Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

journal homepage: www.elsevier.com/locate/cam

Convergence analysis and adaptive strategy for the certified quadrature over a set defined by inequalities

Alexandre Goldsztejn[a,*], Jorge Cruz[b], Elsa Carvalho[b]

[a] CNRS, Laboratoire d'Informatique de Nantes Atlantique, Nantes, France

[b] CENTRIA, Departamento de Informática, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Portugal

ARTICLE INFO

Article history:
Received 14 December 2011
Received in revised form 30 July 2013

Keywords:
Numerical quadrature
Interval analysis
Adaptive mesh
Convergence analysis
Taylor models

ABSTRACT

This paper investigates the sufficient conditions for the asymptotic convergence of a generic branch and prune algorithm dedicated to the verified quadrature of a function in several variables. Quadrature over domains defined by inequalities, and adaptive meshing strategies are in the scope of this analysis. The framework is instantiated using certified quadrature methods based on Taylor models (i.e. Taylor approximations with rigorously bounded remainder), and reported experiments confirmed the analysis. They also show that the performances of the instantiated algorithm are comparable with current methods for certified quadrature.

1. Introduction

In the present paper, we investigate the asymptotic convergence of an adaptive splitting algorithm dedicated to rigorously enclosing the integral

$$\int_{\Omega} f(\mathbf{x})d\mathbf{x} \tag{1}$$

where $\Omega = \{\mathbf{x} \in [\mathbf{x}^{\text{Init}}] : g(\mathbf{x}) \leq 0\}$ and $[\mathbf{x}^{\text{Init}}]$ a bounded box (boldface symbols represent vectors, bracketed symbols represent intervals, e.g. $[\mathbf{x}] = ([x_1], \ldots, [x_n])$ is a box, and $d\mathbf{x}$ stands for $dx_1 \ldots dx_n$). The studied branch and prune algorithm applies some arbitrary certified quadrature methods to subdomains that are recursively split. Since certified quadrature methods are more and more accurate as the domain size decreases, the overall algorithm is expected to converge asymptotically to the exact quadrature. Such convergence study was already carried out in some restricted case, see e.g. [1] where the asymptotic convergence was proved when a regular meshing with Taylor model based certified quadrature is applied to the quadrature over a box domain. This paper generalizes the convergence analysis by providing sufficient conditions enforcing the branch and prune algorithm asymptotical convergence. These sufficient conditions apply to generic quadrature methods, handle quadrature domains defined by inequalities and adaptive meshing strategies.

Using a branch and prune algorithm clearly presents pros and cons with respect to regular meshing, whose compromise is analyzed in this paper: on the one hand, when a box is split the integral of the function of the resulting sub-boxes is computed, and thus the integral that has been computed for the initial box is useless. This leads to repeated useless computations, doubling the number of calls to the generic integration procedure in the worst case (since the number of nodes in a binary tree is at most twice the number of it leaves). On the other hand, the branch and prune algorithm presents

[*] Corresponding author. Tel.: +33 678049487.
E-mail addresses: Alexandre.Goldsztejn@univ-nantes.fr (A. Goldsztejn), jc@di.fct.unl.pt (J. Cruz), elsac@uma.pt (E. Carvalho).

0377-0427/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.cam.2013.08.035

the key advantage of allowing adaptive mesh refinement,[1] which allows both reducing the overall computational effort and improving the final enclosure (since oversplitting leads to unnecessary computations and rounding errors).

Interval analysis is introduced in Section 2 with emphasis on convergence properties. Generic quadrature inclusion functions are defined in Section 3, together with their properties related to asymptotic convergence. The branch and prune algorithm dedicated to the certified enclosure of (1) is described in Section 4. The convergence analysis of this branch and prune algorithm is investigated theoretically in Section 5 together with its rate of convergence. Finally, the framework is instantiated with Taylor model based quadrature methods: Although other methods allow computing certified quadrature (e.g. [3–5]), Taylor models are efficient and have good asymptotical convergence properties. Experiments on standard benchmarks from the literature [6–9] are presented in Section 6 to support this convergence analysis. A detailed comparison is reported with respect to several algorithms [6,10,9] dedicated to the certified enclosure of (1). Finally, Taylor models are presented in the Appendix, with emphasis on their usage to certified quadrature and their convergence.

2. Interval analysis

Interval analysis is a modern branch of numerical analysis that was born in the 60's. It consists of computing with intervals of reals instead of reals, providing a framework for handling uncertainties and verified computations (see [11–15] for a survey). Section 2.1 presents the basic definitions related to interval extensions, and Section 2.2 some properties related to their convergence.

2.1. Interval extensions

An interval is a connected subset of $\mathbb{R}$. Intervals are denoted by bracketed symbols, e.g. $[x] \subseteq \mathbb{R}$. When no confusion is possible, lower an upper bounds of an interval $[x]$ are denoted by $\underline{x} \in \mathbb{R}$ and $\overline{x} \in \mathbb{R}$, with $\underline{x} \leq \overline{x}$, i.e. $[x] = [\underline{x}, \overline{x}] = \{x \in \mathbb{R} : \underline{x} \leq x \leq \overline{x}\}$. Furthermore, a real number x will be identified with the degenerated interval $[x, x]$. The width, midpoint and magnitude of an interval are respectively defined by $\operatorname{wid}[x] := \overline{x} - \underline{x}$, $\operatorname{mid}[x] := 0.5(\overline{x} + \underline{x})$, $\operatorname{mag}[x] := \max\{|\underline{x}|, |\overline{x}|\}$.

There are two equivalent ways of defining interval vectors. On the one hand, being given two vectors $\underline{\mathbf{x}} \leq \overline{\mathbf{x}} \in \mathbb{R}^n$ (where the inequality is defined componentwise), an interval of vectors is obtained by considering

$$[\mathbf{x}] := \{\mathbf{x} \in \mathbb{R}^n : \underline{\mathbf{x}} \leq \mathbf{x} \leq \overline{\mathbf{x}}\}. \tag{2}$$

On the other hand, being given intervals $[x_i] \in \mathbb{IR}$ for $i \in \{1, \ldots, n\}$, a vector of intervals is obtained by considering

$$[\mathbf{x}] := \{\mathbf{x} \in \mathbb{R}^n : \forall i \in \{1, \ldots, n\}, x_i \in [x_i]\}. \tag{3}$$

These two definitions are obviously equivalent following the notational convention $\underline{\mathbf{x}} = (\underline{x_i})$, $\overline{\mathbf{x}} = (\overline{x}_i)$ and $[x_i] = [\underline{x_i}, \overline{x}_i]$, and will be used indifferently. The width, volume and midpoint of an interval vector are respectively defined by $\operatorname{wid}[\mathbf{x}] := \max_i \operatorname{wid}[x_i] \in \mathbb{R}$, $\operatorname{vol}[\mathbf{x}] := \prod_i \operatorname{wid}[x_i] \in \mathbb{R}$ and $\operatorname{mid}[\mathbf{x}] := 0.5(\overline{\mathbf{x}} + \underline{\mathbf{x}}) \in \mathbb{R}^n$.

Operations $\circ \in \{+, \times, -, \div\}$ are extended to intervals in the following way:

$$[x] \circ [y] := \{x \circ y : x \in [x], y \in [y]\}. \tag{4}$$

The division is defined for intervals $[\underline{y}, \overline{y}]$ that do not contain zero. Unary elementary functions $f(x)$ (like exp, ln, sin, etc.) are also extended to intervals similarly:

$$f([x]) = \{f(x) : x \in [x]\}. \tag{5}$$

All these elementary interval extensions form the interval arithmetic. As real numbers are identified to degenerated intervals, the interval arithmetic actually generalizes the real arithmetic, and mixed operations like $1 + [1, 2] = [2, 3]$ are interpreted using (4).

An interval function $[f] : \mathbb{IR}^n \longrightarrow \mathbb{IR}$ is an interval extension of the real function $f : D \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ if for all $[\mathbf{x}] \in \mathbb{IR}^n$ we have $[f]([\mathbf{x}]) \supseteq \{f(\mathbf{x}) : \mathbf{x} \in D \cap [\mathbf{x}]\}$. Thus interval extensions allow computing enclosures of real functions range over boxes. So called natural interval extensions of a function are obtained by evaluating an expression of this function for interval arguments using the interval arithmetic.

Example 1. Let $f(x, y) = x(y - x)$. The interval function $[f]([x], [y]) = [x]([y] - [x])$ is the natural interval extension of f. Hence for example

$$[f]([0, 1], [-1, 1]) = [-2, 1] \supseteq \{f(x, y) : x \in [0, 1], y \in [-1, 1]\}. \tag{6}$$

Note that the exact range is $f([0, 1], [-1, 1]) = \left[-2, \frac{1}{4}\right]$, and the natural interval extension is thus pessimistic. One central issue of interval analysis is to fight this pessimism introduced by the interval evaluation of a function.

There are other interval extensions, in particular the mean-value interval extension which uses the natural extension of the derivatives to try improving the enclosure. See [13] for details. Throughout the paper, we suppose that involved interval extensions are inclusion isotone (i.e. $[\mathbf{x}'] \subseteq [\mathbf{x}]$ implies $[f]([\mathbf{x}']) \subseteq [f]([\mathbf{x}])$), which generally holds.

[1] Such adaptive strategies are known to be of critical importance also in the context of non verified quadrature, see e.g. [2]

2.2. Definitions and properties of convergent interval extensions

An interval function $[f]$ is said convergent inside $[\mathbf{x}]$ if and only if for all sequence $([\mathbf{x}_k])_{k\in\mathbb{N}}$ such that $[\mathbf{x}_k]\subseteq[\mathbf{x}]$ and $\lim_{k\to\infty}\operatorname{wid}[\mathbf{x}_k]=0$ we have $\lim_{k\to\infty}\operatorname{wid}[f]([\mathbf{x}_k])=0$. An interval extension $[f]$ of f has an order of convergence o inside $[\mathbf{x}]$ if there exists $c>0$ such that

$$\operatorname{wid}[f]([\mathbf{x}'])-\operatorname{wid}f([\mathbf{x}'])\leq c\,(\operatorname{wid}[\mathbf{x}'])^{o} \tag{7}$$

holds for all $[\mathbf{x}']\subseteq[\mathbf{x}]$. This means that the overestimation decreases as quick as the width of the interval argument to the power o. The order of convergence 1 is also called linear order of convergence.

The interval evaluation of an expression is convergent inside any bounded $[\mathbf{x}]$ where it is defined (the boundedness hypothesis cannot be removed, e.g. $[x]^2$ is not convergent inside $\mathbb{R}$ for the sequence $\left(\left[k-\frac{1}{k},k+\frac{1}{k}\right]\right)_{k\in\mathbb{N}}$). If furthermore f is Lipschitz continuous inside $[\mathbf{x}]$ then the interval evaluation of its expression has a linear order of convergence inside $[\mathbf{x}]$ (see [13] for details).

An additional consequence of the Lipschitz continuity of f inside $[\mathbf{x}]$ is

$$\operatorname{wid}f([\mathbf{x}'])\leq\max_{\mathbf{x},\mathbf{x}'\in[\mathbf{x}']}|f(\mathbf{x})-f(\mathbf{x}')|\leq\lambda\max_{\mathbf{x},\mathbf{x}'\in[\mathbf{x}']}\|\mathbf{x}-\mathbf{x}'\|\leq\lambda\operatorname{wid}[\mathbf{x}'] \tag{8}$$

for some $\lambda\geq 0$. Therefore, a linearly convergent interval extension $[f]$ of a Lipschitz continuous function f satisfies

$$\exists c'\geq 0,\quad\forall[\mathbf{x}']\subseteq[\mathbf{x}],\ \operatorname{wid}[f]([\mathbf{x}'])\leq c'\operatorname{wid}[\mathbf{x}'], \tag{9}$$

where $c'=c+\lambda$. The Lipschitz continuity hypothesis cannot be removed, e.g. $\sqrt{[x]}$ is an interval extension of $f(x)=\sqrt{x}$ that does not satisfy (9) (consider the sequence $\left(\left[0,\frac{1}{k}\right]\right)_{k\in\mathbb{N}}$).

Roughly speaking, the Proposition 1 below states that a convergent interval extension $[g]$ of g allows computing an arbitrarily sharp enclosure of the set $\{\mathbf{x}\in[\mathbf{x}]:g(\mathbf{x})=0\}$. First, the following lemma is necessary.

Lemma 1. *Consider a continuous function $g:\mathbb{R}^n\longrightarrow\mathbb{R}$, an interval extension $[g]$ of g which is convergent inside a box $[\mathbf{x}]$ and a sequence of boxes $([\mathbf{x}_k])_{k\in\mathbb{N}}$ such that $[\mathbf{x}_k]\subseteq[\mathbf{x}]$ and $\operatorname{wid}[\mathbf{x}_k]$ converges to zero. If there exists a sequence $(\mathbf{x}_k)_{k\in\mathbb{N}}$, $\mathbf{x}_k\in[\mathbf{x}_k]$, that has an accumulation point $\mathbf{x}_*$ satisfying $g(\mathbf{x}_*)\neq 0$ then there exists $k\in\mathbb{N}$ such that $0\notin[g]([\mathbf{x}_k])$.*

Proof. We can pick up a subsequence $(\mathbf{x}_{\pi(k)})_{k\in\mathbb{N}}$ (i.e. $\pi:\mathbb{N}\to\mathbb{N}$ is strictly increasing) that converges to $\mathbf{x}_*$, and since g is continuous $y_k:=g(\mathbf{x}_{\pi(k)})$ converges to $y_*:=g(\mathbf{x}_*)$. Now since $[g]$ is an interval extension of g we have $y_k\in[g]([\mathbf{x}_{\pi(k)}])=:[y_k]$. Now for any $y\in[y_k]$ we have

$$|y-y_*|\leq|y-y_k|+|y_k-y_*|\leq\operatorname{wid}[y_k]+|y_k-y_*|. \tag{10}$$

As both terms of the last sum converge to zero, we have proved that $\max_{y\in[y_k]}|y-y_*|$ converges to zero. Finally since $y_*\neq 0$ there must exist k such that $0\notin[y_k]$. □

We can now prove Proposition 1.

Proposition 1. *Consider a continuous function $g:\mathbb{R}^n\to\mathbb{R}$ and an interval extension $[g]$ of g that is convergent inside a bounded $[\mathbf{x}]$. Let $\Omega_0=\{\mathbf{x}\in[\mathbf{x}]:g(\mathbf{x})=0\}$ and consider an arbitrary set $\Omega_0^+\subseteq\mathbb{R}^n$ such that $\Omega_0\subseteq\operatorname{int}\Omega_0^+$, where $\operatorname{int}E$ is the interior of E. Then there exists $\epsilon>0$ such that for all $[\mathbf{x}']\subseteq[\mathbf{x}]$ we have $\operatorname{wid}[\mathbf{x}']\leq\epsilon$ and $0\in[g]([\mathbf{x}'])$ imply $[\mathbf{x}']\subseteq\Omega_0^+$.*

Proof. The proof is carried out by contradiction: Suppose that for all $\epsilon>0$ there exist a box $[\mathbf{x}_\epsilon]\subseteq[\mathbf{x}]$ that satisfies $\operatorname{wid}[\mathbf{x}_\epsilon]\leq\epsilon$, $0\in[g]([\mathbf{x}_\epsilon])$ and $[\mathbf{x}_\epsilon]$ not included inside Ω_0^+ (so there exists furthermore $\mathbf{x}_\epsilon\in[\mathbf{x}_\epsilon]$ such that $\mathbf{x}_\epsilon\notin\Omega_0^+$). Considering $\epsilon=\frac{1}{k}$ we obtain sequences $([\mathbf{x}_k])_{k\in\mathbb{N}}$ and $(\mathbf{x}_k)_{k\in\mathbb{N}}$ satisfying these properties, so in particular $\operatorname{wid}[\mathbf{x}_k]$ converges to zero. The sequence $(\mathbf{x}_k)_{k\in\mathbb{N}}$ being bounded inside the compact set $[\mathbf{x}]$ it has a limit point $\mathbf{x}_*\in[\mathbf{x}]$ which cannot belong to $\operatorname{int}\Omega_0^+$ (since for all $k\in\mathbb{N}$ $\mathbf{x}_k\notin\Omega_0^+$). Therefore $g(\mathbf{x}_*)\neq 0$ and since g is continuous Lemma 1 can be applied showing that there exists $k\in\mathbb{N}$ such that $0\notin[g]([\mathbf{x}_k])$ which is a contradiction. □

3. Quadrature inclusion functions

An interval function $[Q_{f,g}]:\mathbb{IR}^n\to\mathbb{IR}$ is a quadrature inclusion function if and only if

$$[Q_{f,g}]([\mathbf{x}])\supseteq\int_{\{\mathbf{x}\in[\mathbf{x}]:g(\mathbf{x})\leq 0\}}f(\mathbf{x})\,d\mathbf{x}. \tag{11}$$

The symbol g can be dropped when the quadrature is performed inside a simple box. Since the exact quadrature is a real number, the overestimation of $[Q_{f,g}]([\mathbf{x}])$ is $\operatorname{wid}[Q_{f,g}]([\mathbf{x}])$. Based on the convergence properties of the most obvious quadrature inclusion function, namely $\int_{\mathbf{x}\in[\mathbf{x}]}f(\mathbf{x})\,d\mathbf{x}\subseteq(\operatorname{vol}[\mathbf{x}])\,[f]([\mathbf{x}])$ whose overestimation is $(\operatorname{vol}[\mathbf{x}])\operatorname{wid}[f]([\mathbf{x}])$, convergence related properties will involve the excess of the quadrature inclusion function, defined by:

$$\operatorname{exc}[Q_{f,g}]([\mathbf{x}]):=\frac{\operatorname{wid}[Q_{f,g}]([\mathbf{x}])}{\operatorname{vol}[\mathbf{x}]}. \tag{12}$$

Definition 1. Consider a domain $[\mathbf{x}^{\text{Init}}] \in \mathbb{IR}^n$.

- $[Q_{f,g}]$ is weakly convergent inside $[\mathbf{x}^{\text{Init}}]$ if and only if there exists $c \geq 0$ such that for all $[\mathbf{x}] \subseteq [\mathbf{x}^{\text{Init}}]$

$$\operatorname{exc}[Q_{f,g}]([\mathbf{x}]) \leq c. \tag{13}$$

- $[Q_{f,g}]$ is convergent inside $[\mathbf{x}^{\text{Init}}]$ if and only if for all sequences $([\mathbf{x}_k])_{k\in\mathbb{N}}$ such that $[\mathbf{x}_k] \subseteq [\mathbf{x}^{\text{Init}}]$ we have

$$\lim_{k\to\infty} \operatorname{wid}[\mathbf{x}_k] = 0 \Longrightarrow \lim_{k\to\infty} \operatorname{exc}[Q_{f,g}]([\mathbf{x}_k]) = 0. \tag{14}$$

- $[Q_{f,g}]$ has a convergence order $o \geq 1$ inside $[\mathbf{x}^{\text{Init}}]$ if and only if there exists c such that for all $[\mathbf{x}] \subseteq [\mathbf{x}^{\text{Init}}]$

$$\operatorname{exc}[Q_{f,g}]([\mathbf{x}]) \leq c\,(\operatorname{wid}[\mathbf{x}])^o. \tag{15}$$

Roughly speaking, the weak convergence states that $\operatorname{wid}[Q_{f,g}]([\mathbf{x}])$ behaves like $\operatorname{vol}[\mathbf{x}]$, while the convergence states that $\operatorname{wid}[Q_{f,g}]([\mathbf{x}])$ converges quicker to zero than $\operatorname{vol}[\mathbf{x}]$. It can be proved that under mild hypothesis convergence implies weak convergence, but this result will not be used here. On the other hand, having a convergence order $o \geq 1$ clearly implies convergence. The obvious quadrature inclusion function mentioned above is weakly convergent inside $[\mathbf{x}^{\text{Init}}]$ provided that $[f]([\mathbf{x}^{\text{Init}}])$ is bounded; it is convergent inside $[\mathbf{x}^{\text{Init}}]$ provided that $[f]$ is convergent inside $[\mathbf{x}^{\text{Init}}]$; it has a linear order of convergence inside $[\mathbf{x}^{\text{Init}}]$ provided that $[f]$ has a linear order of convergence inside $[\mathbf{x}^{\text{Init}}]$.

On the one hand, the weak convergence is not strong enough to enforce the branch and prune algorithm convergence (see Remark 2). On the other hand, although convergence is typically verified for quadrature over a box, the authors know no certified method achieving it for the quadrature over a constrained domain (see e.g. Appendix A.2). In practice, quadrature inclusion functions will use different techniques depending on the box being proved or not to be inside the constraint domain: When the box is proved to be inside, some efficient quadrature method dedicated to box domain can be used. Formally, such a quadrature inclusion function has the following form: Given an interval extension $[g]$ of the function g,

$$[Q_{f,g}^{[g]}]([\mathbf{x}]) := \begin{cases} [Q_{f,g}^1]([\mathbf{x}]) & \text{if } \underline{g} \leq 0 < \overline{g} \\ [Q_f^2]([\mathbf{x}]) & \text{if } \overline{g} \leq 0 \\ 0 & \text{otherwise,} \end{cases} \tag{16}$$

where $[\underline{g}, \overline{g}] = [g]([\mathbf{x}])$, $[Q_{f,g}^1]$ is a quadrature inclusion function dedicated to constrained domains, and $[Q_f^2]$ is a quadrature inclusion function with box domain. This class of quadrature inclusion functions based on an interval extension of the constraint will be called $[g]$-quadrature inclusion functions ($[g]$ will be implicitly supposed to be an interval extension of g). The following definition allows enforcing two different convergence requirements for $[g]$-quadrature inclusion functions.

Definition 2. A $[g]$-quadrature inclusion function $[Q_{f,g}^{[g]}]$ is $[g]$-convergent inside $[\mathbf{x}^{\text{Init}}]$ if and only if $[Q_{f,g}^1]([\mathbf{x}])$ is weakly convergent inside $[\mathbf{x}^{\text{Init}}]$ and $[Q_f^2]$ is convergent inside $[\mathbf{x}^{\text{Init}}]$.

This definition of $[g]$-convergence captures quite accurately the sufficient conditions for the branch and prune algorithm to be asymptotically convergent. Finally, the ϵ-maximum excess $\overline{r}_\epsilon$ of a quadrature inclusion function $[Q_{f,g}]$ inside $[\mathbf{x}^{\text{Init}}]$ is defined by:

$$\overline{r}_\epsilon := \sup\left\{\operatorname{exc}[Q_{f,g}]([\mathbf{x}]) : [\mathbf{x}] \subseteq [\mathbf{x}^{\text{Init}}], \operatorname{wid}[\mathbf{x}] \leq \epsilon\right\}. \tag{17}$$

The following proposition provides an equivalence to the convergence of quadrature inclusion functions.

Proposition 2. *$[Q_{f,g}]$ is convergent inside $[\mathbf{x}^{\text{Init}}]$ if and only if*

$$\lim_{\epsilon\to 0} \overline{r}_\epsilon = 0. \tag{18}$$

Proof. Condition (18) obviously implies convergence. The converse is proved by contradiction. Suppose that Condition (18) does not hold, that is, there exist $\delta > 0$ and $(\epsilon_k)_{k\in\mathbb{N}}$ such that $\lim_{k\to\infty} \epsilon_k = 0$ and $\overline{r}_{\epsilon_k} > \delta$. That is, there exists $([\mathbf{x}_k])_{k\in\mathbb{N}}$, $[\mathbf{x}_k] \subseteq [\mathbf{x}^{\text{Init}}]$, such that $\operatorname{wid}[\mathbf{x}_k] \leq \epsilon_k$, and thus $\lim_{k\to\infty} \operatorname{wid}[\mathbf{x}_k] = 0$, and $\operatorname{exc}[Q_{f,g}]([\mathbf{x}_k]) \geq \delta$, which contradicts the convergence of $[Q_{f,g}]$ inside $[\mathbf{x}^{\text{Init}}]$. □

4. The branch and prune algorithm

The branch and prune algorithm maintains a list $\mathcal{L}$ which contains pairs $([\mathbf{x}], [y])$ that satisfy

$$\int_{[\mathbf{x}]\cap\Omega} f(\mathbf{x})d\mathbf{x} \subseteq [y].$$

The function $\operatorname{Split}([\mathbf{x}])$ produces a finite set of boxes $\{[\mathbf{x}_1], \ldots, [\mathbf{x}_s]\}$ which satisfies $[\mathbf{x}] = \cup_i [\mathbf{x}_i]$ and $\sum_i \operatorname{vol}[\mathbf{x}_i] = \operatorname{vol}[\mathbf{x}]$. For each box obtained by splitting, the interval enclosure of the quadrature over this box is updated at Line 6. The updated

Please cite this article in press as: A. Goldsztejn, et al., Convergence analysis and adaptive strategy for the certified quadrature over a set defined by inequalities, Journal of Computational and Applied Mathematics (2013), http://dx.doi.org/10.1016/j.cam.2013.08.035

Algorithm 1: Branch and Prune Algorithm for Quadrature.

```
   Input: $f : \mathbb{R}^n \longrightarrow \mathbb{R}, g : \mathbb{R}^n \longrightarrow \mathbb{R}, [\mathbf{x}^{\mathrm{Init}}] \in \mathbb{IR}^n, \delta \geq 0$
   Output: $[y] \in \mathbb{IR}$
1  $\mathcal{L} \leftarrow \{([\mathbf{x}^{\mathrm{Init}}], [Q_{f,g}]([\mathbf{x}^{\mathrm{Init}}]))\}$;
2  while wid $\sum_{([\mathbf{x}],[y])\in\mathcal{L}}[y] \geq \delta$ do
3      $([\mathbf{x}], [y]) \leftarrow$ Extract($\mathcal{L}$);
4      $\mathcal{S} \leftarrow$ Split($[\mathbf{x}]$);
5      foreach $\mathbf{x} \in \mathcal{S}$ do
6          $[y] \leftarrow [Q_{f,g}]([\mathbf{x}])$;
7          if $[y] \neq 0$ then Insert$\big(\mathcal{L}, ([\mathbf{x}], [y])\big)$;
8      end
9  end
10 return $\sum_{([\mathbf{x}],[y])\in\mathcal{L}}[y]$;
```

pair $([\mathbf{x}], [y])$ is inserted again in $\mathcal{L}$ for further processing provided that its contribution is not null. In particular, if a $[g]$-quadrature inclusion function is used then the contribution is null whenever $\inf[g]([\mathbf{x}]) > 0$, thus boxes proved to be outside $\Omega = \{\mathbf{x} \in [\mathbf{x}^{\mathrm{Init}}] : g(\mathbf{x}) \leq 0\}$ are discarded. From now on, $[\mathbf{x}^{\mathrm{Init}}]$ is supposed to be bounded.

Remark 1. Since wid $([y_1] + [y_2]) =$ wid $[y_1] +$ wid $[y_2]$ the sum at Line 2 in Algorithm 1 does not need to be recomputed at each loop but is instead updated.

Theorem 1 below shows that the algorithm is correct.

Theorem 1. *Algorithm* 1 *is correct.*

Proof. We prove that Algorithm 1 maintains the list $\mathcal{L}$ in such a way that

$$\int_{\Omega} f(\mathbf{x})d\mathbf{x} \in \sum_{([\mathbf{x}],[y])\in\mathcal{L}} [y] \tag{19}$$

holds at each while iteration. This holds starting the first iteration thanks to the initialization at Line 1. Then this holds inductively for all iterations: During one iteration, a pair $([\mathbf{x}], [y])$ is replaced by a set of pairs $\{([\mathbf{x}_1], [y_1]), \ldots, ([\mathbf{x}_s], [y_s])\}$, $[\mathbf{x}_i] \in \mathcal{S}$, while

$$\int_{[\mathbf{x}]\cap\Omega} f(\mathbf{x})d\mathbf{x} = \int_{([\mathbf{x}_1]\cup\cdots\cup[\mathbf{x}_s])\cap\Omega} f(\mathbf{x})d\mathbf{x} = \sum_{k=1}^{k=s} \int_{[\mathbf{x}_k]\cap\Omega} f(\mathbf{x})d\mathbf{x} \tag{20}$$

$$\in \sum_{k=1}^{k=s} [y_i], \tag{21}$$

the first and second equalities holding respectively because $[\mathbf{x}] = \cup_i [\mathbf{x}_i]$ and $\sum_i \mathrm{vol}\,[\mathbf{x}_i] = \mathrm{vol}\,[\mathbf{x}]$ (hence the measure of the intersection is null). □

5. Convergence analysis

In this section, we study the asymptotical convergence of Algorithm 1: We suppose that the algorithm is called with a $[g]$-quadrature inclusion function (16), $\delta = 0$ and implemented with an infinite precision interval arithmetic in order to produce asymptotic convergence statements. In practice, the finiteness of the interval arithmetic precision will produce a limitation to the reachable precision of the final quadrature (hence the requirement of some additional safe guard stopping criteria to enforce halting). We present two results: First we prove that Algorithm 1 converges asymptotically in the case of fair splitting strategies and of the worst first strategy (Corollaries 1 and 2 in Section 5.2). Second, in the case where the quadrature is performed over a box, we provide some additional result about the rate of convergence of Algorithm 1 (Theorem 3 in Section 5.3). The rate of convergence for domains Ω defined by inequalities will be finally informally discussed in Section 5.4.

5.1. Preliminaries

For the investigations that follow, let us split the list $\mathcal{L}$ into two sublists $\mathcal{B}$ and $\mathcal{I}$, such that $\mathcal{L} = \mathcal{B} \cup \mathcal{I}$, and $([\mathbf{x}], [y]) \in \mathcal{I}$ if and only if $\sup[g]([\mathbf{x}]) \leq 0$. Therefore pairs from $\mathcal{B}$ and $\mathcal{I}$ are processed respectively by the first and the second quadrature

Please cite this article in press as: A. Goldsztejn, et al., Convergence analysis and adaptive strategy for the certified quadrature over a set defined by inequalities, Journal of Computational and Applied Mathematics (2013), http://dx.doi.org/10.1016/j.cam.2013.08.035

inclusion function of the $[g]$-quadrature inclusion function. Since boxes satisfying $\inf[g]([\mathbf{x}]) > 0$ are simply rejected we have:

$$[\mathbf{x}] \in \mathcal{B} \Longrightarrow 0 \in [g]([\mathbf{x}]). \tag{22}$$

Boxes from $\mathcal{I}$ (resp. $\mathcal{B}$) are called inner (resp. boundary) boxes. We define $\mathcal{L}_k$, $\mathcal{I}_k$ and $\mathcal{B}_k$ the lists built by Algorithm 1 at the k^{th} execution of the while loop. For $([\mathbf{x}_k], [y_k]) \in \mathcal{L}_k$ the height of $[\mathbf{x}_k]$ is the number of splits it went through. It is denoted by $\text{height}([\mathbf{x}_k])$. Also we denote $\sum_{([\mathbf{x}],[y])\in\mathcal{L}_k}[y]$, the output that would be produced at step k, by $[y_k^{\text{Tot}}]$. Finally, we define the maximum width of boxes of $\mathcal{L}_k$ by

$$\epsilon_k := \max_{([\mathbf{x}],[y])\in\mathcal{L}_k} \text{wid}\,[\mathbf{x}]. \tag{23}$$

Box splitting and selection strategies. The strategy for selecting the box to be processed at Line 3 is of central importance for the convergence and the efficiency of Algorithm 1. Some common search strategies are Deep-First (DFS), Breadth-First (BFS), FIFO, etc. The strategy that consists of extracting the box of Largest width First (LFS) is also widely used in branch and prune algorithms. In order to study the convergence of the algorithm, we introduce fair selection strategies: A selection strategy is fair if and only if all sequences $\big(([\mathbf{x}_k], [y_k])\big)_{k\in\mathbb{N}}$, with $([\mathbf{x}_k], [y_k]) \in \mathcal{L}_k$, satisfy $\text{height}([\mathbf{x}_k]) \to \infty$. It is easily seen that the BFS, the FIFO strategy, and the LFS are fair, while the DFS is not fair.

In the context of global optimization, it was noticed that a dedicated box selection strategy can drastically improve the efficiency of the algorithm. Similarly, we define the worst-first strategy (WFS) for box selection, which extracts from $\mathcal{L}$ the pair $([\mathbf{x}], [y])$ which has the maximum $\text{wid}\,[y]$. In this way, the box that has the largest contribution to the overall quadrature uncertainty will be processed first. Note that the WFS is not fair in general: Indeed, if a pair $([\mathbf{x}], [y])$ satisfies $\text{wid}\,[\mathbf{x}] > 0$ and $\text{wid}\,[y] = 0$ (e.g. f has a polynomial expression inside $[\mathbf{x}]$) then the WFS will not consider it anymore, hence not satisfying the fair strategy definition. Nevertheless, convergence will be proved for the WFS too.

Finally, we define a box splitting strategy to be fair if and only if

$$\lim_{k\to\infty} \text{height}([\mathbf{x}_k]) = \infty \Longrightarrow \lim_{k\to\infty} \text{wid}\,[\mathbf{x}_k] = 0. \tag{24}$$

The standard strategy that consists of bisecting a box at the midpoint of its largest component is easily seen to be fair. This splitting strategy will be used in the experiments of Section 6.

5.2. Convergence

The convergence of Algorithm 1 is now investigated. As noted before, the WFS box selection strategy is not fair in general. So, Theorem 2 first shows that, under some appropriate hypotheses, Algorithm 1 is asymptotically convergent. Then, two corollaries are derived for the cases of fair box selection strategies and the WFS. This also proves that under those hypotheses Algorithm 1 always stops when called with $\delta > 0$.

In the following we define $\mathcal{L}'_k = \{([\mathbf{x}], [y]) \in \mathcal{L}_k : \text{wid}\,[y] > 0\}$, $\mathcal{I}'_k = \mathcal{I}_k \cap \mathcal{L}'_k$, $\mathcal{B}'_k = \mathcal{B}_k \cap \mathcal{L}'_k$, so that prime symbols concern only pairs that have a positive contribution to the overall enclosure width, and

$$\epsilon'_k := \max_{([\mathbf{x}],[y])\in\mathcal{L}'_k} \text{wid}\,[\mathbf{x}]. \tag{25}$$

Note that $\epsilon'_k \le \epsilon_k$ obviously holds. Also, $\{\mathbf{x} \in \Omega : g(\mathbf{x}) = 0\}$ will be denoted by Ω_0. The next theorem will have two corollaries that actually show the asymptotic convergence of Algorithm 1 for fair and worst first selection strategies.

Theorem 2. *Suppose that* $\text{vol}\,(\Omega_0) = 0$, g *is continuous in* $[\mathbf{x}^{\text{Init}}]$, $[g]$ *is convergent inside* $[\mathbf{x}^{\text{Init}}]$, *and the* $[g]$*-quadrature inclusion function is* $[g]$*-convergent inside* $[\mathbf{x}^{\text{Init}}]$. *Then Algorithm 1 satisfies:*

$$\lim_{k\to\infty} \epsilon'_k = 0 \Longrightarrow \lim_{k\to\infty} \text{wid}\,[y_k^{\text{Tot}}] = 0. \tag{26}$$

Proof. We have

$$[y_k^{\text{Tot}}] = \sum_{([\mathbf{x}],[y])\in\mathcal{I}_k} [y] + \sum_{([\mathbf{x}],[y])\in\mathcal{B}_k} [y] \tag{27}$$

and therefore

$$\text{wid}\,[y_k^{\text{Tot}}] = \sum_{([\mathbf{x}],[y])\in\mathcal{I}_k} \text{wid}\,[y] + \sum_{([\mathbf{x}],[y])\in\mathcal{B}_k} \text{wid}\,[y]. \tag{28}$$

$$= \sum_{([\mathbf{x}],[y])\in\mathcal{I}'_k} \text{wid}\,[y] + \sum_{([\mathbf{x}],[y])\in\mathcal{B}'_k} \text{wid}\,[y]. \tag{29}$$

We now prove that the last two summations both converge to zero.

The quadrature enclosures in the first summation use $[Q_f^2]([\mathbf{x}])$, which is convergent inside $[\mathbf{x}^{\text{Init}}]$ and whose ϵ-maximum excess is denoted by $\overline{r}_\epsilon$. Therefore,

$$\begin{aligned}\sum_{([\mathbf{x}],[y])\in\mathcal{I}'_k} \text{wid}\,[y] &= \sum_{([\mathbf{x}],[y])\in\mathcal{I}'_k} \text{exc}\,[Q_{f,g}]([\mathbf{x}])\,\text{vol}\,[\mathbf{x}]\\ &\le \overline{r}_{\epsilon'_k} \sum_{([\mathbf{x}],[y])\in\mathcal{I}'_k} \text{vol}\,[\mathbf{x}] \le \overline{r}_{\epsilon'_k}\text{vol}\,[\mathbf{x}^{\text{Init}}]. \qquad (30)\end{aligned}$$

Since $[Q_f^2]([\mathbf{x}])$ is convergent inside $[\mathbf{x}^{\text{Init}}]$ and ϵ'_k converges toward zero, Proposition 2 proves that $\overline{r}_{\epsilon'_k}$ also converge toward zero. Therefore so does this summation.

The quadrature enclosures in the second summation are computed using $[Q_f^1]([\mathbf{x}])$, which is weakly convergent inside $[\mathbf{x}^{\text{Init}}]$. Therefore there exists $c \ge 0$ such that

$$\begin{aligned}\sum_{([\mathbf{x}],[y])\in\mathcal{B}'_k} \text{wid}\,[y] &= \sum_{([\mathbf{x}],[y])\in\mathcal{B}'_k} \text{exc}\,[Q_{f,g}]([\mathbf{x}])\,\text{vol}\,[\mathbf{x}]\\ &\le c \sum_{([\mathbf{x}],[y])\in\mathcal{B}'_k} \text{vol}\,[\mathbf{x}] \le c\,\text{vol}\,(\cup\mathcal{B}'_k). \qquad (31)\end{aligned}$$

Therefore it just remains to prove that $\text{vol}\,(\cup\mathcal{B}'_k)$ converges to zero, which is done using Proposition 1. By definition of a null-volume set, for all $\delta > 0$ there exists a set of boxes $\mathcal{C}$ such that $\text{vol}\,(\cup\mathcal{C}) \le \frac{1}{2^n}\delta$ and $\cup\mathcal{C} \supseteq \Omega_0$. Now define

$$\mathcal{C}^+ = \{\text{mid}\,[\mathbf{x}'] + 2([\mathbf{x}'] - \text{mid}\,[\mathbf{x}']) : \mathbf{x}' \in \mathcal{C}\}, \qquad (32)$$

informally each box of $\mathcal{C}$ sees each of its dimension inflated of a ratio 2. Thus we now have $\text{vol}\,(\cup\mathcal{C}^+) \le \delta$ and $\Omega_0 \subseteq \text{int}\,(\cup\mathcal{C}^+)$. By Proposition 1 there exists ϵ''_δ such that $[\mathbf{x}] \subseteq [\mathbf{x}^{\text{Init}}]$, $\text{wid}\,[\mathbf{x}] \le \epsilon''_\delta$ and $0 \in [g]([\mathbf{x}])$ implies $[\mathbf{x}] \subseteq \cup\mathcal{C}^+$. Since the boxes $[\mathbf{x}] \in \mathcal{B}'_k$ satisfy $[\mathbf{x}] \subseteq [\mathbf{x}^{\text{Init}}]$, $\text{wid}\,[\mathbf{x}] \le \epsilon'_k$ and $0 \in [g]([\mathbf{x}])$ we obtain that $\epsilon'_k \le \epsilon''_\delta$ implies $(\cup\mathcal{B}'_k) \subseteq (\cup\mathcal{C}^+)$ and thus $\text{vol}\,(\cup\mathcal{B}'_k) \le \delta$. As this holds for an arbitrary $\delta > 0$ and ϵ'_k converges to zero, we have $\text{vol}\,(\cup\mathcal{B}'_k)$ also converges to zero. □

Corollary 1. *Under the hypotheses of Theorem 2, suppose furthermore that both the splitting and the box selection strategies are fair, then Algorithm* 1 *is asymptotically convergent.*

Proof. As the selection strategy is supposed fair, all sequences $\big(([\mathbf{x}_k], [y_k])\big)_{k\in\mathbb{N}}$, $([\mathbf{x}_k], [y_k]) \in \mathcal{L}_k$, satisfy $\text{height}([\mathbf{x}_k]) \to \infty$. So they also satisfy $\text{wid}\,([\mathbf{x}_k]) \to 0$ because the splitting strategy is fair. As a consequence ϵ_k converges to zero and so does ϵ'_k. Theorem 2 finally proves that Algorithm 1 is asymptotically convergent. □

Lemma 2. *For an arbitrary $\delta > 0$, let $\mathcal{S}_\delta$ be the set of all boxes of width greater than δ that can be produced by Algorithm* 1 *using a fair box splitting strategy. Then $\mathcal{S}_\delta$ is finite.*

Proof. Suppose that $\mathcal{S}_\delta$ is infinite. Since each new splitting produces finitely many boxes and increases the height of each newly produced boxes of 1, the height cannot be bounded inside $\mathcal{S}_\delta$. So there exists $([\mathbf{x}_k])_{k\in\mathbb{N}}$, $[\mathbf{x}_k] \in \mathcal{S}_\delta$, such that $\text{height}([\mathbf{x}_k])$ converges to infinity. However, since the splitting strategy is fair, this implies that $[\mathbf{x}_k]$ converges to zero and contradicts the lower bound δ for the width of the boxes of $\mathcal{S}_\delta$. □

Corollary 2. *Under the hypotheses of Theorem 2, suppose furthermore that the splitting strategy is fair and the WFS is used, then Algorithm* 1 *is asymptotically convergent.*

Proof. Note that there is at least one sequence $\big(([\mathbf{x}_k], [y_k])\big)_{k\in\mathbb{N}}$, $([\mathbf{x}_k], [y_k]) \in \mathcal{L}_k$, such that $\text{height}([\mathbf{x}_k])$ is unbounded, and since the splitting strategy is fair we have $\text{wid}\,[\mathbf{x}_k]$ converges to zero. Now since the $[g]$-quadrature inclusion function is $[g]$-convergent, $\text{wid}\,[y_k]$ also converges to zero. Finally, since the element of $\mathcal{L}$ with largest $[y]$ is selected at each step, we have all sequences $(\text{wid}\,[y_k])_{k\in\mathbb{N}}$, $([\mathbf{x}_k], [y_k]) \in \mathcal{L}_k$, do actually converge to zero.

Now, let us suppose that ϵ'_k is not converging to zero and show a contradiction. So there exist $\delta > 0$ and a sequence $\big(([\mathbf{x}_k], [y_k])\big)_{k\in\mathbb{N}}$, $([\mathbf{x}_k], [y_k]) \in \mathcal{L}'_k$, such that $\text{wid}\,[\mathbf{x}_k] \ge \delta$. Note that by definition of $\mathcal{L}'_k$ we have $\text{wid}\,[y_k] \ne 0$ for all $k \in \mathbb{N}$. However, by Lemma 2 there are only finitely many different boxes of widths larger than δ so we can pickup a constant subsequence $\big(([\mathbf{x}_{\pi(k)}], [y_{\pi(k)}])\big)_{k\in\mathbb{N}}$ from $\big(([\mathbf{x}_k], [y_k])\big)_{k\in\mathbb{N}}$. But since we have previously shown that $(\text{wid}\,[y_k])_{k\in\mathbb{N}}$ converges to zero, $(\text{wid}\,[y_{\pi(k)}])_{k\in\mathbb{N}}$ also converges to zero. Therefore since it is constant we have $\text{wid}\,[y_{\pi(k)}] = 0$, which contradicts $\text{wid}\,[y_k] \ne 0$ for all $k \in \mathbb{N}$. Finally we have proved that ϵ'_k converges to zero, and therefore Theorem 2 proves that Algorithm 1 is asymptotically convergent. □

Remark 2. Note that the weak convergence of the quadrature inclusion function is not enough to enforce the asymptotic convergence of the branch and prune algorithm. E.g. using the weakly convergent quadrature function $[Q_{f,g}]([\mathbf{x}]) = [f]([\mathbf{x}^{\text{Init}}])\,\text{vol}\,[\mathbf{x}]$, the algorithm will output $[f]([\mathbf{x}^{\text{Init}}])\,\text{vol}\,[\mathbf{x}^{\text{Init}}]$.

5.3. Convergence rate for box domains

In this section, we suppose that quadrature inclusion function has an order of convergence $o \geq 1$. Up to the authors' knowledge, state of the art techniques for certified quadrature can achieve this only when the quadrature is performed over a box domain (see e.g. Appendix). The convergence rate is first computed with respect to ϵ_k, i.e. the maximal width of the boxes in $\mathcal{L}_k$.

Theorem 3. *Suppose that the quadrature inclusion function has an order of convergence $o \geq 1$. Then there exists $c > 0$ such that*

$$\text{wid}\,[y_k^{\text{Tot}}] \leq c\,(\epsilon_k')^o. \tag{33}$$

Proof. Similarly to Eqs. (29) and (30) we have:

$$\text{wid}\,[y_k^{\text{Tot}}] = \sum_{([\mathbf{x}],[y])\in\mathcal{L}_k'} \text{wid}\,[y] \leq \bar{r}_{\epsilon_k'} \sum_{([\mathbf{x}],[y])\in\mathcal{L}_k'} \text{vol}\,[\mathbf{x}] \leq \bar{r}_{\epsilon_k'} \text{vol}\,[\mathbf{x}^{\text{Init}}]. \tag{34}$$

Finally, since the quadrature inclusion function has an order of convergence o, there exist $c' > 0$ such that $\bar{r}_{\epsilon_k'} \leq c'\,(\epsilon_k')^o$ and we obtain $\text{wid}\,[y_k^{\text{Tot}}] \leq c\,(\epsilon_k')^o$ with $c = c'\,\text{vol}\,[\mathbf{x}^{\text{Init}}]$. □

Since $\epsilon_k' \leq \epsilon_k$, Theorem 3 trivially entails

$$\text{wid}\,[y_k^{\text{Tot}}] \leq c\,(\epsilon_k)^o. \tag{35}$$

Finally, strategies that homogeneously bisect a domain within a n dimensional space satisfy $k \sim (\epsilon_k)^{-n}$ (e.g. this obvious holds for the strategy that consists in selecting the larger box and splitting its largest component at the midpoint, for which dividing ϵ_k by 2 requires subdividing all boxes into 2^n, hence multiplying k by 2^n). Then Theorem 3 gives rise to

$$\text{wid}\,[y_k^{\text{Tot}}] \sim c\,k^{-\frac{o}{n}}. \tag{36}$$

As the most important part of the computations are the computation of the Taylor models, the execution time t should also be proportional and we should have $\text{wid}\,[y_k^{\text{Tot}}] \sim c\,t^{-\frac{o}{n}}$, which is quite accurately verified in Section 6.

5.4. Convergence rate for domains defined by inequalities

In this section, we provide an informal study of the convergence rate of Algorithm 1 in the case of a domain defined by inequalities. When the domain of integration is defined by inequalities, both inner and boundary boxes contribute to the overall width of the overall quadrature enclosure. However, quadrature enclosures provided for boundary boxes are much less accurate than those provided for inner boxes. So, if the order of convergence of the quadrature inclusion function used to compute the quadrature of inner boxes is high enough, we can neglect the contribution of these boxes to the overall quadrature and focus only on the contribution of boundary boxes. We can also suppose that asymptotically the boundary boxes are homogeneously balanced with size ϵ, so the volume of such a box is ϵ^n.

Now, provided that the order of convergence of the interval extension of g is high enough,[2] the total volume of the boundary boxes is asymptotically proportional to ϵ (it is proportional to the surface of the boundary of Ω times epsilon). And provided that the quadrature inclusion function is weakly convergent inside $[\mathbf{x}^{\text{Init}}]$, we obtain

$$\text{wid}\,[y_k^{\text{Tot}}] \sim \epsilon. \tag{37}$$

Finally, if the box selection strategy is the LFS then the boxes accumulate on the whole Ω, whose measure in $\mathbb{R}^n$ is denoted by $M_n(\Omega)$. The algorithm produces k boxes after k bisections, which we suppose for this asymptotic analysis to have a measure of ϵ^n. Therefore, asymptotically $k = \frac{M_n(\Omega)}{\epsilon^n}$ such boxes are required to cover Ω, leading to the relation $\epsilon^{-n} \sim k$. In this case we obtain

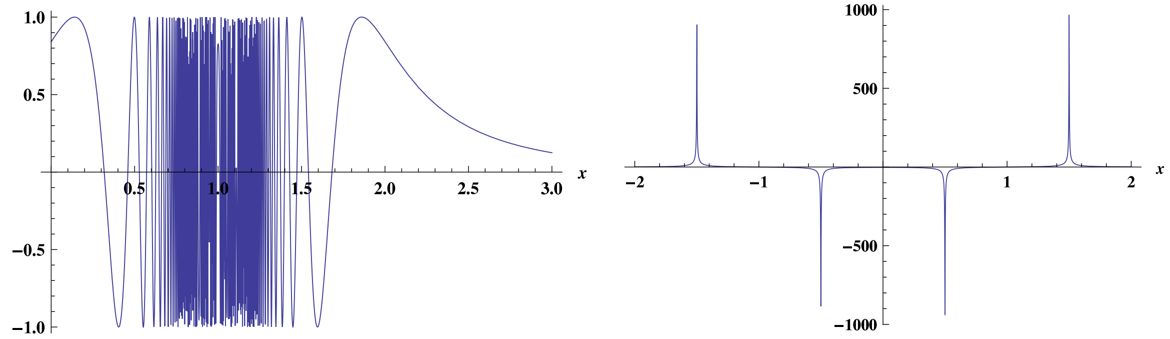
$$\text{wid}\,[y_k^{\text{Tot}}] \sim k^{-\frac{1}{n}}. \tag{38}$$

On the other hand, if the box selection strategy is the WFS then the boxes accumulate on the boundary of Ω, whose measure in $\mathbb{R}^{n-1}$ is denoted by $M_{n-1}(\partial\Omega)$. As mentioned previously, these boxes cover a set of measure proportional to $\epsilon M_{n-1}(\partial\Omega)$, and therefore $\epsilon\,\epsilon^{-n} = \epsilon^{-(n-1)} \sim k$. In this case we obtain

$$\text{wid}\,[y_k^{\text{Tot}}] \sim k^{-\frac{1}{n-1}}. \tag{39}$$

As in the case of unconstrained integration, the execution time t should also be proportional to the iteration steps k, resulting in the approximations $\text{wid}\,[y_k^{\text{Tot}}] \sim t^{-\frac{1}{n}}$ for the LFS and $\text{wid}\,[y_k^{\text{Tot}}] \sim t^{-\frac{1}{n-1}}$ for the WFS, which are in accordance with the experiments of Section 6.

[2] E.g. a linear order of convergence is enough provided that $g(\mathbf{x}) = 0 \Longrightarrow \nabla g(\mathbf{x}) \neq 0$.



(a) $f_1(x)$.

(b) $f_2(x)$.

Fig. 1. Graphs of the integrand functions $f_1(x)$ and $f_2(x)$ within the integration bounds [0, 3] and [−2, 2] respectively.

6. Experiments

The following experiments were carried out on an Intel Core 2 Duo at 2.53 GHz. Algorithm 1 was implemented in C++ using the interval library PROFIL/BIAS [16] and the C++ multimap container for keeping the sub-boxes and the sub-integrals ordered accordingly to the selection strategy. In the following Algorithm 1 will be designated by Q_s^o where o is a number specifying the order of expansion of the Taylor model and s is a letter denoting the adopted box selection strategy (f → largest width first strategy[3]; a → worst first strategy).

The computation of Taylor models of arbitrary expansion orders for generic multidimensional integrand functions required the evaluation of higher order derivative tensors performed in interval arithmetic. There are a few available tools for computing higher order partial derivatives in C++, such as Adol-C [17] and Rapsodia [18], but they are based on floating-point arithmetic and its extension to interval arithmetic is not trivial.

We have adopted a implementation of the recursive calculation in the forward mode of the chain-rule based technique known as automatic differentiation [19]. The idea is to pre-compile the integrand expression into a program code that, from a domains box, computes a large array with all the derivative enclosures required by the Taylor model. More sophisticated alternative approaches include: differential algebraic techniques [20]; the computation of hyperpyramid structures with the generalized Leibniz's rule [21]; and the interpolation of multidimensional Taylor coefficients from the propagation of univariate Taylor series along fixed directions [22,23].

Remark 3. Timings of the proposed methods are compared to others from the literature. They are not meant assessing which method is the best, but just demonstrating that the proposed method is comparable to the state of the art in certified quadrature.

6.1. One-dimensional integrals

Consider the following one dimensional definite integrals extracted from [6]:

$$I_1 = \int_0^3 f_1(x)dx \quad \text{and} \quad I_2 = \int_{-2}^{2} f_2(x)dx$$

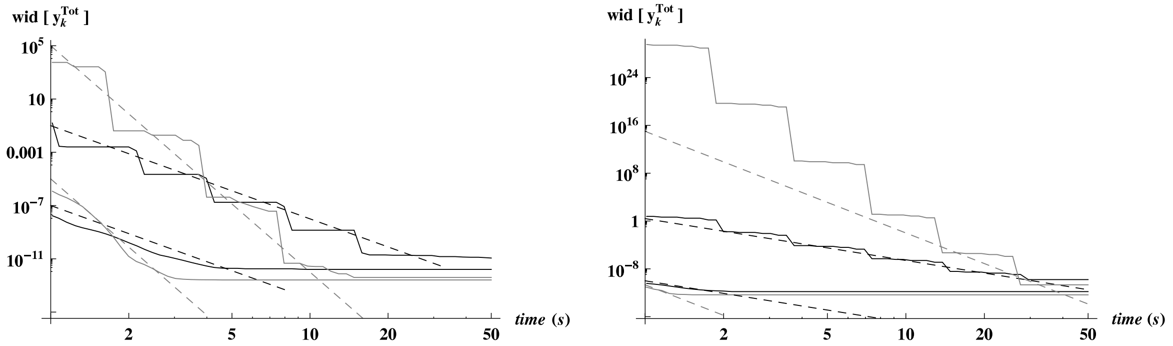
$$\text{with} : f_1(x) = \sin\left(\frac{1}{10^{-3} + ((1-x)^2)^{1.5}}\right)$$

$$f_2(x) = +\frac{1}{\sqrt{(10^{-6} + (x+1.5)^2)}} - \frac{1}{\sqrt{(10^{-6} + (x+0.5)^2)}} - \frac{1}{\sqrt{(10^{-6} + (x-0.5)^2)}} + \frac{1}{\sqrt{(10^{-6} + (x-1.5)^2)}}$$

Fig. 1 plots the graph of the integrand functions within the integration bounds. Clearly, f_1 is highly oscillatory whereas f_2 presents a number of distinct peaks that stand out from the smooth graph. Such difficulties are apparent in [6] where the interval enclosures obtained required much more computation time than the other one-dimensional experiments presented in [6].

Algorithm 1 was tested with different expansion orders and different box selection strategies for computing enclosures of I_1 and I_2. Fig. 2 shows how the width of the interval enclosures of the integral (wid $[y_k^{Tot}]$) decreases with respect to CPU time. Q^5 and Q^{15} are represented respectively by black and gray lines. Fair box selection strategies produced stair shaped lines whereas adaptive strategies evolved as smooth lines. All the tests are plotted in a logarithmic scale and so, accordingly

[3] LFS is a fair box selection strategy which is not adaptive.



(a) I_1.

(b) I_2.

Fig. 2. Time plots for the width of the enclosures of I_1 and I_2 obtained with Q_f^5 (black stair shape lines), Q_a^5 (black smooth lines), Q_f^{15} (gray stair shape lines) and Q_a^{15} (gray smooth lines). The dashed strait lines represent the respective theoretically predicted convergence slopes.

Table 1
Enclosures of I_1 and I_2 obtained by Algorithm 1 and in [6].

	I_1		I_2	
Q_f^5	$0.7578918118\binom{559}{443}$	47.9 s	$-1.524278365\binom{6}{4}$	31.6 s
Q_a^5	$0.7578918118\binom{511}{490}$	4.4 s	$-1.52427836550\binom{4}{1}$	1.9 s
Q_f^{15}	$0.7578918118\binom{534}{464}$	265.6 s	$-1.524278365\binom{6}{4}$	27.2 s
Q_a^{15}	$0.7578918118\binom{503}{499}$	2.9 s	$-1.524278365502\binom{9}{4}$	1.3 s
[6]	$0.7578918118\binom{503}{499}$	13.2 s	$-1.52427836550\binom{4}{1}$	35.5 s

to the theoretical convergence rate results of Section 5.3, they should asymptotically converge to strait lines. Each dashed strait line represents a predicted convergence slope which is conveniently superimposed to the experimental results for the respective order of the Taylor model expansion.

All variants of Algorithm 1 succeeded in computing tight enclosures for the integrals and the most competitive computation times were achieved by adaptive strategies. The results obtained by Algorithm 1 are summarized in Table 1. The last line of the table presents the results obtained in [6].

Notice that due to differences in hardware and software (e.g. the interval arithmetic library), execution time comparisons are not meaningful to support the superiority of any of these approaches.

Now consider the definite integrals I_3 and I_4 extracted respectively from [7,8]:

$$I_3 = \int_0^1 f_3(x)dx \quad \text{and} \quad I_4 = \int_{-1}^1 f_4(x)dx$$

$$\text{with} : f_3(x) = \sqrt{0.01 + x + x^2}(\cos x + \sin x)$$

$$f_4(x) = \sin(10x) + \sum_{i=1}^{15} g_i(x) \quad \text{where } g_i(x) = \begin{cases} 1 - 2\sin(10x) & \text{if } i = 1 \\ 1 - \dfrac{81}{128} g_{i-1}^4(x) & \text{if } i > 1. \end{cases}$$

I_3 illustrates an integral with an nearly endpoint singularity which affects the convergence rate of Gaussian quadrature methods [7]. I_4 is a definite integral over a complicated function f_4 which is very ill behaved outside the integration bounds [8]. Fig. 3 plots the graph of f_4 within the integration bounds $[-1, 1]$.

Again all variants of Algorithm 1 succeeded in computing tight enclosures for the integrals and the adaptive strategies were clearly superior to fair strategies. The enclosures obtained by Algorithm 1 together with the respective execution times are summarized in Table 2. The last line of the table shows the output of Mathematica computed numerically [10] $_n$ (with the standard options of *NIntegrate*, i.e. global adaptive strategy with cartesian product rule).

In both problems, I_3 and I_4, Mathematica [10] failed to compute the integral symbolically and the approximations obtained by the *NIntegrate* method were 5 orders of magnitude less accurate than the enclosures provided by Q_a^{15}.

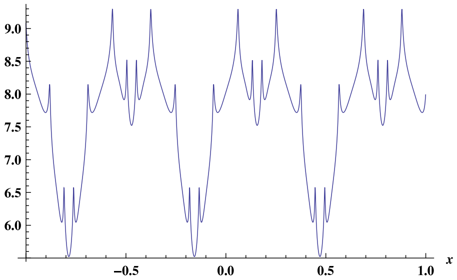


Fig. 3. Graph of function $f_4(x)$ within the integration bounds $[-1, 1]$.

Table 2
Enclosures of I_3 and I_4 obtained by Algorithm 1 and their numerical computations with Mathematica [10] $_n$ (the correct digits are underlined).

	I_3		I_4	
Q_f^5	$1.144540250039\binom{3}{1}$	0.157 s	$15.319813554\binom{7}{5}$	15.74 s
Q_a^5	$1.1445402500391\binom{9}{4}$	0.062 s	$15.3198135546\binom{3}{1}$	8.40 s
Q_f^{15}	$1.1445402500391\binom{8}{5}$	0.250 s	$15.3198135546\binom{3}{1}$	66.59 s
Q_a^{15}	$1.1445402500391\binom{8}{6}$	0.016 s	$15.31981355461\binom{9}{6}$	11.64 s
$[10]_n$	$\underline{1.1445402}49506109$	0.016 s	$\underline{15.319813}13779004$	0.11 s

From the different tests performed on integrals I_1, I_2, I_3 and I_4 several general conclusions can be drawn which are corroborated by the experiments performed in the next section for multidimensional integrals.

The stair shape pattern of fair strategies versus the smooth line of adaptive strategies can be verified in all experiments. In fair strategies, the imprecision of the larger boxes dominates the overall enclosure computations and so, significant improvements are only achieved when all boxes become the same size. On the other hand, in adaptive strategies, the next selected box is always the one that contributes the most for the overall imprecision, which can be significantly improved with a tighter enclosure of such contribution.

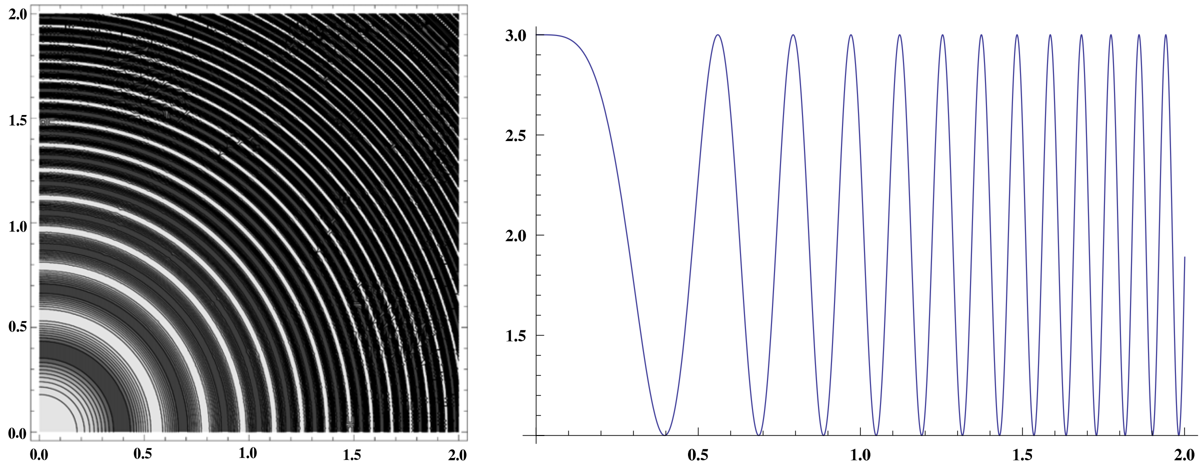
With the same order of the Taylor model expansion, the adaptive strategy always outperforms the fair strategy. Adaptive strategies induce an uneven partitioning of the domains where non-smooth regions are intensively subdivided. To achieve the same accuracy, strategies that induce a fair partitioning of the domains, require much more subdivisions. For example, to achieve an enclosure $I_1 \in 0.7578918118\binom{7}{3}$, Q_f^5 partitioned the x domain into about 175 thousand subintervals whereas Q_a^5 required only about 14 thousand subintervals. Such proportion of about 12.5 times more boxes is reflected also in the computation times required to obtain the enclosures, 28.7 and 2.3 s, respectively.

Because with the same Taylor model order, fair and adaptive strategies, approach asymptotically to the same convergence rate (same slope in the graphics), the ratio between number of boxes (or execution times) tend to be kept constant and proportional to the distance between the respective parallel lines.

The experimental convergence rates observed at different Taylor model expansion orders are pretty in accordance with the theoretical results for both, fair and adaptive strategies. However, due to interval arithmetic accuracy limits [16], errors cannot be decreased behind those limits and, when the domain is intensively partitioned into small boxes, their cumulative effect decelerate the convergence rate and limits the overall accuracy attainable. This limit appears on the graphics when the curves become flat.

Note that limits on the attainable overall accuracy results from the accumulation of unavoidable small errors on each domain subdivision. Therefore, it is expected that methods requiring less partitioning may provide better overall accuracy. This is observed in all experiments where, in the limit, higher order expansions reach better accuracy than lower order expansions and adaptive strategies reach better accuracy than fair strategies.

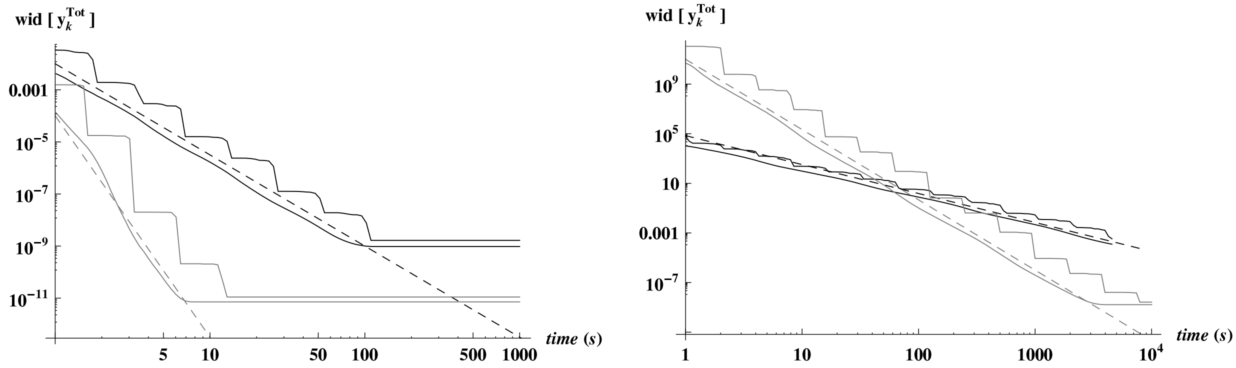
Experiments show also that, in some cases, lower order Taylor model expansions may be more efficient for larger boxes than higher order expansions. This is because the computation of higher order tensors, which is more time consuming, may provide larger enclosures, when computed on large boxes, due to extensive interval arithmetic operations on large interval domains. In cases such as I_1 (and D_3 in the next section), despite of the better convergence rate of higher order expansions, better overall enclosures can only be obtained when boxes get small enough. Graphically, the better slope of higher order expansions is balanced by their worse starting points (larger initial enclosures).



(a) $f(\mathbf{x})$ contour.

(b) $f(\mathbf{x})$ slice.

Fig. 4. (a) Contour plot of the integrand function $f(\mathbf{x})$ for $n = 2$ within the integration bounds $[0, 2]^2$ and (b) a slice of the function crossing **0**.



(a) I_5.

(b) I_6.

Fig. 5. Time plots for the width of the enclosures of I_5 and I_6 obtained with Q_f^5 (black stair shape lines), Q_a^5 (black smooth lines), Q_f^{15} (gray stair shape lines) and Q_a^{15} (gray smooth lines). The dashed strait lines represent the respective theoretically predicted convergence slopes.

6.2. Multidimensional integrals

Consider the two- and three-dimensional definite integrals $I_5 = B_2$ and $I_6 = B_3$ where:

$$B_n = \int_{\mathbf{x}\in[0,2]^n} f(\mathbf{x})d\mathbf{x} \quad \text{with } f(\mathbf{x}) = 2 + \cos(20\|\mathbf{x}\|_2^2)$$

Function f is quite oscillatory as illustrated in Fig. 4.

Fig. 5 shows tests performed with Q_f^5, Q_a^5, Q_f^{15} and Q_a^{15} whose results are summarized in Table 3. The last lines of the table shows the outputs of Mathematica computed numerically $[10]_n$ (with the standard options of NIntegrate) and formally $[10]_f$ (the result is first computed symbolically and then rounded to the nearest so all digits are correct). The numerical procedure detects some *highly oscillatory integrand* in both cases.

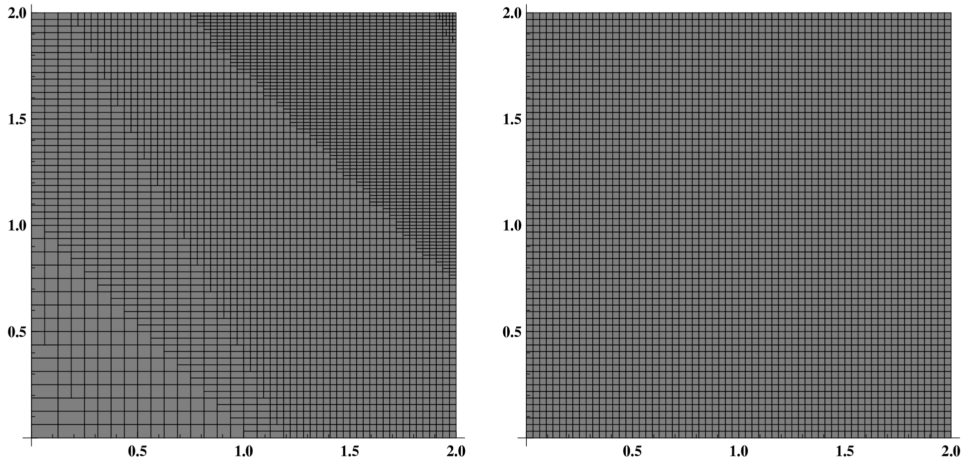
Again Algorithm 1 succeeded in computing tight enclosures for the integrals. For example, Q_a^{15} computed interval enclosures $I_5 \in 7.9962646819\binom{6}{4}$ and $I_6 \in 15.99440289\binom{3}{0}$. However, the execution time was much faster for I_5 (about 5.6 s) than for I_6 (about 48 min). Moreover, whereas in I_5 tight enclosures could also be attained with lower order expansions, in the case of I_6, only high orders could obtain such results.

The general conclusions drawn in the previous section for the one-dimensional cases can also be applied to the multidimensional cases shown in Fig. 5. The convergence rates observed experimentally are in accordance with the theoretical results and the adaptive strategies always outperforms the fair strategies (with the same order of expansion).

The difference between adaptive and fair strategies with respect to the domain partitioning is illustrated in Fig. 6 for the two dimensional case. With a domain partitioning into 4096 boxes obtained after about 3 s of CPU time, in the adaptive strategy, the smaller boxes are further away from the origin because, as can be seen in Fig. 4, the oscillatory effect in f increases with the distance from the origin.

Table 3
Enclosures of I_5 and I_6 obtained by Algorithm 1 and their numerical ([10] $_n$) and formal ([10] $_f$) computations with Mathematica.

	I_5		I_6	
Q_f^5	$7.99626468\begin{pmatrix}3\\1\end{pmatrix}$	114.7 s	$15.994\begin{pmatrix}6\\2\end{pmatrix}$	4165.9 s
Q_a^5	$7.99626468\begin{pmatrix}3\\1\end{pmatrix}$	75.7 s	$15.994\begin{pmatrix}5\\3\end{pmatrix}$	3674.0 s
Q_f^{15}	$7.9962646819\begin{pmatrix}6\\4\end{pmatrix}$	12.4 s	$15.99440289\begin{pmatrix}3\\0\end{pmatrix}$	7924.0 s
Q_a^{15}	$7.9962646819\begin{pmatrix}6\\4\end{pmatrix}$	5.6 s	$15.99440289\begin{pmatrix}3\\0\end{pmatrix}$	2877.5 s
$[10]_n$	$\underline{7.99626468}055753$	3.6 s	$\underline{15.9}8756388887314$	3.1 s
$[10]_f$	$\underline{7.99626468194751}$	0.5 s	$\underline{15.99440289141569}$	44.5 s



(a) Adaptive strategy.

(b) Fair strategy.

Fig. 6. I_5 Domain partitioning into 4096 boxes obtained after about 3 s of CPU time with (a) Q_a^{15} and (b) Q_f^{15}.

The additional complexity of multidimensional integrals seems to magnify the effect of the expansion order into the attainable overall accuracy. As illustrated in Table 3, lower expansion orders may be significantly less accurate than higher expansion orders. With Q^5 the best attainable enclosures for I_5 and I_6 are respectively 2 and 5 orders of magnitude less accurate than with Q^{15}.

6.3. Constrained multidimensional integrals

Consider the constrained two- and three-dimensional integrals $I_7 = C_2$ and $I_8 = C_3$ where:

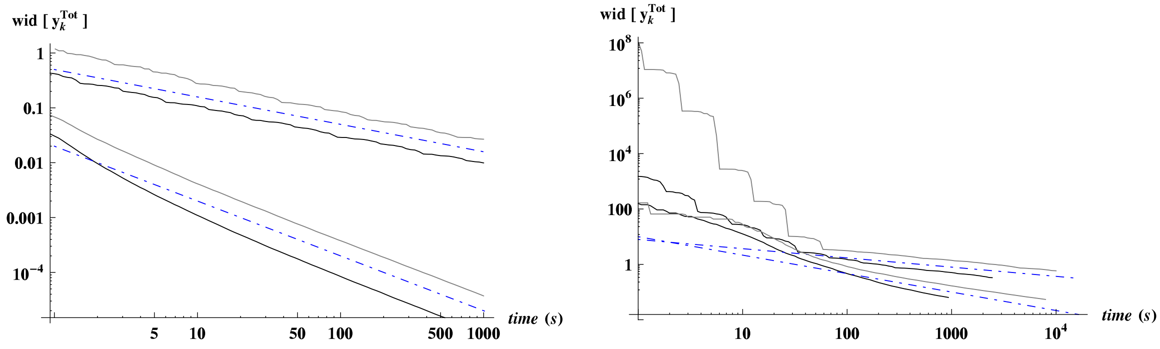
$$C_n = \int_{\|\mathbf{x}-1\|_2^2 \le 1} f(\mathbf{x})d\mathbf{x} \quad \text{with}\, f(\mathbf{x}) = 2 + \cos(20\|\mathbf{x}\|_2^2).$$

Function f is the same as illustrated in Fig. 4 for the two dimensional case, but now the domain of integration is not the entire box but rather a unit circle (ball in the 3D case) centered in the box midpoint.

Fig. 7 and Table 4 shows the results of Algorithm 1 for computing enclosures of I_7 and I_8. The last lines in the table summarize the results obtained with Mathematica (to enforce the satisfaction of the inequality constraints, the integrand function is now $f(\mathbf{x})$ multiplied by the boolean evaluation of $\|\mathbf{x} - 1\|_2^2 \le 1$).

Clearly I_7 and I_8 presented some difficulties to solve with Mathematica. The numerical computations were fast but without any guarantees (the imaginary part is meaningless in case I_8) and the formal computations were difficult (failed for case I_8). On these constrained integrals, Algorithm 1 is not able to compute enclosures as tight as in the case of integrals computed in box domains.

In the graphics we kept the theoretical convergence slopes of the respective unconstrained problems I_5 and I_6 presented in the previous section. Clearly the convergence rate in no longer related with the order of the Taylor model expansion. In fact, it is in accordance with the theoretical rates for the constrained problems explained in Section 5.4: for n-dimensional problems, the order of convergence for fair and adaptive strategies is $1/n$ and $1/(n-1)$, respectively.



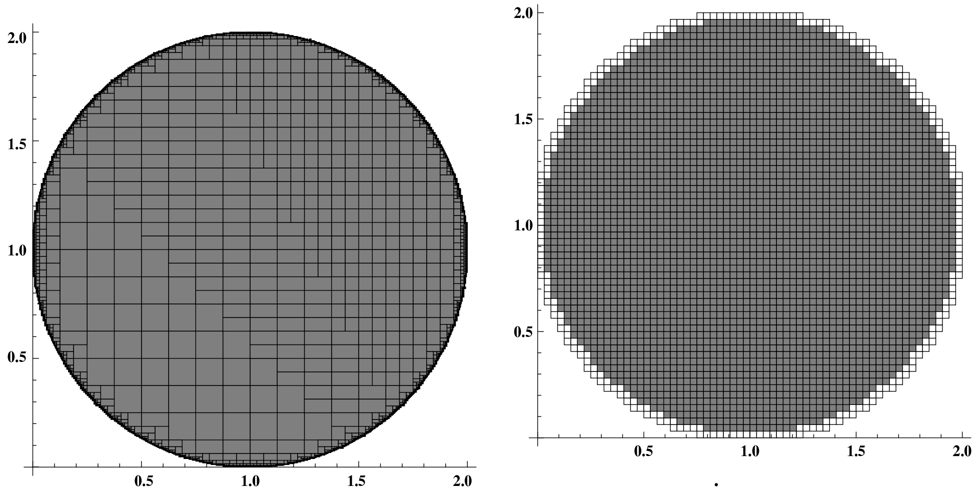
(a) I_7.

(b) I_8.

Fig. 7. Time plots for the width of the enclosures of I_7 and I_8 obtained with Q_f^5 (black stair shape lines), Q_a^5 (black smooth lines), Q_f^{15} (gray stair shape lines) and Q_a^{15} (gray smooth lines). The dot dashed strait lines represent the convergence rates predicted theoretically for constrained integrals.

Table 4
Enclosures of I_7 and I_8 obtained by Algorithm 1 and their numerical ([10] $_n$) and formal ([10] $_f$) computations with Mathematica.

	I_7		I_8	
Q_f^5	[6.296, 6.304]	1352.8 s	[8.25, 8.54]	3758.0 s
Q_a^5	$6.3001\binom{3}{1}$	414.5 s	[8.34, 8.41]	905.8 s
Q_f^{15}	[6.296, 6.304]	9072.5 s	[8.25, 8.54]	58512.3 s
Q_a^{15}	$6.3001\binom{3}{1}$	1911.6 s	[8.34, 8.41]	5686.7 s
$[10]_n$	$\underline{6.30011890}5502621$	2.3 s	$8.37845 + 7.3 \times 10^{-31} i$	3.5 s
$[10]_f$	$\underline{6.300118904167976}$	111.5 s	Fail	–



(a) Adaptive strategy.

(b) Fair strategy.

Fig. 8. I_7 Domain partitioning after about 3 s with (a) Q_a^{15} and (b) Q_f^{15}. Inner boxes are gray, and boundary boxes are white.

Fig. 8 illustrates the difference between adaptive and fair strategies with respect to domain partitioning for the constrained two dimensional case. In Q_a^{15} the number of boundary boxes (1654) is much higher than in Q_f^{15} (252) but the boundary area in Q_a^{15} is much smaller than in Q_f^{15} because the boundary boxes are very small (we cannot even distinguish white boxes).

Another remarkable difference relatively to the unconstrained integrals is the advantage of lower expansion orders when the same strategy is chosen. The reason is that the boundary boxes are by large the major contributors for the overall uncertainty and so, the improvement on the inner boxes enclosures obtained by an higher order expansion does not compensate the additional computational effort required. We can observe in Fig. 7 that, for the same selection strategy,

Table 5
Enclosures of I_9, I_{10} and I_{11} obtained by Algorithm 1 and their numerical ([10] $_n$) and formal ([10] $_f$) computations with Mathematica. For the problem I_{11}, order 0 Taylor models are used instead of order 5 since the function to be integrated is constant. The computations shown on the 5th and 6th lines have been stopped after 5 s so as to compare to [9] on the 7th line.

	I_9		I_{10}		I_{11}	
Q_f^5	$1.25\binom{7}{5}$	1582.5 s	$1.47\binom{6}{3}$	923.8 s	$1.0\binom{9}{7}$	152.9 s
Q_a^5	$1.25620\binom{6}{4}$	757.3 s	$1.47482\binom{5}{1}$	498.8 s	$1.79\binom{9}{1}$	134.7 s
Q_f^{15}	$1.25\binom{7}{5}$	44543.8 s	$1.47\binom{6}{3}$	5393.9 s	NA	NA
Q_a^{15}	$1.25620\binom{6}{4}$	14754.1 s	$1.47482\binom{5}{1}$	1928.1 s	NA	NA
Q_a^5	$1.256\binom{4}{0}$	5 s	$1.47\binom{6}{4}$	5 s	$1.0\binom{9}{7}$	5 s
Q_a^{15}	$1.25\binom{9}{3}$	5 s	$1.47\binom{6}{4}$	5 s	NA	NA
[9]	$1.2\binom{655}{471}$	5 s	$1.0\binom{5049}{4451}$	5 s	$1.0\binom{8778}{6972}$	5 s
$[10]_n$	$\underline{1.25620523}9807379$	0.3 s	$\underline{1.4748226264}573348$	0.2 s	$\underline{1.79474347}83982987$	0.7 s
$[10]_f$	$\underline{1.2562052338296295}$	201.9 s	$\underline{1.4748226264150006}$	0.8 s	$\underline{1.7947434754477893}$	7.2 s

gray lines (higher order) are always above black lines (lower order). In the limit the gray lines continue evolving further than the black lines because the space limits are only reached latter in time (to process the same number of boxes higher orders need more time).

Now consider the following multidimensional integrals (taken from [9]) which are constrained by multiple inequalities:

$$I_9 = \int_{\Omega_9} \arctan(\|\mathbf{x}\|_2^2)\, d\mathbf{x}, \qquad I_{10} = \int_{\Omega_{10}} |1 + x_1 + x_2|\, d\mathbf{x}, \qquad I_{11} = \int_{\Omega_{11}} 1\, d\mathbf{x}$$

$$\text{with} : \Omega_9 = \{\mathbf{x} \in [0, 2]^2 : \|\mathbf{x}\|_2^2 \le 4,\ x_2 \le \sin x_1\}$$

$$\Omega_{10} = \{\mathbf{x} \in [-2, 2] \times [-1, 1] : x_1^2 + x_1 \le x_2,\ x_2 \le \cos x_1\}$$

$$\Omega_{11} = \{\mathbf{x} \in [0, 2] \times [0, 2] \times [-1, 2] : \|\mathbf{x}\|_2^2 \le 4,\ (x_1 - 1)^2 + x_2^2 \le x_3\}.$$

When the domain of integration is subject to multiple inequality constraints, function g in Algorithm 1 is defined as the maximum value with respect to each individual inequality (e.g.: $g_9(\mathbf{x}) = \max(\|\mathbf{x}\|_2^2 - 4, x_2 - \sin x_1)$).

Table 5 presents the enclosures of I_9 and I_{10} computed by Algorithm 1 together with the results obtained with Mathematica [10]. Again, enclosures are not as tight as in the case of integrals computed on box domains, and the adaptive strategies are several orders of magnitude better than fair strategies both in the accuracy of the results and in execution time (e.g: the best results obtained by Q_f^5 are obtained 1000 times faster by Q_a^5). The enclosures obtained with Q_a^5 seem to outperformed those obtained in [9]. On the other hand, this has to be balanced by the fact that [9] uses an older computer. Nevertheless, the results obtained here follow the theoretical convergence results presented in the previous section.

The last problem I_{11} illustrates that, even with a trivial integrand function, tight enclosures cannot be obtained by Algorithm 1 due to the crude enclosures (A.6) computed for each boundary box. In this problem, the integral of each inner box is given by its volume (which can be sharply computed) and so, the overall width of the integral enclosure is exactly the volume of the constraint boundary (the sum of the volumes of each boundary box in the domain partitioning). Despite the trivial integrand function, I_{11} was apparently the most difficult problem from [9] where the best enclosure computed was [1.6972, 1.8778]. The enclosure computed by Algorithm 1 with the adaptive strategy (the order of the Taylor model expansion is irrelevant in this case) was $1.79\binom{9}{1}$ and took about 2 min of CPU time. The best enclosure obtained with the fair strategy was [1.77, 1.82] after about 3 min of CPU time. Clearly, the limiting factor is not the execution time but rather the covering of the boundary region with an exceedingly large set of small boxes.

7. Discussion

The experiments carried out have confirmed the theoretical convergence and convergence rate established in Section 5. As expected, experiments show that the adaptive strategy generally strongly decrease the computational time and improve the final enclosure, hence confirming the claim that in spite of repeated useless computations mentioned in introduction, the branch and prune algorithm with adaptive strategy outperforms regular meshes used in [24,1]. The algorithm under study is comparable with recent algorithms proposed for certified quadrature, and has the advantage of being intrinsically anytime: The enclosure precision can be monitored during the algorithm execution, and stopping the algorithm can be decided on this information (e.g. when a prescribed accuracy is reached, or when the accuracy improvement with respect to computational effort is not worth anymore).

Experiments also confirm that the crude enclosure of the integral on the constraint boundary dramatically decreases the efficiency of the method. Future work will focus on improving this crude enclosure, e.g. by forming higher order models of the constraint boundary that can be used in the quadrature formula.

Acknowledgment

The first author is grateful to the National Institute of Informatics for having partially funded this work through the MOU program.

Appendix. Taylor models for certified quadrature

Taylor expansions with rigorously bounded remainder (also called Taylor models) is one key object used in interval analysis. Among several fruitful applications, they allow computing certified enclosures of the integral of real functions over box domains (see [24,1]). Appendix A.1 provides the basic definitions related to Taylor models, and Appendix A.2 their usage for the quadrature over box and inequality constrained domains, as well as their necessary convergence properties.

A.1. Taylor models of real functions

A Taylor model of $f : \mathbb{R}^n \to \mathbb{R}$ inside a bounded $[\mathbf{x}]$ is a pair $(p, [r])$ where p is a polynomial with variables $x_1, \dots, x_n$ and $[r]$ an interval satisfying $f(\mathbf{x}) \in p(\mathbf{x}) + [r]$ for all $\mathbf{x} \in [\mathbf{x}]$. The degree of the Taylor model is the degree of p. There are at least two different ways of building Taylor models: The most obvious is using a multivariate Taylor expansion of f and using the interval evaluation of the highest order derivatives in order to compute a rigorous bound on the remainder. Using the multi-index notation the order m Taylor expansion of f inside $[\mathbf{x}]$ expanded at $\tilde{\mathbf{x}} \in [\mathbf{x}]$ is

$$f(\mathbf{x}) = \sum_{|\alpha|=0}^{m} \frac{1}{\alpha!} \frac{\partial^\alpha f(\tilde{\mathbf{x}})}{\partial \mathbf{x}^\alpha} (\mathbf{x} - \tilde{\mathbf{x}})^\alpha + \sum_{|\alpha|=m+1} r_\alpha(\mathbf{x}) (\mathbf{x} - \tilde{\mathbf{x}})^\alpha \tag{A.1}$$

where the remainder can be bounded as follows:

$$r_\alpha(\mathbf{x}) = \frac{1}{\alpha!} \frac{\partial^\alpha f(\mathbf{x})}{\partial \mathbf{x}^\alpha}\bigg|_{\mathbf{y}} \in \frac{1}{\alpha!} \frac{\partial^\alpha f(\mathbf{x})}{\partial \mathbf{x}^\alpha}\bigg|_{[\mathbf{x}]} =: [r_\alpha], \tag{A.2}$$

the first equality holding for some $\mathbf{y} \in \tilde{\mathbf{x}} \vee \mathbf{x} \subseteq [\mathbf{x}]$ hence the final enclosure using some interval extensions of the $(m+1)^{th}$ derivatives (computed e.g. using automatic interval differentiation). The above rigorously bounded Taylor series gives rise to the following Taylor model of f inside $[\mathbf{x}]$: $f(\mathbf{x}) \in p(\mathbf{x}) + [r]$ with

$$p(\mathbf{x}) = \sum_{|\alpha|=0}^{m} \frac{1}{\alpha!} \frac{\partial^\alpha f(\tilde{\mathbf{x}})}{\partial \mathbf{x}^\alpha} (\mathbf{x} - \tilde{\mathbf{x}})^\alpha + \sum_{|\alpha|=m+1} \operatorname{mid}[r_\alpha] (\mathbf{x} - \tilde{\mathbf{x}})^\alpha \tag{A.3}$$

and

$$[r] = \sum_{|\alpha|=m+1} ([r_\alpha] - \operatorname{mid}[r_\alpha])([\mathbf{x}] - \tilde{\mathbf{x}})^\alpha. \tag{A.4}$$

Thus, an order m Taylor series with rigorous enclosure of the remainder gives rise to a $m+1$ degree Taylor model.

Another way of computing Taylor models is to use the Taylor model arithmetic [25]. Up to our knowledge, these two methods for computing Taylor models have not been compared w.r.t. complexity, enclosure quality or easiness of implementation. We do not provide details of the Taylor model arithmetic since our focus here is the study of the convergence of the Taylor model based quadrature algorithm while no convergence analysis of the Taylor model arithmetic is available in the literature.

A.2. Quadrature using Taylor models

Taylor models have been shown to provide accurate enclosures of the quadrature of a function over a box. The following lemma is used in [24] and is proved here for completeness.

Lemma 3. *Suppose that $f(\mathbf{x}) \in p(\mathbf{x}) + [r]$ holds for all $\mathbf{x} \in [\mathbf{x}]$. Then*

$$\int_{[\mathbf{x}]} f(\mathbf{x})\, d\mathbf{x} \in \int_{[\mathbf{x}]} p(\mathbf{x})\, d\mathbf{x} + [r] \operatorname{vol}[\mathbf{x}].$$

Proof. We have $p(\mathbf{x}) + \underline{r} \le f(\mathbf{x}) \le p(\mathbf{x}) + \overline{r}$ for all $\mathbf{x} \in [\mathbf{x}]$. Therefore

$$\int_{[\mathbf{x}]} p(\mathbf{x}) + \underline{r}\, d\mathbf{x} \le \int_{[\mathbf{x}]} f(\mathbf{x})\, d\mathbf{x} \le \int_{[\mathbf{x}]} p(\mathbf{x}) + \overline{r}\, d\mathbf{x}.$$

Please cite this article in press as: A. Goldsztejn, et al., Convergence analysis and adaptive strategy for the certified quadrature over a set defined by inequalities, Journal of Computational and Applied Mathematics (2013), http://dx.doi.org/10.1016/j.cam.2013.08.035

Finally, $\int_{[\mathbf{x}]} r\, d\mathbf{x} = r \operatorname{vol}[\mathbf{x}]$ so we obtain

$$\int_{[\mathbf{x}]} p(\mathbf{x})\, d\mathbf{x} + \underline{r} \operatorname{vol}[\mathbf{x}] \le \int_{[\mathbf{x}]} f(\mathbf{x})\, d\mathbf{x} \le \int_{[\mathbf{x}]} p(\mathbf{x})\, d\mathbf{x} + \overline{r} \operatorname{vol}[\mathbf{x}]$$

which corresponds to the statement. □

Now, the following lemma provides a simple enclosure of the quadrature of a function over a inequality constrained set $\Omega = \{\mathbf{x} \in [\mathbf{x}] : g(\mathbf{x}) \le 0\}$:

Lemma 4. *Suppose that $[f]$ is an interval extension of f. Then*

$$\int_{\Omega} f(\mathbf{x})\, d\mathbf{x} \in \big(0 \vee [f]([\mathbf{x}])\big) \operatorname{vol}[\mathbf{x}].$$

Proof. Let $[r] := [f]([\mathbf{x}])$. We have $\underline{r} \le f(\mathbf{x}) \le \overline{r}$ for all $\mathbf{x} \in [\mathbf{x}]$. Therefore we have

$$\underline{r} \operatorname{vol} \Omega \le \int_{\Omega} f(\mathbf{x})\, d\mathbf{x} \le \overline{r} \operatorname{vol} \Omega. \tag{A.5}$$

It remains to notice that since $\operatorname{vol}\Omega \le \operatorname{vol}[\mathbf{x}]$ we have $\min\{0, \underline{r}\} \operatorname{vol}[\mathbf{x}] \le \underline{r} \operatorname{vol}\Omega$ and $\overline{r} \operatorname{vol}\Omega \le \max\{0, \overline{r}\} \operatorname{vol}[\mathbf{x}]$ while $[\min\{0, \underline{r}\}, \max\{0, \overline{r}\}]$ is equal by definition to $0 \vee [r]$. □

Lemmas 3 and 4 prove that the following expression is a $[g]$-quadrature inclusion function:

$$[TM^{[g]}_{f,g}]([\mathbf{x}]) := \begin{cases} \big(0 \vee [f]([\mathbf{x}])\big) \operatorname{vol}[\mathbf{x}] & \text{if } \underline{g} \le 0 < \overline{g} \\ \displaystyle\int_{[\mathbf{x}]} p(\mathbf{x})\, d\mathbf{x} + [r] \operatorname{vol}[\mathbf{x}] & \text{if } \overline{g} \le 0 \\ 0 & \text{otherwise,} \end{cases} \tag{A.6}$$

where $[\underline{g}, \overline{g}] = [g]([\mathbf{x}])$, $[f]$ is an interval extension of f, and $p(\mathbf{x})$ and $[r]$ are the Taylor model defined by (A.3) and (A.4). The integral $\int_{[\mathbf{x}]} p(\mathbf{x})\, d\mathbf{x}$ is easily computed using formal computations (and rigorously enclosed using interval arithmetic for practical implementation), and the remainder can be computed using automatic interval differentiation.

Remark 4. The automatic interval differentiation can be thought as an interval extension of the derivatives, and therefore has a linear order of convergence inside any bounded domain.

Remark 5. Some higher degree Taylor expansion can also be used in Lemma 4. However, the main source of overestimation is actually the interval joint operation $0 \vee \cdot$. Our experiments have shown that the computational cost of increasing the degree of this Taylor model is not worthwhile.

The next two propositions provide the convergent properties needed for Theorems 2 and 3 to hold using $[TM^{[g]}_{f,g}]$.

Proposition 3. *The $[g]$-quadrature inclusion function (A.6) is $[g]$-convergent inside $[\mathbf{x}^{\text{Init}}]$ provided that $[f]([\mathbf{x}^{\text{Init}}]) =: [y^{\text{Init}}]$ is bounded, and that the interval extension used for computing $[r]$ is convergent inside $[\mathbf{x}^{\text{Init}}]$.*

Proof. First, $\big(0 \vee [f]([\mathbf{x}])\big) \operatorname{vol}[\mathbf{x}]$ is weakly convergent inside $[\mathbf{x}^{\text{Init}}]$ since

$$\big(0 \vee [f]([\mathbf{x}])\big) \operatorname{vol}[\mathbf{x}] \subseteq (0 \vee [y^{\text{Init}}]) \operatorname{vol}[\mathbf{x}], \tag{A.7}$$

whose width is $(\operatorname{mag}[y^{\text{Init}}]) \operatorname{vol}[\mathbf{x}]$. Second, the width of $\int_{[\mathbf{x}]} p(\mathbf{x})\, d\mathbf{x} + [r] \operatorname{vol}[\mathbf{x}]$ is $(\operatorname{wid}[r]) \operatorname{vol}[\mathbf{x}]$, and therefore its excess is $\operatorname{wid}[r]$. As a consequence, $\int_{[\mathbf{x}]} p(\mathbf{x})\, d\mathbf{x} + [r] \operatorname{vol}[\mathbf{x}]$ is convergent inside $[\mathbf{x}^{\text{Init}}]$ provided that so is the interval extension used for computing $[r]$. □

Proposition 4. *Provided that $\sup [g]([\mathbf{x}^{\text{Init}}]) \le 0$, i.e. the quadrature is actually performed over a box, and that the interval extension used for computing $[r]$ has a linear order of convergence and is Lipschitz continuous inside $[\mathbf{x}^{\text{Init}}]$, the quadrature inclusion function (A.6) has a order of convergence $m+2$ inside $[\mathbf{x}^{\text{Init}}]$.*

Proof. Since $\sup [g]([\mathbf{x}^{\text{Init}}]) \le 0$, the Taylor model of order $m+1$ is used. Therefore,

$$\operatorname{exc}[Q_{f,g}]([\mathbf{x}]) = \frac{\operatorname{wid}[r] \operatorname{vol}[\mathbf{x}]}{\operatorname{vol}[\mathbf{x}]} = \operatorname{wid}[r] \tag{A.8}$$

$$= \sum_{|\alpha|=m+1} \operatorname{wid}\big(([r_\alpha] - \operatorname{mid}[r_\alpha])([\mathbf{x}] - \tilde{\mathbf{x}})^{\alpha}\big). \tag{A.9}$$

$$\le \sum_{|\alpha|=m+1} (\operatorname{wid}[r_\alpha])(\operatorname{wid}[\mathbf{x}])^{m+1}. \tag{A.10}$$

Since $[r_\alpha]$ has a linear order of convergence and is Lipschitz continuous, by Eq. (9) there exists $c_\alpha \geq 0$ such that

$$\operatorname{exc}[Q_{f,g}]([\mathbf{x}]) \leq \sum_{|\alpha|=m+1} c_\alpha (\operatorname{wid}[\mathbf{x}])(\operatorname{wid}[\mathbf{x}])^{m+1} \tag{A.11}$$

$$= (\operatorname{wid}[\mathbf{x}])^{m+2} \sum_{|\alpha|=m+1} c_\alpha, \tag{A.12}$$

which proves the claim. □

References

[1] R. Moore, R. Kearfott, M. Cloud, Introduction to Interval Analysis, SIAM, 2009.
[2] M.A. Malcolm, R.B. Simpson, Local versus global strategies for adaptive quadrature, ACM Transactions on Mathematical Software 1 (2) (1975) 129–146.
[3] C.-Y. Chen, Verified computed peano constants and applications in numerical quadrature, BIT Numerical Mathematics 47 (2) (2007) 297–312. http://dx.doi.org/10.1007/s10543-007-0128-x.
[4] C.-Y. Chen, Bivariate product cubature using peano kernels for local error estimates, J. Sci. Comput. 36 (2008) 69–88. http://dx.doi.org/10.1007/s10915-007-9178-0.
[5] C.-Y. Chen, On the properties of sard kernels and multiple error estimates for bounded linear functionals of bivariate functions with application to non-product cubature, Numer. Math. 122 (4) (2012) 603–643. http://dx.doi.org/10.1007/s00211-012-0471-y.
[6] C.-Y. Chen, Computing interval enclosures for definite integrals by application of triple adaptive strategies, Computing 78 (2006) 81–99. http://dx.doi.org/10.1007/s00607-006-0172-4. URL http://portal.acm.org/citation.cfm?id=1152729.1152734.
[7] D. Huybrechs, R. Cools, On generalized gaussian quadrature rules for singular and nearly singular integrals, SIAM Journal on Numerical Analysis 47 (2009) 719–739. http://dx.doi.org/10.1137/080723417.
[8] N. Hale, L.N. Trefethen, New quadrature formulas from conformal maps, SIAM Journal on Numerical Analysis 46 (2008) 930–948. http://dx.doi.org/10.1137/07068607X. URL http://portal.acm.org/citation.cfm?id=1350722.1350741.
[9] F. Kálovics, Zones and integrals, J. Comput. Appl. Math. 182 (2005) 243–251. http://dx.doi.org/10.1016/j.cam.2004.10.023.
[10] Wolfram Research Inc., Mathematica, Champaign, Illinois, Version 7.0, 2008.
[11] R. Moore, Interval Analysis, Prentice-Hall, 1966.
[12] G. Alefeld, J. Herzberger, Introduction to Interval Computations, in: Computer Science and Applied Mathematics, 1974.
[13] A. Neumaier, Interval Methods for Systems of Equations, Cambridge Univ. Press, 1990.
[14] L. Jaulin, M. Kieffer, O. Didrit, E. Walter, Applied Interval Analysis with Examples in Parameter and State Estimation, Robust Control and Robotics, Springer-Verlag, 2001.
[15] R.B. Kearfott, Interval computations: introduction, uses, and resources, Euromath Bulletin 2 (1) (1996) 95–112.
[16] O. Knueppel, PROFIL/BIAS — a fast interval library, Computing 53 (3–4) (1994) 277–287.
[17] A. Griewank, D. Juedes, J. Utke, Algorithm 755: ADOL-C: A package for the automatic differentiation of algorithms written in C/C++, ACM Transactions on Mathematical Software 22 (2) (1996) 131–167.
[18] I. Charpentier, J. Utke, Fast higher-order derivative tensors with Rapsodia, Optimization Methods Software 24 (1) (2009) 1–14. http://dx.doi.org/10.1080/10556780802413769.
[19] C.H. Bischof, P. Hovland, B. Norris, On the implementation of automatic differentiation tools, Higher-Order and Symbolic Computation 21 (3) (2008) 311–331. http://dx.doi.org/10.1007/s10990-008-9034-4.
[20] M. Berz, Forward algorithms for high orders and many variables with application to beam physics, in: A. Griewank, G.F. Corliss (Eds.), Automatic Differentiation of Algorithms: Theory, Implementation, and Application, SIAM, Philadelphia, PA, 1991, pp. 147–156.
[21] R.D. Neidinger, An efficient method for the numerical evaluation of partial derivatives of arbitrary order, ACM Transactions on Mathematical Software 18 (2) (1992) 159–173.
[22] A. Griewank, J. Utke, A. Walther, Evaluating higher derivative tensors by forward propagation of univariate Taylor series, Math. Comp. 69 (2000) 1117–1130.
[23] R.D. Neidinger, Directions for computing truncated multivariate Taylor series, Math. Comp. 74 (249) (2005) 321–340.
[24] M. Berz, K. Makino, New methods for high-dimensional verified quadrature, Reliable Computing 5 (1999) 13–22.
[25] M. Berz, G. Hoffstatter, Computation and application of Taylor polynomials with interval remainder bounds, Reliable Computing 4 (1998) 83–97.

Please cite this article in press as: A. Goldsztejn, et al., Convergence analysis and adaptive strategy for the certified quadrature over a set defined by inequalities, Journal of Computational and Applied Mathematics (2013), http://dx.doi.org/10.1016/j.cam.2013.08.035