COMPUTING DISCRETE CONVOLUTIONS WITH VERIFIED ACCURACY VIA BANACH ALGEBRAS AND THE FFT

JEAN-PHILIPPE LESSARD, Montreal

Received March 26, 2018. Published online May 23, 2018.

Dedicated to the 50th birthday of Sergey Korotov

Abstract. We introduce a method to compute rigorous component-wise enclosures of discrete convolutions using the fast Fourier transform, the properties of Banach algebras, and interval arithmetic. The purpose of this new approach is to improve the implementation and the applicability of computer-assisted proofs performed in weighed ℓ^1 Banach algebras of Fourier/Chebyshev sequences, whose norms are known to be numerically unstable. We introduce some application examples, in particular a rigorous aposteriori error analysis for a steady state in the quintic Swift-Hohenberg PDE.

Keywords: discrete convolutions; Banach algebras; fast Fourier transform; interval arithmetic; rigorously verified numerics; quintic Swift-Hohenberg PDE

MSC 2010: 65G40, 65T50, 46J15, 46B45, 42B05

1. INTRODUCTION

In this paper, we introduce a method to compute discrete convolutions with verified accuracy using the fast Fourier transform (FFT), the properties of Banach algebras of bi-infinite complex valued sequences and interval arithmetic. Our motivation comes from the field of rigorously verified numerics in dynamics (see e.g. [9], [16], [19], [21], [23], [25], [27], [30]), which aims at obtaining computer-assisted proofs (CAPs) of existence of solutions of differential equations (ODE, PDEs, delay equations, etc.) and discrete dynamical systems (iterations of finite and infinite dimensional maps). One of the common approaches in obtaining the CAPs for differential equations is to represent the solutions using Fourier/Chebyshev series, and to apply the contraction

The research has been supported by NSERC.

mapping theorem on a ball centered at a numerical approximation in a Banach space of Fourier/Chebyshev coefficients with weighed ℓ^1 -norms (see e.g. [8], [15], [26], [28]). More explicitly, the Banach spaces considered in the aforementioned papers are sequence spaces of the form

(1.1)
$$\ell_{\omega}^{1} \stackrel{\text{def}}{=} \{a = \{a_k\}_{k \in \mathbb{Z}} \colon a_k \in \mathbb{C} \text{ and } \|a\|_{\omega} < \infty\}$$

with norm

(1.2)
$$||a||_{\omega} \stackrel{\text{def}}{=} \sum_{k \in \mathbb{Z}} |a_k| \omega_k,$$

where $\omega = \{\omega_k\}_{k \in \mathbb{Z}}$ is a sequence of positive weights satisfying the sub-multiplicative property $\omega_{n+k} \leq \omega_n \omega_k$ for all $n, k \in \mathbb{Z}$. The advantage of using the sequence space (1.1) in the CAPs is essentially twofold. First, ℓ_{ω}^1 has a known dual (a weighed ℓ^{∞} space) which facilitates estimating norms of functionals and operators. Second, the sub-multiplicative property of the weights ensures that ℓ_{ω}^1 is a Banach algebra under discrete convolutions (see Section 3), which facilites performing the nonlinear analysis. However, choosing the Banach space ℓ_{ω}^1 in performing the CAPs has the disadvantage that its norm $\|\cdot\|_{\omega}$ can be unstable with respect to a given sequence of weights. For instance, in the papers [12], [17], [26] the numerical instability of the ℓ^1 -norms with geometrically growing weights prevented obtaining some CAPs (see e.g. the discussion in Section 6 of [12] or the discussion in Section 6.3 in [26]). The goal of this paper is precisely to address this issue. In order to formalize the problem, we need a bit of notation and background on discrete convolutions.

Assume that two functions $u_1, u_2: \mathbb{R} \to \mathbb{R}$ with period $2\pi/L$ (that is with frequency L > 0) have absolutely converging Fourier series expansions

$$u_1(t) = \sum_{k \in \mathbb{Z}} a_k^{(1)} \mathrm{e}^{\mathrm{i}kLt}$$
 and $u_2(t) = \sum_{k \in \mathbb{Z}} a_k^{(2)} \mathrm{e}^{\mathrm{i}kLt}.$

Then their product $u_1(t)u_2(t)$ also has an absolutely converging Fourier series expansion, which is given by

$$u_1(t)u_2(t) = \sum_{k \in \mathbb{Z}} (a^{(1)} * a^{(2)})_k \mathrm{e}^{\mathrm{i}kLt},$$

where, given two sequences $a = \{a_k\}_{k \in \mathbb{Z}}$ and $b = \{b_k\}_{k \in \mathbb{Z}}$, a * b denotes their discrete convolution defined component-wise by

(1.3)
$$(a*b)_k = \sum_{\substack{k_1+k_2=k\\k_1,k_2 \in \mathbb{Z}}} a_{k_1} b_{k_2}$$

R e m a r k 1.1. The approach presented in the present paper can also help controlling the coefficients of product of functions represented by Chebyshev series. Indeed, as Chebyshev series are Fourier series "in disguise" [1], their product naturally leads to discrete convolutions (see e.g. [18]).

For many applications and for some computer-assisted proofs, we are interested in computing the product of trigonometric polynomials with high but finite degree, that is Fourier series which have only finitely many non zero coefficients. More explicitly, we consider $2\pi/L$ -periodic trigonometric polynomials $u_i(t)$ (i = 1, ..., p) of degree M - 1 defined by

(1.4)
$$u_i(t) = \sum_{|k| < M} a_k^{(i)} e^{ikLt}, \quad i = 1, \dots, p.$$

As the following result demonstrates, the product of *p*-periodic trigonometric polynomials of degree M - 1 has degree p(M - 1).

Lemma 1.2. For i = 1, ..., p, let u_i be the $2\pi/L$ -periodic trigonometric polynomials of degree M - 1 with expansion (1.4). The Fourier expansion of the product $u_1u_2...u_p$ satisfies

(1.5)
$$u_1(t)u_2(t)\dots u_p(t) = \sum_{|k| \le p(M-1)} (a^{(1)} * a^{(2)} * \dots * a^{(p)})_k e^{ikLt}.$$

In other words, the function $u_1u_2...u_p$ is a trigonometric polynomial of degree p(M-1).

Proof. Let $k \in \mathbb{Z}$ be such that $k = k_1 + \ldots + k_p$ for some $k_1, \ldots, k_p \in \mathbb{Z}$ with $|k_i| < M$ $(i = 1, \ldots, p)$. Hence,

$$k = k_1 + \ldots + k_p \in \{-p(M-1), \ldots, p(M-1)\},\$$

and therefore $|k| \leq p(M-1)$. We conclude that

$$u_1(t)u_2(t)\dots u_p(t) = \sum_{k\in\mathbb{Z}} (a^{(1)} * a^{(2)} * \dots * a^{(p)})_k e^{ikLt}$$
$$= \sum_{|k|\leqslant p(M-1)} (a^{(1)} * a^{(2)} * \dots * a^{(p)})_k e^{ikLt}.$$

We are ready to state the goal of the present paper.

Statement of the problem: Given p finite sequences of Fourier/Chebyshev coefficients $\{a_k^{(i)}\}_{|k| < M}$ (i = 1, ..., p) combine interval arithmetic, the fast Fourier

transform (FFT) and Banach algebras to obtain rigorous component-wise enclosures for the discrete convolution

(1.6)
$$(a^{(1)} * a^{(2)} * \dots * a^{(p)})_k = \sum_{\substack{k_1+k_2+\dots+k_p=k\\|k_i|< M}} a^{(1)}_{k_1} a^{(2)}_{k_2} \dots a^{(p)}_{k_p} \text{ for } |k| \leq p(M-1).$$

The statement of the problem comes from the rather common need in rigorous numerics to compute interval enclosures for terms of the form (1.6) (see e.g. [5], [4], [6], [10], [31], [32]). While the FFT algorithm [29] is a fantastic tool to evaluate quickly discrete convolutions (see e.g. Section 2), it often fails to recover the geometric/algebraic decay of the tail (i.e. the terms in (1.6) corresponding to $M \leq |k| \leq p(M-1)$) of a convolution (see e.g. the example in Section 5.1) due to round-off errors. This property implies that the weighted ℓ^1 -norms of the discrete convolutions may blow-up for geometrically growing weights (e.g. when $\omega_k = \nu^{|k|}$ for some $\nu > 1$) (see Tables 4 and 5). This is a major hurdle in obtaining the CAPs in the category of analytic functions, as sometimes taking ν rather large is necessary to obtain a contraction mapping. In order to fix this issue in the tail, we use the fundamental property of a given Banach algebra (X, *) (i.e. $||x*y||_X \leq ||x||_X ||y||_X$ for all $x, y \in X$) to obtain the proper decay in the tail. This has the effect of stabilizing the sensitivity of the ℓ^1 -norms in obtaining the rigorous bounds for the computer-assisted proofs.

Before proceeding any further, we urge to mention that the present work is by no means the first time that the FFT algorithm and interval arithmetic are combined to rigorously compute discrete convolutions (see e.g. [3], [11], [14], [9]). However, we believe that our new proposed approach of combining the FFT algorithm, interval arithmetic and theoretical Banach algebra estimates to obtain rigorous componentwise enclosure (1.6) is new, and that it could benefit the rigorous numerics community.

R e m a r k 1.3. The idea introduced in this paper can be generalized to rigorously enclose components of discrete convolutions of multidimensional sequences in the Banach space

(1.7)
$$\ell_{\omega}^{1} \stackrel{\text{def}}{=} \{a = \{a_{\alpha}\}_{\alpha \in \mathbb{Z}^{d}} \colon a_{\alpha} \in \mathbb{C} \text{ and } \|a\|_{\omega} < \infty\},$$

where $d \in \mathbb{N}$ is the dimension of the space on which solutions of the differential equations are defined, $||a||_{\omega} \stackrel{\text{def}}{=} \sum_{\alpha \in \mathbb{Z}^d} |a_{\alpha}|\omega_{\alpha}$ and $\omega = \{\omega_{\alpha}\}_{\alpha \in \mathbb{Z}^d}$ is a sequence of positive weights satisfying the sub-multiplicative property $\omega_{\alpha+\beta} \leq \omega_{\alpha}\omega_{\beta}$ for all $\alpha, \beta \in \mathbb{Z}^d$. It can be shown that ℓ^1_{ω} is a Banach algebra under the discrete convolution defined by

$$(a * b)_{\alpha} = \sum_{\substack{\beta + \gamma = \alpha \\ \beta, \gamma \in \mathbb{Z}^d}} a_{\beta} b_{\gamma} = \sum_{\beta \in \mathbb{Z}^d} a_{\beta} b_{\alpha - \beta}.$$

The paper is organized as follows. In Section 2, we present the background of how to use the FFT to compute discrete convolutions. In Section 3, we present how to use the Banach algebra property of the Banach space ℓ^1_{ω} to obtain an alternative method to compute interval enclosures for terms of the form (1.6), which recover a similar decay rate in the tail as the one of the inputs. In Section 4, we combine the two approaches of Sections 2 and 3 to refine the rigorous enclosures of the discrete convolutions. Finally in Section 5, we introduce some application examples, where we compute rigorously discrete convolutions of degree 3, 20 and 100, and present a rigorous aposteriori error analysis for a steady state in the quintic Swift-Hohenberg PDE.

2. Convolution enclosures via the FFT and interval arithmetic

In this section, denote by M and p the order of the Fourier series and the power of the nonlinearity, respectively. Following closely the presentation of [11], we introduce the theory to compute discrete convolutions of the form (1.6) with the discrete Fourier transform (DFT). Once this is done, we combine the FFT algorithm (an efficient implementation of the DFT) and interval arithmetic to compute rigorous enclosures of discrete convolutions of the form (1.6).

Definition 2.1. Given $b = (b_0, \ldots, b_{2M-2}) \in \mathbb{C}^{2M-1}$, define its discrete Fourier transform $\mathcal{F}(b) \in \mathbb{C}^{2M-1}$ by

$$a_k = \mathcal{F}_k(b) \stackrel{\text{def}}{=} \sum_{j=0}^{2M-2} b_j \mathrm{e}^{-2\pi \mathrm{i} j k/(2M-1)} \quad \text{for } k = -M+1, \dots, M-1.$$

Definition 2.2. Given $a = (a_k)_{|k| < M} = (a_{-M+1}, \ldots, a_{M-1}) \in \mathbb{C}^{2M-1}$, define its inverse discrete Fourier transform $\mathcal{F}^{-1}(a) \in \mathbb{C}^{2M-1}$ by

$$b_j = \mathcal{F}_j^{-1}(a) \stackrel{\text{def}}{=} \sum_{k=-M+1}^{M-1} a_k \mathrm{e}^{2\pi \mathrm{i} j k/(2M-1)} \quad \text{for } j = 0, \dots, 2M-2.$$

Given $a^{(i)} = (a_k^{(i)})_{|k| < M} \in \mathbb{C}^{2M-1}$, we extend it to eliminate the so-called *aliasing* error (see the second term in (2.4)). Hence, define $\tilde{a}^{(i)} \in \mathbb{C}^{2pM-1}$ by

(2.1)
$$\widetilde{a}_{j}^{(i)} = \begin{cases} a_{j}^{(i)} & \text{for } |j| < M, \\ 0 & \text{for } M \leq |j| \leq pM - 1, \end{cases}$$

that is we pad the vector $a^{(i)}$ with (p-1)M zeros before and after. Then define $\tilde{b}^{(i)} = (b_0^{(i)}, \ldots, b_{2pM-2}^{(i)}) \in \mathbb{C}^{2pM-1}$ component-wise by

(2.2)
$$\tilde{b}_{j}^{(i)} \stackrel{\text{def}}{=} \mathcal{F}_{j}^{-1}(\tilde{a}^{(i)}) = \sum_{k=-pM+1}^{pM-1} \tilde{a}_{k}^{(i)} \mathrm{e}^{2\pi \mathrm{i} jk/(2pM-1)} \quad \text{for } j = 0, \dots, 2pM-2.$$

Define $\tilde{b}^{(1)} \cdot \ldots \cdot \tilde{b}^{(p)}$ to be the component-wise product of the vectors $\tilde{b}^{(1)}, \ldots, \tilde{b}^{(p)}$, that is

(2.3)
$$(\tilde{b}^{(1)} * \dots * \tilde{b}^{(p)})_j \stackrel{\text{def}}{=} \tilde{b}^{(1)}_j \dots \tilde{b}^{(p)}_j, \quad j = 0, \dots, 2pM - 2.$$

Hence, for $k = -pM + 1, \dots, pM - 1$,

$$\mathcal{F}_{k}(\tilde{b}^{(1)} \,\hat{*} \dots \hat{*} \,\tilde{b}^{(p)}) = \sum_{j=0}^{2pM-2} \tilde{b}_{j}^{(1)} \dots \tilde{b}_{j}^{(p)} \mathrm{e}^{-2\pi \mathrm{i} jk/(2pM-1)}$$
$$= \sum_{j=0}^{2pM-2} \prod_{i=1}^{p} \sum_{k_{i}=-pM+1}^{pM-1} \tilde{a}_{k_{i}}^{(i)} \mathrm{e}^{2\pi \mathrm{i} jk_{i}/(2pM-1)} \,\mathrm{e}^{-2\pi \mathrm{i} jk/(2pM-1)}.$$

Letting

$$S_{k}(j) \stackrel{\text{def}}{=} \prod_{i=1}^{p} \sum_{\substack{k_{i}=-pM+1 \\ k_{i}=-pM+1}}^{pM-1} \tilde{a}_{k_{i}}^{(i)} e^{2\pi i j k_{i}/(2pM-1)} e^{-2\pi i j k/(2pM-1)}$$

$$= \sum_{\substack{k_{1}+\ldots+k_{p}=k \\ |k_{i}|$$

we obtain that

$$(2.4) \quad \mathcal{F}_{k}(\tilde{b}^{(1)} \, \hat{\ast} \, \dots \hat{\ast} \, \tilde{b}^{(p)}) = \sum_{j=0}^{2pM-2} S_{k}(j) = (2pM-1) \sum_{\substack{k_{1}+\dots+k_{p}=k \\ |k_{i}| < M}} \sum_{\substack{k_{1}+\dots+k_{p} \neq k \\ |k_{i}| < M}} a_{k_{1}}^{(1)} \dots a_{k_{p}}^{(p)} + (2pM-1) \sum_{l=1}^{p} \sum_{\substack{k_{1}+\dots+k_{p}=k \\ |k_{i}| < M}} a_{k_{1}}^{(1)} \dots a_{k_{p}}^{(p)} \sum_{j=0}^{2pM-2} e^{2\pi i \frac{k_{1}+\dots+k_{p}-k}{2pM-1}j} .$$

Euler's formula gives that for $k_1 + \ldots + k_p - k \not\equiv 0 \pmod{2pM - 1}$,

$$\sum_{j=0}^{2pM-1} e^{2\pi i j(k_1 + \dots + k_p - k)/(2pM-1)} = 0.$$

Hence, the third sum in (2.4) is zero. As far as the second sum in (2.4) is concerned, observe that $|k_1|, \ldots, |k_p| < M$ and that $|k| \leq p(M-1)$, and therefore

$$k_1 + \ldots + k_p - k \in \{-2p(M-1), \ldots, 2p(M-1)\}$$

Hence, for the equality $k_1 + \ldots + k_p = k \pm l(2pM - 1)$ to be satisfied for a choice of $l \in \{1, \ldots, p\}$, one must have that

$$l(2pM - 1) = k_1 + \ldots + k_p - k \in \{-2p(M - 1), \ldots, 2p(M - 1)\},\$$

which is impossible, because 2pM - 1 > 2p(M - 1). Hence, the second sum of (2.4) is zero. Therefore, we can conclude that

(2.5)
$$\sum_{\substack{k_1+\ldots+k_p=k\\|k_1|,\ldots,|k_p|$$

Remark 2.3. In [11], we padded the vectors $a^{(i)}$ as

$$\widetilde{a}_{j}^{(i)} = \begin{cases} a_{j}^{(i)} & \text{for } |j| < M, \\ 0 & \text{for } M \leqslant |j| \leqslant \delta_{p}M - 1, \end{cases}$$

where

$$\delta_p \stackrel{\text{def}}{=} \begin{cases} \frac{p+1}{2} & \text{if } p \text{ is odd,} \\ \\ \frac{p+2}{2} & \text{if } p \text{ is even,} \end{cases}$$

because we only considered the cases |k| < M in (1.6).

The inverse discrete Fourier transforms and the discrete Fourier transforms required in the computations of (2.2) and (2.5) can be computed efficiently and rigorously using the FFT algorithm (see e.g. [2]) and interval arithmetic (e.g. [20]) for instance using the function verifyfft.m in INTLAB [22].

3. Convolution enclosures via Banach Algebras

We begin this section by defining the notion of a Banach algebra.

Definition 3.1. A *Banach algebra* is a Banach space X with a multiplication operation $*: X \times X \to X$ that satisfies

(3.1)

$$\begin{aligned} x * (y * z) &= (x * y) * z, \\ (x + y) * z &= x * z + y * z, \quad x * (y + z) = z * y + x * z, \\ \alpha(x * y) &= (\alpha x) * y = x * (\alpha y), \\ \|x * y\| &\leq \|x\| \|y\|, \end{aligned}$$

for all $x, y, z \in X$ and all scalars α . The Banach algebra is *commutative* if x * y = y * x, for all $x, y \in X$.

Recall the definition of the Banach space ℓ_{ω}^1 in (1.1) endowed with the norm $\|\cdot\|_{\omega}$ defined in (1.2), and the definition of the discrete convolution in (1.3).

The following definition introduces a property about the weight ω which makes ℓ^1_{ω} a commutative Banach algebra under discrete convolutions.

Definition 3.2. A sequence of positive real numbers $\omega = {\{\omega_k\}_{k \in \mathbb{Z}} \text{ is an admissible sequence of weights if } \omega_{n+k} \leq \omega_n \omega_k \text{ for all } n, k \in \mathbb{Z}.$

Lemma 3.3. Given ω an admissible sequence of weights, denote by * the discrete convolution. Then the pair $(\ell_{\omega}^1, *)$ is a commutative Banach algebra.

Proof. We only demonstrate that $||a * b||_{\omega} \leq ||a||_{\omega} ||b||_{\omega}$ for all $a, b \in \ell_{\omega}^1$. Indeed,

$$\begin{aligned} \|a * b\|_{\omega} &= \sum_{k \in \mathbb{Z}} |(a * b)_{k}| \omega_{k} = \sum_{k \in \mathbb{Z}} \left| \sum_{k_{1} \in \mathbb{Z}} a_{k_{1}} b_{k-k_{1}} \right| \frac{\omega_{k}}{\omega_{k-k_{1}}} \omega_{k-k_{1}} \\ &\leqslant \sum_{k \in \mathbb{Z}} \sum_{k_{1} \in \mathbb{Z}} |a_{k_{1}}| \|b_{k-k_{1}}\| \omega_{k_{1}} \omega_{k-k_{1}} \\ &\leqslant \sum_{k_{1} \in \mathbb{Z}} |a_{k_{1}}| \omega_{k_{1}} \sum_{k_{2} \in \mathbb{Z}} |b_{k_{2}}| \omega_{k_{2}} = \|a\|_{\omega} \|b\|_{\omega}. \end{aligned}$$

Example 3.4. Consider a geometric decay rate $\nu \ge 1$ and an algebraic decay rate s > 0. Let $\omega = {\omega_k}_{k \in \mathbb{Z}}$ be the sequence of positive real numbers defined component-wise by

(3.2)
$$\omega_k \stackrel{\text{def}}{=} (|k|+1)^s \nu^{|k|}.$$

Then ω is an admissible sequence of weights. Indeed, for any $n, k \in \mathbb{Z}$,

$$\begin{split} \omega_{n+k}^{1/s} &= (|n+k|+1)\nu^{|n+k|/s} \\ &\leqslant (|n||k|+|n|+|k|+1)\nu^{|n|/s}\nu^{|k|/s} \\ &= (|n|+1)\nu^{|n|/s} \cdot (|k|+1)\nu^{|k|/s} = \omega_n^{1/s}\omega_k^{1/s} = (\omega_n\omega_k)^{1/s} \end{split}$$

which implies that $\omega_{n+k} \leq \omega_n \omega_k$, since the function $f(x) \stackrel{\text{def}}{=} x^{1/s}$ is strictly monotone increasing on $(0, \infty)$ for any s > 0.

The following consequence of Lemma 3.3 allows to obtain rigorous theoretical bounds on any component of a discrete convolution.

Corollary 3.5. Fix $p \in \mathbb{N}$. Let $a^{(1)}, \ldots, a^{(p)} \in \ell^1_{\omega}$. Then for any $k \in \mathbb{Z}$

(3.3)
$$|(a^{(1)} * a^{(2)} * \dots * a^{(p)})_k| \leq ||a^{(1)}||_{\omega} ||a^{(2)}||_{\omega} \dots ||a^{(p)}||_{\omega} \frac{1}{\omega_k}.$$

Proof. Fix $k \in \mathbb{Z}$. Then by Lemma 3.3,

$$|(a^{(1)} * a^{(2)} * \dots * a^{(p)})_k|\omega_k \leq \sum_{j \in \mathbb{Z}} |(a^{(1)} * a^{(2)} * \dots * a^{(p)})_j|\omega_j$$

= $||a^{(1)} * a^{(2)} * \dots * a^{(p)}||_{\omega} \leq ||a^{(1)}||_{\omega} ||a^{(2)}||_{\omega} \dots ||a^{(p)}||_{\omega}.$

4. RIGOROUS CONVOLUTION ENCLOSURES: A REFINEMENT

We are now ready to present a refinement of the rigorous enclosures of the discrete convolutions presented in Section 2. Note that in practice, the method of Section 2 is essentially used for the finite part (i.e. |k| < M), while the method of Section 3 is essentially used for the tail part (i.e. for $M \leq |k| \leq p(M-1)$).

Combining the FFT algorithm with interval arithmetic (see e.g. the function verifyfft.m in INTLAB [22]), we can use formula (2.5) to compute a rigorous enclosure $\mathcal{B}_{k}^{(\text{FFT})} \in \mathbb{C}$ such that

(4.1)
$$\sum_{\substack{k_1+k_2+\ldots+k_p=k\\|k_i|< M}} a_{k_1}^{(1)} a_{k_2}^{(2)} \ldots a_{k_p}^{(p)} = \frac{1}{2pM-1} \cdot \mathcal{F}_k(\tilde{b}^{(1)} \hat{*} \ldots \hat{*} \tilde{b}^{(p)}) \in \mathcal{B}_k^{(\text{FFT})}$$

for all $k = -p(M-1), \ldots, p(M-1)$, where the vectors $\tilde{b}^{(1)}, \ldots, \tilde{b}^{(p)}$ are defined in (2.2). Now, choose an admissible sequence of weights $\omega \stackrel{\text{def}}{=} {\{\omega_k\}_{k \in \mathbb{Z}}}$. For all integers $k = -p(M-1), \ldots, p(M-1)$, let

(4.2)
$$\beta_k \stackrel{\text{def}}{=} \|a^{(1)}\|_{\omega} \|a^{(2)}\|_{\omega} \dots \|a^{(p)}\|_{\omega} \omega_k^{-1},$$

and set

(4.3)
$$\mathcal{B}_{k}^{(\omega)} \stackrel{\text{def}}{=} \{ z \in \mathbb{C} \colon |z| \leq \beta_{k} \}$$

From Corollary 3.5, we get that

(4.4)
$$\sum_{\substack{k_1+k_2+\ldots+k_p=k\\|k_i|< M}} a_{k_1}^{(1)} a_{k_2}^{(2)} \ldots a_{k_p}^{(p)} \in \mathcal{B}_k^{(\omega)}.$$

R e m a r k 4.1. We see from (4.4) that the quality of the enclosure of $\mathcal{B}_{k}^{(\omega)}$ depends on the choice of the admissible sequence of weights $\omega \stackrel{\text{def}}{=} \{\omega_k\}_{k \in \mathbb{Z}}$. In practice, we may know in advance that the type of functions we are approximating is analytic, in which case setting s = 0 and $\nu > 1$ in the weights (3.2) is most suitable. A least square fit can be used to determine the optimal geometric decay rate $\nu > 1$ fitting the data. Once this choice has been identified, set $\omega_k = \nu^{|k|}$, compute β_k using (4.2) and define $\mathcal{B}_k^{(\omega)}$. Similarly, we may know in advance that the type of functions we are approximating is only C^s , in which case setting s > 0 and $\nu = 1$ in the weights (3.2) is most suitable. A least square fit can be used to determine the optimal algebraic decay rate s > 1 fitting the data. Once this choice has been identified, set $\omega_k = (|k| + 1)^s$, compute β_k using (4.2) and define $\mathcal{B}_k^{(\omega)}$.

For each $k \in \mathbb{Z}$ such that $|k| \leq p(M-1)$, we refine the enclosure of the discrete convolution by setting

(4.5)
$$\mathcal{B}_{k} \stackrel{\text{def}}{=} \mathcal{B}_{k}^{(\text{FFT})} \cap \mathcal{B}_{k}^{(\omega)}.$$

5. Examples

In this final section, we show case how the refinement (4.5) of Section 4 helps to stabilize computing ℓ_{ω}^1 norms of discrete convolutions of the form (1.6). We present four application examples. We compute rigorously convolutions of degree 3 (Section 5.1), 20 (Section 5.2) and 100 (Section 5.3), and demonstrate how our new approach improves dramatically the standard method of Section 2. Finally in Section 5.4, we present a rigorous aposteriori error analysis for a steady state in the quintic Swift-Hohenberg PDE, and again compare the two approaches.

5.1. Cubic discrete convolutions. Let us consider a vector of Fourier/Chebyshev coefficients $\{a_k\}_{|k| < M}$ with a prescribed geometric decay rate $\rho > 1$. We want to compute a rigorous enclosure for cubic discrete convolutions of the form

(5.1)
$$(a^3)_k = (a * a * a)_k = \sum_{\substack{k_1 + k_2 + k_3 = k \\ |k_i| \le M}} a_{k_1} a_{k_2} a_{k_3} \quad \text{for } |k| \le 3(M-1).$$

Fix p = 3, M = 30 and $\varrho = 4$. Consider a symmetric vector $\{\bar{a}_k\}_{|k| < M}$ such that $\bar{a}_{-k} = \bar{a}_k$ and $\bar{a}_k \in \mathbb{R}$. For the coefficients \bar{a}_k $(k = 0, \ldots, 29)$, we consider $\bar{a}_k = \alpha_k \varrho^{-k}$, where $\alpha_k \in [-1, 1]$ are chosen randomly. We compute rigorous componentwise bounds of $(\bar{a}^3)_k$ given in (5.1) for all $|k| \leq 3(M-1) = 87$ using the standard approach of Section 2 (rigorous FFT only) and the refinement of Section 4 (rigorous FFT intersected with the Banach algebra bounds). In Figure 1 on the left, we show the plot of both rigorous component-wise enclosures for $\ln(|(\bar{a}^3)_k|)$ for $|k| \leq p(M-1) = 3(29) = 87$. In Table 1, we consider the interval vectors $\mathcal{B}^{(\text{FFT})}$ and $\mathcal{B} \stackrel{\text{def}}{=} \mathcal{B}^{(\text{FFT})} \cap \mathcal{B}^{(\omega)}$ (having each 87 interval components) and compute a rigorous upper bound for their norm in ℓ^1_{ω} with $\omega = \{\nu^{|k|}\}_{|k| \leq 87}$ for different values of $\nu > 1$. We also compare the bounds with the theoretical upper bound $\|\bar{a}^3\|_{\omega} \leq \|\bar{a}\|_{\omega}^3$.



Figure 1. Two rigorous component-wise enclosures for: (left) $\ln(|(\bar{a}^3)_k|)$ for $|k| \leq 87$; (right) $\ln(|(\bar{a}^{20})_k|)$ for $|k| \leq 980$.

5.2. Discrete convolution of degree 20. We now want to compute a rigorous enclosure for discrete convolutions of degree 20 of the form

(5.2)
$$(a^{20})_k = (\underbrace{a*\ldots*a}_{20 \text{ times}})_k = \sum_{\substack{k_1+\ldots+k_{20}=k\\|k_i|< M}} a_{k_1}\ldots a_{k_{20}} \text{ for } |k| \leq 20(M-1).$$

ν	$\ \mathcal{B}^{(\mathrm{FFT})}\ _{\omega}$	$\ \mathcal{B}^{(\mathrm{FFT})}\cap\mathcal{B}^{(\omega)}\ _{\omega}$	$\ \bar{a}\ ^3_{\omega}$
1	9.9559×10^{-2}	9.9559×10^{-2}	1.6643×10^{-1}
1.25	1.8288×10^{-1}	1.8287×10^{-1}	3.9832×10^{-1}
1.5	2.5805×10^{0}	2.7271×10^{-1}	8.7256×10^{-1}
1.6	5.6391×10^2	2.7271×10^{-1}	1.1745×10^{0}
1.7	1.0032×10^5	4.0439×10^{-1}	1.5701×10^{0}
1.8	1.3436×10^7	4.9123×10^{-1}	2.0881×10^{0}
1.9	1.3926×10^9	5.9573×10^{-1}	2.7661×10^{0}
2	1.1450×10^{11}	7.2135×10^{-1}	3.6548×10^{0}
2.5	2.5847×10^{19}	1.8758×10^{0}	1.4778×10^{1}
3	1.8058×10^{26}	1.4199×10^1	6.8826×10^{1}
3.5	1.1272×10^{32}	2.8576×10^3	5.5624×10^2
4	1.1936×10^{37}	1.9558×10^6	2.8233×10^4

Table 1. Three different upper bounds for $\|\bar{a}^3\|_{\omega}$. In the second and third columns, rigorous upper bounds for the ℓ^1_{ω} norms of the interval vectors $\mathcal{B}^{(\text{FFT})}$ and $\mathcal{B} \stackrel{\text{def}}{=} \mathcal{B}^{(\text{FFT})} \cap \mathcal{B}^{(\omega)}$ (having each 87 interval components) with $\omega = \{\nu^{|k|}\}_{|k| \leq 87}$ for different values of $\nu \geq 1$. In the fourth column, we compare the bounds with the theoretical upper bound $\|\bar{a}^3\|_{\omega} \leq \|\bar{a}\|_{\omega}^3$.

Fix p = 20, M = 50 and $\rho = 3$. Consider a symmetric vector $\{\bar{a}_k\}_{|k| < M}$ such that $\bar{a}_{-k} = \bar{a}_k$ and $\bar{a}_k \in \mathbb{R}$. For the coefficients \bar{a}_k $(k = 0, \ldots, 49)$, we consider $\bar{a}_k = \alpha_k \rho^{-k}$, where $\alpha_k \in [-1/2, 1/2]$ are chosen randomly. We compute rigorous component-wise bounds of $(\bar{a}^{20})_k$ given in (5.2) for all $|k| \leq p(M-1) = 980$ using the standard approach of Section 2 and the refinement of Section 4. In Figure 1 on the right, we show the plot of both rigorous component-wise enclosures for $\ln(|(\bar{a}^{20})_k|)$ for $|k| \leq p(M-1) = p(M-1) = 980$. As in Section 5.1, we provide comparisons in Table 2.

ν	$\ \mathcal{B}^{(\mathrm{FFT})}\ _{\omega}$	$\ \mathcal{B}^{(\mathrm{FFT})} \cap \mathcal{B}^{(\omega)}\ _{\omega}$	$\ \bar{a}\ _{w}^{20}$
1	3.6095×10^{-11}	3.6095×10^{-11}	2.7979×10^{-2}
1.05	2.0193×10^{-3}	4.7327×10^{-11}	3.3327×10^{-2}
1.1	6.6628×10^{16}	$6.3808 imes 10^{-11}$	3.9653×10^{-2}
1.2	3.9165×10^{53}	1.2718×10^{-10}	5.5972×10^{-2}
1.5	1.8339×10^{148}	2.0877×10^{-6}	1.5603×10^{-1}
1.7	2.7672×10^{201}	1.0043×10^{0}	3.1316×10^{-1}

Table 2. Three different upper bounds for $\|\bar{a}^{20}\|_{\omega}$. In the second and third columns, rigorous upper bounds for the ℓ_{ω}^1 norms of the interval vectors $\mathcal{B}^{(\text{FFT})}$ and $\mathcal{B} \stackrel{\text{def}}{=} \mathcal{B}^{(\text{FFT})} \cap \mathcal{B}^{(\omega)}$ (having each 980 interval components) with $\omega = \{\nu^{|k|}\}_{|k| \leq 980}$ for different values of $\nu \geq 1$. In the fourth column, we compare the bounds with the theoretical upper bound $\|\bar{a}^{20}\|_{\omega} \leq \|\bar{a}\|_{\omega}^{20}$.

5.3. Discrete convolutions of degree 100. We now compute a rigorous enclosure for discrete convolutions of degree 100 of the form

(5.3)
$$(a^{100})_k = (\underbrace{a * \dots * a}_{100 \text{ times}})_k = \sum_{\substack{k_1 + \dots + k_{100} = k \\ |k_i| < M}} a_{k_1} \dots a_{k_{100}} \text{ for } |k| \leq 100(M-1).$$

Fix p = 100, M = 11 and $\rho = 30$. Consider a symmetric vector $\{\bar{a}_k\}_{|k| < M}$ such that $\bar{a}_{-k} = \bar{a}_k$ and $\bar{a}_k \in \mathbb{R}$. For the coefficients \bar{a}_k (k = 0, ..., 1000), we consider $\bar{a}_k = \alpha_k \rho^{-k}$, where $\alpha_k \in [-1, 1]$ are chosen randomly. We compute rigorous component-wise bounds of $(\bar{a}^{100})_k$ given in (5.2) for all $|k| \leq p(M-1) = 1000$ using the standard approach of Section 2 and the refinement of Section 4. We provide comparisons in Table 3.

ν	$\ \mathcal{B}^{(\mathrm{FFT})}\ _{\omega}$	$\ \mathcal{B}^{(\mathrm{FFT})} \cap \mathcal{B}^{(\omega)}\ _{\omega}$	$\ \bar{a}\ _{\omega}^{20}$
1	4.0676×10^{-31}	4.0676×10^{-31}	1.2570×10^{-1}
1.1	3.1261×10^{-2}	4.7431×10^{-31}	1.2812×10^{-1}
1.2	1.0476×10^{36}	5.5291×10^{-31}	1.3060×10^{-1}
1.3	4.3735×10^{70}	6.6089×10^{-31}	1.3315×10^{-1}
1.4	5.4031×10^{102}	2.4845×10^{-29}	1.3576×10^{-1}
1.5	4.2542×10^{132}	2.3256×10^{-26}	1.3845×10^{-1}
1.6	4.0393×10^{160}	1.0333×10^1	1.4121×10^{-1}

Table 3. Three different upper bounds for $\|\bar{a}^{100}\|_{\omega}$.

5.4. Aposteriori error analysis in the quintic Swift-Hohenberg PDE. The quintic Swift-Hohenberg PDE (see e.g. [13], [24]) defined on a bounded interval with even periodic boundary conditions is given by

(5.4)
$$u_t = (\lambda - 1)u - 2u_{xx} - u_{xxxx} + \mu u^3 - u^5 \quad \text{in } \Omega = [0, 2\pi/L],$$
$$u(x, t) = u(x + 2\pi/L, t), \quad u(x, t) = u(-x, t) \quad \text{on } \partial\Omega.$$

The solutions can be expressed via the Fourier expansion

(5.5)
$$u(x,t) = \sum_{k=-\infty}^{\infty} a_k(t) e^{ikLx} = a_0 + 2\sum_{k=1}^{\infty} a_k(t) \cos(kLx),$$

where $a_k \in \mathbb{R}$ and $a_{-k} = a_k$. Plugging (5.5) in (5.4) results in the infinite set of ODEs given by

(5.6)
$$\dot{a}_k = F_k(a) \stackrel{\text{def}}{=} (\lambda - (1 - k^2 L^2)^2) a_k + \mu(a^3)_k - (a^5)_k,$$

where $a^3 = a * a * a$ and $a^5 = a * a * a * a * a$ are cubic and quintic discrete convolutions, respectively. Since $a_{-k} = a_k$, hence $F_{-k} = F_k$. Looking for equilibria of the Swift-Hohenberg PDE (5.4) is equivalent to computing solutions of F(a) = 0 in ℓ_{ω}^1 , for some admissible sequence of weights ω with geometric growth $\omega_k = \nu^{|k|}$ for some $\nu > 1$. As in [13], we consider the parameter values L = 0.1, $\mu = 3$ and $\lambda < 0$. In particular, we fix $\lambda = -1/2$. Equilibria u = u(x) of (5.4) correspond to solutions of F(a) = 0, where $a = (a_k)_{k \in \mathbb{Z}}$ is the infinite sequence of Fourier coefficients and $F = (F_k)_{k \in \mathbb{Z}}$ is given component-wise by (5.6). We consider a Galerkin projection of (5.6) of dimension M = 230 and apply Newton's method to find a numerical approximation $\bar{a} = {\bar{a}_k}_{|k|<230} \in \mathbb{R}^{459}$ ($\bar{a}_{-k} = \bar{a}_k$) such that $F(\bar{a}) \approx 0$ (see Figure 2). Applying a least square fit, we compute that the numerical solution \bar{a} has a geometric decay rate of about 1.213, that is $|\bar{a}_k| \leq C/1.213^{|k|}$.



Figure 2. Steady state of the quintic Swift-Hohenberg equation (5.4) at parameter values L = 0.1, $\mu = 3$, and $\lambda = -1/2$, where M = 230 Fourier coefficients are used to represent the solution.

One of the approaches in the field of rigorous numerics is called the radii polynomial approach [7], [15] and its goal is to demonstrate the existence (and local uniqueness) of solutions of some infinite dimensional nonlinear problems F(a) = 0posed on a given Banach space X. The idea, based on a Newton-Kantorovich type argument, is to prove that a certain Newton-like operator T(a) = a - AF(a) is a contraction mapping on a carefully chosen ball centered at a numerical approximation $\bar{a} \in X$, where A is an approximate inverse for $DF(\bar{a})$ in the sense that $\|I - ADF(\bar{a})\|_{B(X)} \ll 1$. In the works [8], [15], [26], [27], [28], the radii polynomial approach is applied on the Banach space ℓ^1_{ω} , and in the process, the quantity $\|AF(\bar{a})\|_{\omega}$ has to be bounded rigorously. The bound for $\|AF(\bar{a})\|_{\omega}$ is often denoted by Y_0 . In Table 4, we present a list of rigorous upper bounds for $\|F(\bar{a})\|_{\omega}$ with $\omega = \{\omega_k\}_{k\in\mathbb{Z}} = \{\nu^{|k|}\}_{k\in\mathbb{Z}}$ for different values of $\nu \ge 1$, where the rigorous computation of $F(\bar{a})$ is performed using the two approaches presented in this paper (1. FFT alone as introduced in Section 2; and 2. FFT and Banach algebras as introduced in Section 4). We do a similar comparison for Y_0 satisfying $||AF(\bar{a})||_{\omega} \leq Y_0$ in Table 5. In these two tables, the refinement enclosure of Section 4 is a dramatic improvement over the standard approach. As $||T(\bar{a}) - \bar{a}||_{\omega} = ||AF(\bar{a})||_{\omega} \leq Y_0$, the bounds presented in Table 5 are rigorous aposteriori error bounds for a steady state in the quintic Swift-Hohenberg PDE (5.4).

ν	$ F(\bar{a}) _{\omega}$ (FFT only)	$ F(\bar{a}) _{\omega}$ (FFT + Banach algebras)
1	7.9453×10^{-10}	2.3107×10^{-10}
1.01	4.4067×10^{-6}	1.2869×10^{-9}
1.02	1.7514×10^{-1}	1.1573×10^{-8}
1.03	8.3728×10^3	1.3328×10^{-7}
1.04	4.0427×10^8	1.7290×10^{-6}
1.05	1.8729×10^{13}	2.3820×10^{-5}
1.06	8.1458×10^{17}	3.3840×10^{-4}
1.07	3.2908×10^{22}	4.8769×10^{-3}
1.08	$1.2279 imes 10^{27}$	7.0628×10^{-2}
1.09	4.2188×10^{31}	1.0221×10^0

Table 4. Two different upper bounds for $||F(\bar{a})||_{\omega}$. In the second column, a rigorous upper FFT and interval arithmetic whereas in the third column, the bound is obtained using the refinement of Section 4.

ν	Y_0 (FFT only)	Y_0 (FFT + Banach algebras)
1	1.8068×10^{-6}	1.8068×10^{-6}
1.02	2.2227×10^{-6}	2.2227×10^{-6}
1.04	2.5844×10^{0}	2.7805×10^{-6}
1.05	1.1736×10^5	3.1380×10^{-6}
1.1	8.0385×10^{27}	2.7180×10^{-5}
1.11	2.3253×10^{32}	2.9418×10^{-4}
1.12	6.1836×10^{36}	3.9776×10^{-3}
1.13	1.5120×10^{41}	5.4766×10^{-2}

Table 5. Two different upper bounds for Y_0 satisfying $||AF(\bar{a})||_{\omega} \leq Y_0$. In the second column, a rigorous upper FFT and interval arithmetic whereas in the third column, the bound is obtained using the refinement of Section 4.

References

[1]	J. P. Boyd: Chebyshev and Fourier Spectral Methods. Dover Publications, Mineola,	
	2001.	$\mathrm{zbl}\ \mathbf{MR}$
[2]	E.Brigham: Fast Fourier Transform and Its Applications. Prentice Hall, Upper Saddle River, 1988.	
[3]	J. Cyranka: Efficient and generic algorithm for rigorous integration forward in time of dPDEs. I. J. Sci. Comput. 59 (2014), 28–52.	zbl MR doi
[4]	S. Day, Y. Hiraoka, K. Mischaikow, T. Ogawa: Rigorous numerics for global dynamics: a study of the Swift-Hohenberg equation. SIAM J. Appl. Dyn. Syst. 4 (2005), 1–31.	zbl MR doi
[5]	S. Day, O. Junge, K. Mischaikow: A rigorous numerical method for the global analysis of infinite-dimensional discrete dynamical systems. SIAM J. Appl. Dyn. Syst. 3 (2004), 117–160.	zbl <mark>MR</mark> doi
[6]	S. Day, W. D. Kalies: Rigorous computation of the global dynamics of integrodifference equations with smooth nonlinearities. SIAM J. Numer. Anal. 51 (2013), 2957–2983.	zbl <mark>MR</mark> doi
[7]	S. Day, JP. Lessard, K. Mischaikow: Validated continuation for equilibria of PDEs. SIAM J. Numer. Anal. 45 (2007), 1398–1424.	zbl <mark>MR</mark> doi
[8]	<i>JL. Figueras, R. de la Llave</i> : Numerical computations and computer assisted proofs of periodic orbits of the Kuramoto-Sivashinsky equation. SIAM J. Appl. Dyn. Syst. <i>16</i> (2017) 834–852	zhl MB doi
[9]	JL. Figueras, A. Haro, A. Luque: Rigorous computer-assisted application of KAM the-	zbl MR doi
[10]	<i>M. Gameiro, JP. Lessard</i> : Analytic estimates and rigorous continuation for equilibria of higher dimensional PDFs. J. Differ. Equations 2/0 (2010) 2327 2326	
[11]	M. Gameiro, JP. Lessard, K. Mischaikow: Validated continuation over large parameter	
[12]	<i>M. Gameiro, JP. Lessard, Y. Ricaud:</i> Rigorous numerics for piecewise-smooth systems: a functional analytic approach based on Chebyshev series. J. Comput. Appl. Math. 292	
[13]	(2016), 654–673. Y. Hiraoka, T. Ogawa: Rigorous numerics for localized patterns to the quintic Swift-	zbl MR doi
[14]	Hohenberg equation. Japan J. Ind. Appl. Math. 22 (2005), 57–75. Y. Hiraoka, T. Ogawa: An efficient estimate based on FFT in topological verification	zbl MR doi
[15]	method. J. Comput. Appl. Math. 199 (2007), 238–244. A. Hungria, JP. Lessard, J. D. Mireles James: Rigorous numerics for analytic solutions	zbl <mark>MR</mark> doi
	of differential equations: the radii polynomial approach. Math. Comput. 85 (2016), 1427–1459.	zbl MR doi
[16]	H. Koch, A. Schenkel, P. Wittwer: Computer-assisted proofs in analysis and programming in logic: A case study. SIAM Rev. 38 (1996), 565–604.	zbl MR doi
[17]	<i>JP. Lessard, J. D. Mireles James, J. Ransford</i> : Automatic differentiation for Fourier series and the radii polynomial approach. Physica D 334 (2016), 174–186.	MR doi
[18]	JP. Lessard, C. Reinhardt: Rigorous numerics for nonlinear differential equations using Chebyshev series. SIAM J. Numer. Anal. 52 (2014), 1–22.	zbl <mark>MR</mark> doi
[19]	J. D. Mireles James, K. Mischaikow: Computational proofs in dynamics. Encyclopedia of Applied and Computational Mathematics (B. Engquist, ed.). Springer, Berlin, 2015.	doi
[20] [21]	<i>R. E. Moore</i> : Interval Analysis. Prentice-Hall, Englewood Cliffs, 1966. <i>M. T. Nakao</i> : Numerical verification methods for solutions of ordinary and partial dif-	zbl MR
[22]	ferential equations. Numer. Funct. Anal. Optimization 22 (2001), 321–356. S. M. Rump: INTLAB - INTerval LABoratory. Developments in Reliable Computing	zbl <mark>MR</mark> doi
r 1	(T. Csendes, ed.). Kluwer Academic Publishers, Dordrecht, 1999, pp. 77-104. Available at http://www.ti3.tu-harburg.de/rump/intlab/.	zbl



Author's address: Jean-Philippe Lessard, McGill University, Department of Mathematics and Statistics, 805 Sherbrooke Street West, Montreal, QC, H3A 0B9, Canada, e-mail: jp.lessard@mcgill.ca.