Acad. Sci. USA, 79 (1982), pp. 2554–2558.
- [21] A. HUTT AND F. M. ATAY, *Analysis of nonlocal neural fields for both general and gamma-distributed connectivities*, Phys. D, 203 (2005), pp. 30–54.
- [22] V. K. JIRSA AND H. HAKEN, *Field theory of electromagnetic brain activity*, Phys. Rev. Lett., 77 (1996), pp. 960–963.
- [23] R. KRESS, *Linear Integral Equations*, Springer, Berlin, 1989.
- [24] R. KRESS, *Numerical Analysis*, Springer, Berlin, 1998.
- [25] T. LIEBSCHER, *Modeling reaction times with neural networks using leaky integrator units*, in Proceedings of the 18th Twente Workshop on Language Technology (TWLT 18), K. Jokinen, D. Heylen, and A. Nijholt, eds., University of Twente, Twente, The Netherlands, 2000, pp. 81–94.
- [26] M. MINSKY AND S. PAPERT, *Perceptrons*, MIT Press, Cambridge, MA, 1969; partly reprinted in [2].
- [27] B. A. PEARLMUTTER, *Learning state space trajectories in recurrent neural networks*, Neural Comput., 1 (1989), pp. 263–269.

- [28] B. A. PEARLMUTTER, *Gradient calculations for dynamic recurrent neural networks: A survey*, IEEE Trans. Neural Networks, 6 (1995), pp. 1212–1228.
- [29] R. POTTHAST, *Point-Sources and Multipoles in Inverse Scattering Theory*, Chapman & Hall, London, 2001.
- [30] R. POTTHAST, *Topical review: A survey on sampling and probe methods for inverse problems*, Inverse Problems, 22 (2006), pp. R1–R47.
- [31] R. POTTHAST AND P. BEIM GRABEN, *Existence and properties of solutions for neural field equations*, Math. Methods Appl. Sci., to appear; DOI: 10.1002/mmm.1199.
- [32] K. A. RICHARDSON, S. J. SCHIFF, AND B. J. GLUCKMAN, *Control of traveling waves in the mammalian cortex*, Phys. Rev. Lett., 94 (2005), paper 028103.
- [33] P. A. ROBINSON, C. J. RENNIE, AND J. J. WRIGHT, *Propagation and stability of waves of electrical activity in the cerebral cortex*, Phys. Rev. E, 56 (1997), pp. 826–840.
- [34] F. ROSENBLATT, *The perceptron: A probabilistic model for information storage and organization in the brain*, Phys. Rev., 65 (1958), pp. 386–408; also reprinted in [2].
- [35] G. SCHÖNER AND E. THELEN, *Using dynamic field theory to rethink infant habituation*, Psych. Rev., 113 (2006), pp. 273–299.
- [36] L. G. SOTELINO, M. SAERENS, AND H. BERSINI, *Classification of temporal trajectories by continuous-time recurrent nets*, Neural Networks, 7 (1994), pp. 767–776.
- [37] R. B. STEIN, K. V. LEUNG, D. MANGERON, AND M. N. OĞUZTÖRELI, *Improved neuronal models for studying neural networks*, Kybernetik, 15 (1974), pp. 1–9.
- [38] G.-Z. SUN, H.-H. CHEN, AND Y.-C. LE, *A fast online learning algorithm for recurrent neural networks*, in Proceedings of the International Joint Conference on Neural Networks (IJCNN 91), Vol. 2, 1991, pp. 13–18.
- [39] E. THELEN, G. SCHÖNER, C. SCHEIER, AND L. B. SMITH, *The dynamics of embodiment: A field theory of infant perseverative reaching*, Behavioral Brain Sci., 24 (2001), pp. 1–86.
- [40] N. A. VENKOV, S. COOMBES, AND P. C. MATTHEWS, *Dynamic instabilities in scalar neural field equations with space-dependent delays*, Phys. D, 232 (2007), pp. 1–15.
- [41] H. R. WILSON AND J. D. COWAN, *Excitatory and inhibitory interactions in localized populations of model neurons*, Biophys. J., 12 (1972), pp. 1–24.
- [42] H. R. WILSON AND J. D. COWAN, *A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue*, Kybernetik, 13 (1973), pp. 55–80.