Evaluation of peak functions on ultra-coarse grids

Petrovskaya, Natalia; Embleton, Nina

Document Version
Peer reviewed version

Link to publication on Research at Birmingham portal

Evaluation of Peak Functions on Ultra-Coarse Grids

Journal:	*Proceedings A*
Manuscript ID:	Draft
Article Type:	Research
Date Submitted by the Author:	n/a
Complete List of Authors:	Petrovskaya, Natalia; University of Birmingham, School of Mathematics
Subject:	Computer modelling and simulation < COMPUTER SCIENCE, Computational physics < COMPUTER SCIENCE, Applied mathematics < MATHEMATICS
Keywords:	sampled data, sparse data, peak function, coarse grid, midpoint rule

Evaluation of Peak Functions on Ultra-Coarse Grids

Natalia Petrovskaya$^{a\#}$ and Nina Embleton^{b}

^{a} School of Mathematics, University of Birmingham,
Birmingham, B15 2TT, U.K.
Email: n.b.petrovskaya@bham.ac.uk

^{b} School of Mathematics, University of Birmingham,
Birmingham, B15 2TT, U.K.
Email: embleton@for.mat.bham.ac.uk

Abstract

Integration of sampled data arises in many practical applications, where the integrand function is available from experimental measurements only. One extensive field of research is the problem of pest monitoring and control where an accurate evaluation of the population size from the spatial density distribution is required for a given pest species. High aggregation population density distributions (peak functions) are an important class of data that often appear in this problem. The main difficulty associated with the integration of such functions is that the function values are usually only available at a few locations therefore new techniques are required to evaluate the accuracy of integration as the standard approach based on convergence analysis does not work when the data are sparse. Thus in our paper we introduce the new concept of ultra-coarse grids for high aggregation density distributions. Integration of the density function on ultra-coarse grids cannot provide the prescribed accuracy because of the insufficient information (uncertainty) about the integrand function. Instead, the results of the integration should be treated probabilistically by considering the integration error as a random variable and we show how the corresponding probabilities can be calculated. Handling the integration error as a random variable allows us to evaluate the accuracy of integration on very coarse grids where asymptotic error estimates cannot be applied.

Keywords: sampled data, sparse data, peak function, coarse grid, midpoint rule

---

$^{\#}$ Corresponding author

Introduction

Integration of sampled data arises in a wide class of problems, as in many practical applications the integrand function is available from experimental measurements only and is therefore given by a discrete set of function values. Numerous examples include acoustics and signal processing [1, 2], image reconstruction [3], microbiology [4], ecological applications [5], etc. A general problem of numerical integration has a long and successful history and many accurate and efficient methods for integration of sampled data have been designed and documented in the literature [6, 7, 8, 9]. Meanwhile the increasing complexity of practical problems resulted in the recent need to revise existing algorithms as new problems have emerged. One such problem is that of insect pest monitoring where an accurate evaluation of the population size from the spatial density distribution is required for a given biological species [10].

Let us consider a particular problem of pest insect monitoring in a single agricultural field as an instructive example. An accurate estimate of the total number of pest insects is crucial for making a reliable decision about the use of pesticides in the area where the crop is grown [11]. The data for evaluation of the pest population size is usually collected by trapping. Insect traps are installed at the nodes of a uniform Cartesian grid in the agricultural field, they are exposed for a certain time, and then the traps are emptied and the caught insects are counted [12]. Under the assumption that the number of insects in each trap gives us the true value of the population density obtained at the location of the trap [13], the methods of numerical integration on uniform grids can be employed to estimate the total number of pest insects from the discrete density distribution [14]. However, as we show it below, the application of well-known methods, such as the Newton -Cotes integration rules on a uniform grid, is not straightforward at all.

The main difficulty associated with the estimation of the total number of pest insects from trap counts is that the number N of traps at each direction cannot be made large enough to ensure that the integral estimate is accurate. Installment of many traps per a unit agricultural area would in itself bring considerable damage to the agricultural product. Also, trapping is costly and labour-consuming and it introduces a disturbance to agricultural procedures. Hence the problem of numerical integration has to be solved for a small number N of traps.

From a computational viewpoint, the evaluation of the population size when the number N of traps is small presents the problem of numerical integration of the integrand function on a very coarse uniform Cartesian grid. Moreover, grid adaptation is not possible in the problem as the grid should be generated only once and N cannot be further increased. This restriction appears because a repeated trapping with an increased number of traps is not available in ecological applications due to the impossibility to reproduce the initial conditions.

Under the restrictions outlined above two following questions arise:

- What is the minimum number N_t of traps required to achieve desirable accuracy?

- What can be an alternative measure of accuracy on a coarse grid of traps where $N < N_t$?

Although the processing of sparse data has intensively been studied in various problems from physics and engineering (e.g., see [15, 16, 17, 18]), to the best of our knowledge numerical integration of sparse data has not been discussed in the literature. An attempt to address the questions above has been made in our recent work [19, 20, 14]. In particular, it has been demonstrated in [14] that the answers to those questions depend strongly on the integrand function under consideration. Namely, the accuracy of integration remains acceptable even on very coarse grids if the pest population is distributed more or less over the whole area, no matter whether this distribution is close to homogeneous or has a complex heterogeneous structure. On the contrary, the accuracy is very poor when one has to integrate a high aggregation density distribution on a coarse grid.

The high aggregation density distributions (peak functions) are an important class of data that may appear in ecosystems under various conditions. For instance, one common scenario of biological invasion is that the

2

pest species starts spreading from a small localized area and invades the entire domain as time progresses [21]. Obviously the decision about the application of pesticides is best made before the 'spot' of high density will spread over the whole agricultural field. Hence timely and accurate evaluation of the total number of pest insects at earlier stages of biological invasion is very beneficial for the cultivation of the agricultural product. At the same time the application of numerical integration methods is significantly hampered by the fact that the exact location of the high density sub-domain is not known in the problem. Thus, instead of installing the traps locally (i.e., in the initially infested area) to increase the accuracy of integration, a uniform grid of traps has to be generated over the entire domain. That makes numerical integration of peak functions a very difficult task, as the information about the density function on such a sparse grid may be insufficient for reasonably accurate integration.

The aim of our paper is to investigate this problem in order to answer the questions highlighted above. In doing this, we develop a new approach to work with sparse data and show that integration of peak functions on coarse grids may lead to a paradigm shift. Our research of ultra-coarse grids will be focused on the one-dimensional case but the approach we present in the paper can be extended to two-dimensional density distributions.

The paper is organized as follows. In Section 1 we briefly revisit the problem of numerical integration and introduce the concept of ultra-coarse grids for high aggregation density distributions. It will be shown that integration on ultra-coarse grids cannot provide the prescribed accuracy because of the insufficient information (uncertainty) about the integrand function. Instead, the results of the integration should be treated probabilistically by considering the integration error as a random variable. We calculate the probability of an accurate integral estimate in Section 2, while in Section 3 we consider the transition from ultra-coarse grids to coarse grids where the integration error becomes deterministic. Numerical examples are considered in Section 4. The discussion of our results is provided in Section 5.

1 Numerical integration of sparse data

The invasion regime when a spreading pest population forms the strongly heterogeneous patchy spatial distribution, has its one-dimensional counterpart when a peak density function appears somewhere in a sub-domain D_u at the unit interval $D = [0, 1]$ [19]. One computational problem related to the study of ecologically meaningful density distributions is that the integral is not available in closed form. Hence before moving to the discussion of a more general case we consider several examples where the exact value of the integral is available and the integration error can therefore be computed.

A convenient example of a peak function is given by the normal distribution

$$u(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\frac{(x-x^*)^2}{\sigma^2}\right), \tag{1}$$

shown in Fig. 1b for the peak width $\delta = 6\sigma = 0.25$. Let us numerically integrate (1) over the unit interval $D = [0, 1]$. Consider a uniform grid generated in the domain D as $x_{i+1} = x_i + h$, $i = 1, \ldots, N-1$, where the grid step size is $h = 1/(N-1)$. We employ the midpoint rule of integration [26] as a baseline integration method in our problem. Once a computational grid has been generated, the integral I is computed by the compound midpoint rule as

$$I = \int\limits_0^1 u(x)dx \approx \tilde{I} = \sum_i w_i u(x_i), \tag{2}$$

where $w_i = h$ for the interior nodes $i = 2, \ldots, N-1$ and $w_i = h/2$ at boundary points $i = 1$, $i = N$.

In our discussion of the accuracy of numerical integration on coarse grids we will also compute the integral by the compound Simpson method where the number N is required to be an odd number, $N = 2m + 1$, and the
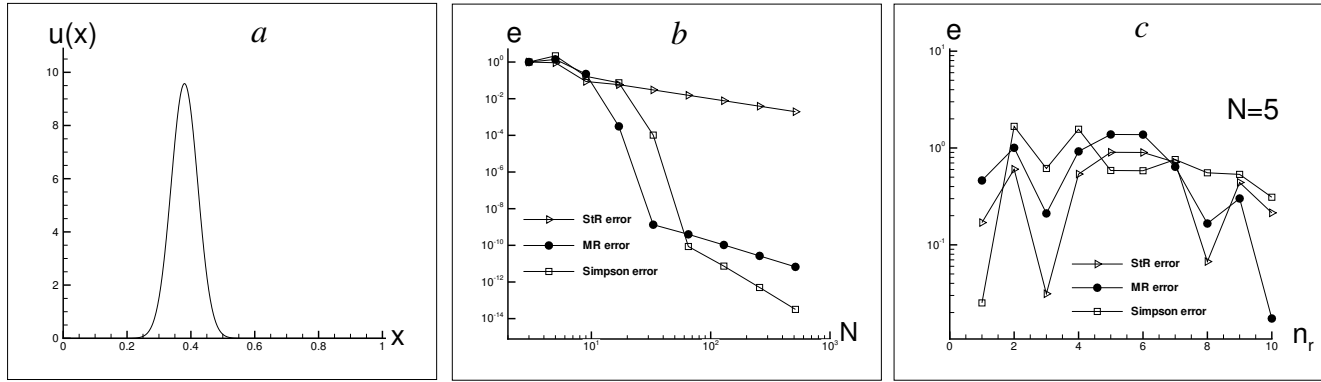
3

Figure 1: *Numerical integration of a peak function. (a) The peak function (1), where the peak width is $\delta = 0.25$ and the peak is located at $x^* = 0.38$. (b) The integration error (4) computed for the peak function of Fig. 1a on a sequence of uniformly refined grids. The midpoint rule (MR): solid line, closed circle; the Simpson rule: solid line, open square; the statistical rule (3) (StR): solid line, open right triangle. The uniform refinement of the original coarse grid of $N = 3$ nodes does not decrease the integration error, unless the domain of the non-zero density is resolved. c) The integration error computed when the peak (1) is randomly located on the uniform grid of $N = 5$ nodes. The error (4) is shown for the ten realizations n_r of the random variable x^*. The midpoint rule (MR): solid line, closed circle; the Simpson rule: solid line, open square; the statistical rule (3) (StR): solid line, open right triangle.*

weights in (2) are given by $w_i = 4h/3$, $i = 2, 4, \ldots, 2m$, $w_i = 2h/3$, $i = 3, 5, \ldots, 2m - 1$, and $w_i = h/3$, $i = 1$, $i = N$. Finally, the third method we employ for numerical integration is a so called 'statistical method' widely used in ecological applications [34]. The method evaluates the integral as

$$I \approx A\bar{u} = \frac{A}{N} \sum_{i=1}^{N} u(x_i), \tag{3}$$

where A is the given area ($A = 1$ when we integrate over the unit interval).

We define the integration error as

$$e = \frac{|I - \tilde{I}|}{|I|}, \tag{4}$$

where I is the exact integral and $\tilde{I}$ is the approximate integral computed by the chosen method of numerical integration. The accuracy of integration should be

$$e \leq \tau, \tag{5}$$

where τ is specified tolerance. It is important to note here that in ecological applications the accuracy requirements on coarse grids are essentially different from those arising in the conventional problem of numerical integration, as the tolerance $0.2 < \tau < 0.5$ is considered as acceptable [27, 28]. However, we will see further in the text that even such relatively low accuracy of computations is not always achievable when the number of grid nodes is small.

Let us compute the integration error (4) for the density distribution (1). Consider first the midpoint rule (2) of integration. The error (4) of the midpoint rule is shown as a function of the number N of grid nodes in Fig. 1b. It can be seen from the figure that the integration error of the midpoint rule is not controlled by uniform grid refinement on very coarse grids that are in the focus of ecological research. Namely, the error $e \sim 1$ remains beyond the acceptable range $e \sim 0.2 - 0.5$ on grids with $N < \tilde{N}$, where $\tilde{N} \sim 10$ is the upper bound for the realistic number of traps to be considered in ecological applications [29, 30].

4

As the uniform refinement does not decrease the error on coarse grids (see Fig. 1b), the obvious decision would be to apply a more accurate method of numerical integration (i.e., the Simpson method; see [26]) in order to improve the accuracy of integration on grids with small N. However, it can be seen from Fig. 1b, where the error of the Simpson method is shown on a sequence of uniformly refined grids, that the Simpson method is not more accurate than the midpoint rule on coarse grids. On the other hand, the method (3) is theoretically less accurate than (2) (see [26]), but it cannot be said from Fig. 1b that the method (2) is better than (3) when coarse grids are considered.

Let us recall that the exact location of the 'peak sub-domain' D_u is unknown to us. Meanwhile the value of the approximate integral depends obviously on how many grid nodes are stationed inside the sub-domain D_u on a coarse grid. In order to better understand how the integration error (4) depends on the location of the peak with respect to the position of grid nodes, we now consider the peak location x^* in the function (1) as a uniformly distributed random variable. Let us fix the number of grid nodes as $N = 5$ and randomly move the peak (1) over the domain D 10 times. We integrate the function (1) every time that we move the peak, i.e. for each of the 10 realizations n_r of the random variable x^*. The results of numerical integration are shown in Fig. 1c, where the integration error (4) is computed for each of the 10 locations of the peak on the coarse grid of 5 nodes. For instance, when $n_r = 3$ the peak's location is $x^* = 0.7013$, while for $n_r = 7$ the same peak is located at $x^* = 0.4188$, etc. It is readily seen from the graph of Fig. 1c that for the midpoint rule of integration the error (4) depends essentially on the peak location as the error varies as $0.0173 \leq e \leq 1.379$ on the grid with the fixed number of nodes. The same observation with regard to the integration error is true for the Simpson rule and the integration rule (3).

It follows from the results of the test case (1) that the accuracy of computation cannot be determined when the high aggregation density distributions are considered on grids with a small number of nodes. Since the integration error depends on a random location of the point x^*, the error itself becomes a random variable and a probabilistic approach should be used to evaluate the accuracy. We will refer to the grids, where the accuracy of numerical integration cannot be determined, as *ultra-coarse grids*. Correspondingly, the questions formulated in the Introduction are re-formulated on ultra-coarse grids as follows:

1 Given the number N of grid nodes on a uniform grid (the grid step size $h = const$), what is the probability of the event $e \leq \tau_0$, where the integration error e is given by (4) and τ_0 is the chosen tolerance?

2 What is the threshold number N_t of grid nodes when the error (4) becomes deterministic?

In the next two sections of our paper we answer the questions (1) and (2) when peak functions are integrated using the general midpoint rule (2).

2 Ultra-coarse grids

Consider a high aggregation density distribution $u(x)$ modelled by the following peak function

$$u(x) = \begin{cases} f(x) > 0, & x \in (x_I, x_{II}), \\ 0, & otherwise, \end{cases} \tag{6}$$

where we assume that the function $f(x)$ has a single maximum at point $x^* = 0.5(x_I + x_{II})$. At this stage we are not interested in a more detailed definition of the function $f(x)$, as particular examples of $u(x)$ will be considered further in the text. A sketch of the peak function (6) is shown in Fig. 2.
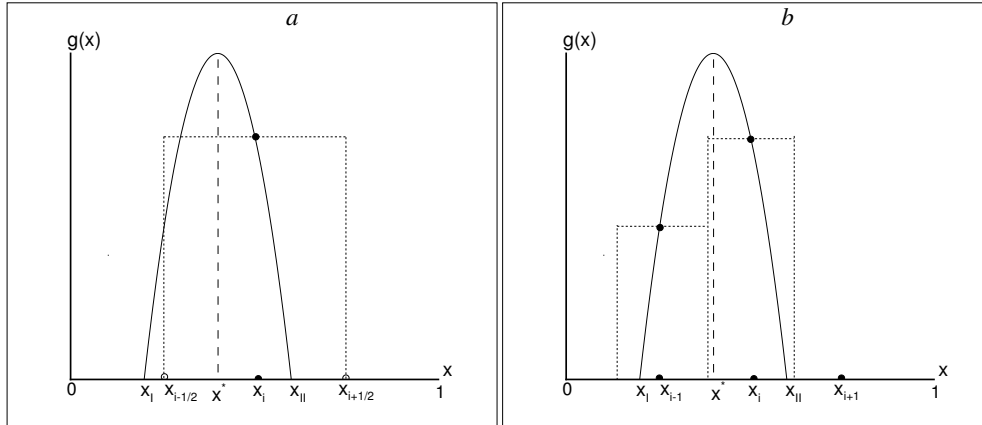
5

Figure 2: *The midpoint integration rule for the peak function. (a) One grid node is located within the sub-domain $[x_I, x_{II}]$. (b) Two grid nodes belong to the peak sub-domain.*

Let us expand $u(x)$ at the maximum point x^* as $u(x) = u(x^*) + \dfrac{1}{2}\dfrac{d^2 u(x^*)}{dx^2}(x - x^*)^2 + R(x)$ and assume that the remainder $R(x)$ can be neglected in the vicinity of the peak, so that the integrand becomes

$$u(x) \approx g(x) = B - A(x - x^*)^2, \quad x \in [x_I, x_{II}],$$
$$u(x) = 0, \quad otherwise, \tag{7}$$

where $A = -\dfrac{1}{2}\dfrac{d^2 u(x^*)}{dx^2} > 0$, $B = u(x^*) > 0$. The integral I is then given by

$$I = \int\limits_0^1 u(x)dx \approx \int\limits_{x_I}^{x_{II}} g(x)dx = \frac{2}{3}B\delta, \tag{8}$$

where the peak width δ is defined as

$$\delta \equiv x_{II} - x_I = 2\sqrt{B/A}. \tag{9}$$

Consider a uniform grid of N nodes generated in the domain $[0, 1]$ as $x_{i+1} = x_i + h$, $i = 1, \ldots, N - 1$, where the grid step size is $h = 1/(N - 1)$. We begin our study of ultra-coarse grids with the case when the grid step size $h > \delta$. In other words, we require that the grid is so coarse that only one grid node x_i can be located in the sub-domain $[x_I, x_{II}]$ (see Fig. 2a).

Let the grid step size be

$$h = \alpha\delta, \tag{10}$$

where the parameter $\alpha > 1$. We also parametrize the location of the grid node x_i in the sub-domain $[x_I, x_{II}]$ as

$$x_i = x^* + \gamma\,h, \quad \gamma \in [0, \frac{1}{2}], \tag{11}$$

in order to compute the approximate integral $\tilde{I}$. The midpoint rule approximation of the integral for the geometry $\alpha > 1$ is shown in Fig. 2a. The integral $\tilde{I}$ computed by the midpoint rule when we use the approximation (7) for $u(x)$ is as follows:

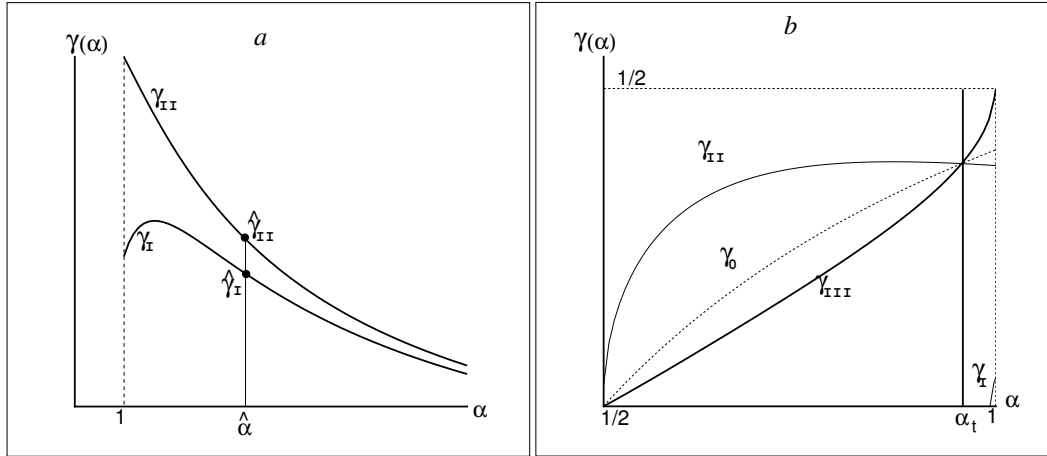$$\tilde{I} = \left(B - A\gamma^2 h^2\right)h. \tag{12}$$

6

Figure 3: *The range of parameter $\gamma = \gamma(\alpha)$ for which the integration error is $e \le \tau_0$. (a) The grid step size is $h > \delta$ where δ is the peak width. (b) The grid step size is $\delta/2 \le h \le \delta$.*

Hence the integration error (4) is

$$e = \frac{|\frac{2}{3}B\delta + A\gamma^2 h^3 - Bh|}{\frac{2}{3}B\delta} \le \tau_0.$$

Without loss of generality let us choose the tolerance as $\tau_0 = 1/4$. Taking into account (9) and (10) and solving the inequalities

$$\frac{3}{4}I \le \tilde{I} \le \frac{5}{4}I \tag{13}$$

for the node location γ, we obtain the following condition

$$\gamma_I(\alpha) \le \gamma(\alpha) \le \gamma_{II}(\alpha), \tag{14}$$

where

$$\gamma_I(\alpha) = \frac{1}{2\alpha}\sqrt{1 - \frac{5}{6\alpha}}, \quad \gamma_{II}(\alpha) = \frac{1}{2\alpha}\sqrt{1 - \frac{1}{2\alpha}}, \tag{15}$$

for the tolerance $\tau_0 = 1/4$.

For the sake of the discussion in Section 3 it is worth noting here that the inequality $\gamma_I(\alpha) \le \gamma(\alpha)$ implies that $\alpha \ge \alpha_I = 5/6$. If $\alpha < \alpha_I$ then $\tilde{I} \le \frac{5}{4}I$ always holds and the inequalities (14) should be replaced as

$$0 \le \gamma(\alpha) \le \gamma_{II}(\alpha). \tag{16}$$

Similarly, the inequality $\gamma(\alpha) \le \gamma_{II}(\alpha)$ requires $\alpha \ge \alpha_{II} = 1/2$.

The curves $\gamma_I(\alpha)$ and $\gamma_{II}(\alpha)$ are shown in Fig. 3a. The conditions (15) define the parameter range where integral is computed with the required accuracy τ_0. Consider a peak of the width δ and let us fix $\alpha = \hat{\alpha} > 1$, so that the grid step size becomes fixed as $h = \hat{h} = \hat{\alpha}\delta$. Compute $\hat{\gamma}_I = \gamma_I(\hat{\alpha})$ and $\hat{\gamma}_{II} = \gamma_{II}(\hat{\alpha})$ on the grid of size $\hat{h}$. The inequalities (14) provide the error $e \le \tau_0$ for any $\hat{\gamma}_I \le \gamma \le \hat{\gamma}_{II}$ (see Fig. 3a).

Let us assume that the location x^* of the peak maximum can be found at any point of the domain $[0, 1]$ with equal probability, that is the random variable x^* is uniformly distributed. It then readily follows from the above consideration that the probability $p(h)$ of achieving the error $e \le \tau_0$ on a given grid of N nodes is given by

$$p(h) = (\gamma_{II}(h) - \gamma_I(h))/(\gamma_{max} - \gamma_{min}) = 2(\gamma_{II}(h) - \gamma_I(h)),$$

where the entire range of γ is given by $\gamma_{min} = 0$, $\gamma_{max} = 1/2$, the grid step size is $h = 1/(N-1)$ and we use the transformation $\alpha = h/\delta$ for the fixed peak width δ. Straightforward analysis of expressions (15) shows that the probability p of achieving the integration error $e \leq \tau_0$ remains $p < 1$ on any ultra-coarse grid where the condition $h > \delta$ holds.

We summarize the findings of this section as follows:

- We have shown that the integration error should be handled is a random variable on ultra-coarse grids where the data available for integration are sparse.

- We have considered a quadratic approximation of the integrand function and found the probability p of an accurate answer when a quadratic polynomial is integrated on an ultra-coarse grid.

Meanwhile, the important question that still remains is: If we gradually increase the number of grid nodes N, what is the threshold number N_t for which the error becomes deterministic? The answer to this question will be found in the next section.

3 Transition from ultra-coarse grids to coarse grids

In this section we investigate the transition from grids where the integration error is a random variable to grids where the condition $e \leq \tau_0$ always holds for the given tolerance τ_0. Let us increase the number N of grid nodes in order to decrease the grid step size h as

$$\delta/2 \leq h \leq \delta, \tag{17}$$

where δ again is the peak width (9). In other words, we now require that $1/2 \leq \alpha \leq 1$ when the parametrization (10) is used. If the condition (17) holds, then either one or two grid nodes belong to the sub-domain $[x_I, x_{II}]$ (see Fig. 2b).

Consider the location (11) of grid node x_i. The minimum value γ_0 that provides the location of two grid nodes x_{i-1} and x_i in the sub-domain $[x_I, x_{II}]$ is defined from the conditions $x_{i-1} = x^* - \delta/2$ and $x_{i-1} = x_i - h$. We have

$$\gamma_0 = 1 - \frac{\delta}{2h} = 1 - \frac{1}{2\alpha}, \tag{18}$$

where the parametrization (10) is taken into account.

For $\gamma \in [0, \gamma_0)$ only one grid point belongs to the interval $[x_I, x_{II}]$, and we can use the result (15) to compute the probability of accurate integration. Hence we now focus on the range $\gamma \in [\gamma_0, 1/2]$ when two grid points are captured by the peak as it is shown in Fig. 2b.

Let us use again the quadratic approximation (7) of the integrand function. Since $g(x_{i-1}) = -A((\gamma-1)h)^2 + B$ and $g(x_i) = -A(\gamma h)^2 + B$, the integral $\tilde{I}$ is computed by the midpoint rule as follows

$$\tilde{I} = (g(x_{i-1}) + g(x_i))h = -A((\gamma-1)^2 + \gamma^2)h^3 + 2Bh.$$

For the sake of simplicity we require again the tolerance $\tau_0 = 1/4$ to arrive at the inequalities (13). Consider first the inequality $\tilde{I} \geq \frac{3}{4}I$. Simple algebraic transformation results in

$$\gamma^2(\alpha) - \gamma(\alpha) + C(\alpha) \leq 0,$$

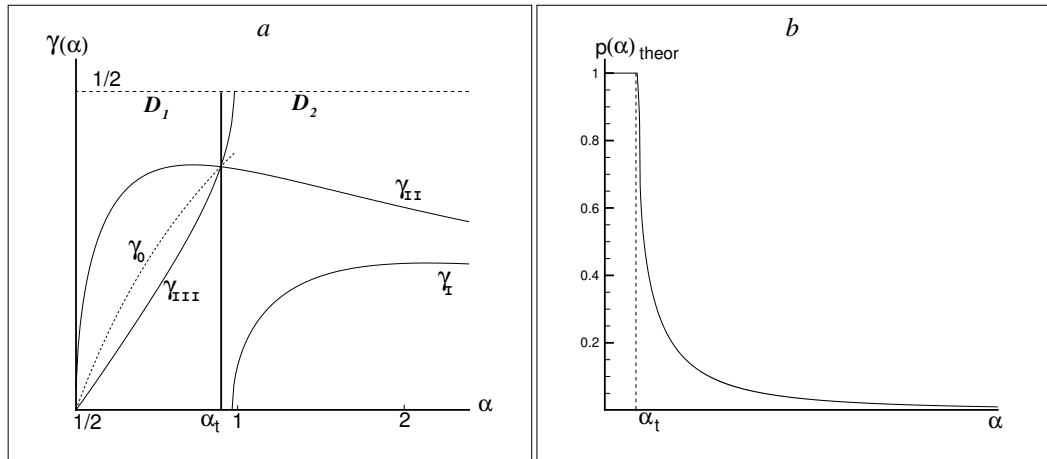where $C(\alpha) = \frac{1}{2} - \frac{1}{16\alpha^3}(4\alpha - 1)$.

8

Figure 4: *(a) The set of parametric curves defining the admissible range of node location γ on grids with the grid step size $h \geq \delta/2$. In the domain $D_1 : \alpha \in [1/2, \alpha_t]$ the condition $e \leq \tau_0$ holds for any γ, while in the domain $D_2 : \alpha > \alpha_t$ the integration error becomes a random variable. (b) The probability $p(\alpha)$ of having the integration error $e \leq \tau_0$ for a peak of the width δ integrated on a grid with the grid step size $h = \alpha\delta$.*

Consider the roots $\gamma_{III} = \dfrac{1 - \sqrt{1 - 4C(\alpha)}}{2}$ and $\gamma_{IV} = \dfrac{1 + \sqrt{1 - 4C(\alpha)}}{2}$ of the equation $\gamma^2 - \gamma + C = 0$. The non-empty range $\gamma \in [\gamma_{III}, \gamma_{IV}]$ exists, if the inequality $4C(\alpha) \leq 1$ holds. Substituting the above expression for $C(\alpha)$ into this inequality we obtain

$$4\alpha^3 - 4\alpha + 1 \leq 0. \tag{19}$$

Numerical solution of the equation (19) gives us the roots $\alpha_1 \approx -1.1072$, $\alpha_2 \approx 0.2696$ and $\alpha_3 \approx 0.8376$. Hence if $\alpha \in [1/2, \alpha_3]$, the range $\gamma \in [\gamma_{III}, \gamma_{IV}]$ will provide us with the integration error $e \leq \tau_0 = 1/4$, where we should also take into account the restriction $\gamma \in [\gamma_0, \dfrac{1}{2}]$. It readily follows from the above computation that $\gamma_{III}(\alpha) < \dfrac{1}{2}$ and $\gamma_{IV}(\alpha) > \dfrac{1}{2}$ for any $\alpha \in (\dfrac{1}{2}, 1)$.

Consider now the lower boundary γ_0. The curve $\gamma_{III}(\alpha)$ intersects the curve $\gamma_0(\alpha)$ at point α_t (see Fig. 3b). We require $\gamma_0(\alpha_t) = \gamma_{III}(\alpha_t)$ to obtain the equation

$$8\alpha_t^3 - 8\alpha_t^2 + 1 = 0. \tag{20}$$

The only root α_t in the subinterval $[1/2, 1]$ is $\alpha_t \approx 0.8090$. Hence the curve $\gamma_{III}(\alpha)$ lies under the curve $\gamma_0(\alpha)$ for $\alpha \in [1/2, \alpha_t]$ and it is above the curve $\gamma_0(\alpha)$ when $\alpha \in [\alpha_t, 1]$ (see Fig. 3b). Let us also note that $\alpha_t < \alpha_3$.

Finally, we consider the inequality $\tilde{I} \leq \dfrac{5}{4}I$ where after some algebraic transformations we arrive at

$$\gamma^2(\alpha) - \gamma(\alpha) + D(\alpha) \geq 0, \tag{21}$$

with $D(\alpha)$ given by $D(\alpha) = \dfrac{1}{2} - \dfrac{1}{48\alpha^3}(12\alpha - 5)$. The requirement $4D(\alpha) \leq 1$ results in the inequality

$$4\alpha^3 - 4\alpha + \frac{5}{3} \leq 0,$$

that does not have any real roots for $\alpha > 0$. Hence the inequality (21) holds for any value of γ.

9

Let us now compute the probability $p = p(\alpha)_{theor}$ of the event that the error is $e \leq \tau_0$ on a grid with fixed grid step size $h = \alpha\delta$, where $\alpha \geq 1/2$. The entire domain $\alpha \geq 1/2$ is shown in Fig. 4a, where the curves $\gamma(\alpha)$ of Fig. 3a and Fig. 3b are now 'glued' together. The following cases should be considered.

1. $\alpha \in [\frac{1}{2}, \alpha_t]$, where $\alpha_t \approx 0.8090$ has been obtained as a solution to the equation (20) derived under the condition that the tolerance $\tau_0 = \frac{1}{4}$. Since $\gamma_{III}(\alpha) < \gamma_0(\alpha)$ and $\gamma_{IV}(\alpha) > \frac{1}{2}$ then for any $\gamma \in [\gamma_0, \frac{1}{2}]$ the conditions (13) hold. In other words, if there are two grid points in the sub-domain $[x_I, x_{II}]$ then the integration error is $e \leq \tau_0$, no matter how those grid points are located with respect to the maximum point x^*. The probability of the accurate answer is $p_2(\alpha) = (1/2 - \gamma_0(\alpha))/(\gamma_{max} - \gamma_{min}) = 1 - 2\gamma_0(\alpha)$.

Consider now $\gamma \in [0, \gamma_0]$, so that just one grid point is located in the peak sub-domain. The admissible range of γ is then given by the inequality (14). Let us investigate the position of the curve $\gamma_{II}(\alpha)$ with respect to the curve $\gamma_0(\alpha)$. Simplifying the equation $\gamma_0(\alpha) = \gamma_{II}(\alpha)$ we obtain the same cubic equation for α as the equation (20). Hence the three curves $\gamma_0(\alpha), \gamma_{II}(\alpha)$ and $\gamma_{III}(\alpha)$ intersect in a single point α_t (see Fig. 2b).

As $\gamma_0(\alpha) < \gamma_{II}(\alpha)$ for $\alpha \in [\frac{1}{2}, \alpha_t]$, the upper bound in the inequality (14) should be replaced as $\gamma_I(\alpha) < \gamma(\alpha) < \gamma_0(\alpha)$. At the same time, the lower bound in (14) requires the restriction $\alpha \geq \alpha_I$ that does not hold for $\alpha \in [\frac{1}{2}, \alpha_t]$ and therefore the condition (16) should be considered. Hence the inequality (14) is transformed as $0 \leq \gamma(\alpha) \leq \gamma_0(\alpha)$, if there is one grid point in the sub-domain $[x_I, x_{II}]$. That means the integration error is $e \leq \tau_0$ for the grid step size $h < \alpha_t\delta$. As the admissible range becomes $0 \leq \gamma(\alpha) \leq \gamma_0(\alpha)$, the probability is $p_1(\alpha) = (\gamma_0(\alpha) - 0)/(\gamma_{max} - \gamma_{min}) = 2\gamma_0(\alpha)$. Hence, the resulting probability of the accurate answer is $p(\alpha)_{theor} = p_1(\alpha) + p_2(\alpha) = 1$ and the condition $e \leq \tau_0$ holds for any γ if the grid step size is $h = \alpha\delta$, where $\alpha \in [1/2, \alpha_t]$ for the fixed peak width δ.

2. $\alpha \in [\alpha_t, \alpha_I]$, where $\alpha_I = \frac{5}{6} \approx 0.8333$ for $\tau_0 = \frac{1}{4}$ (see Section 2). From condition (16) we now have $p_1(\alpha) = (\gamma_{II}(\alpha) - 0)/(\gamma_{max} - \gamma_{min}) = 2\gamma_{II}(\alpha)$ for $\gamma \in [0, \gamma_0]$ and $p_2(\alpha) = (1/2 - \gamma_{III}(\alpha))/(\gamma_{max} - \gamma_{min}) = 1 - 2\gamma_{III}(\alpha)$ if $\gamma \in [\gamma_0(\alpha), \frac{1}{2}]$. The resulting probability is $p(\alpha)_{theor} = p_1(\alpha) + p_2(\alpha) < 1$.

3. $\alpha \in [\alpha_I, \alpha_3]$, where we have $\alpha_3 \approx 0.8376$ from (19) for $\tau_0 = \frac{1}{4}$. For this range of α we have $p_1(\alpha) = (\gamma_{II}(\alpha) - \gamma_I(\alpha))/(\gamma_{max} - \gamma_{min}) = 2(\gamma_{II}(\alpha) - \gamma_I(\alpha))$, as the inequality (14) now holds for any $\gamma \in [0, \gamma_0]$. We also have $p_2(\alpha) = (\frac{1}{2} - \gamma_{III}(\alpha))/(\gamma_{max} - \gamma_{min}) = 1 - 2\gamma_{III}(\alpha), \gamma \in [\gamma_0, \frac{1}{2}]$. The probability $p(\alpha)_{theor} = p_1(\alpha) + p_2(\alpha) < 1$.

4. $\alpha > \alpha_3$. We compute the probability p as $p(\alpha)_{theor} = p_1(\alpha) = (\gamma_{II}(\alpha) - \gamma_I(\alpha))/(\gamma_{max} - \gamma_{min}) = 2(\gamma_{II}(\alpha) - \gamma_I(\alpha))$ (see also Section 2). The probability $p_2 = 0$ because of the restriction $\alpha \in [\frac{1}{2}, \alpha_3]$ required for the inequality $\tilde{I} \geq \frac{3}{4}I$ when $\gamma \in [\gamma_0, 1/2]$.

The function $p(\alpha)_{theor}$ is shown in Fig. 4b. For the fixed width δ of the peak the parameter α_t is the threshold value of the grid step size that provides the transition from ultra-coarse grids to coarse grids. On any grid with $\alpha < \alpha_t$ (domain D_1 in Fig. 4a) the error (4) is deterministic in the sense that the condition $e \leq \tau_0$ always holds. On ultra-coarse grids where $\alpha > \alpha_t$ (domain D_2 in Fig. 4a) the error (4) is a random variable as the probability of getting the accurate answer is $p < 1$.

Below we summarize the main results of the Section 3:

- We have found the threshold number N_t of grid nodes such that the integration error e is always $e < \tau$, if the number of grid nodes $N > N_t$. In other words, the error becomes deterministic when we approach the threshold N_t.

- It has been shown that the grid step size h_t that corresponds to the threshold number N_t on a uniform grid
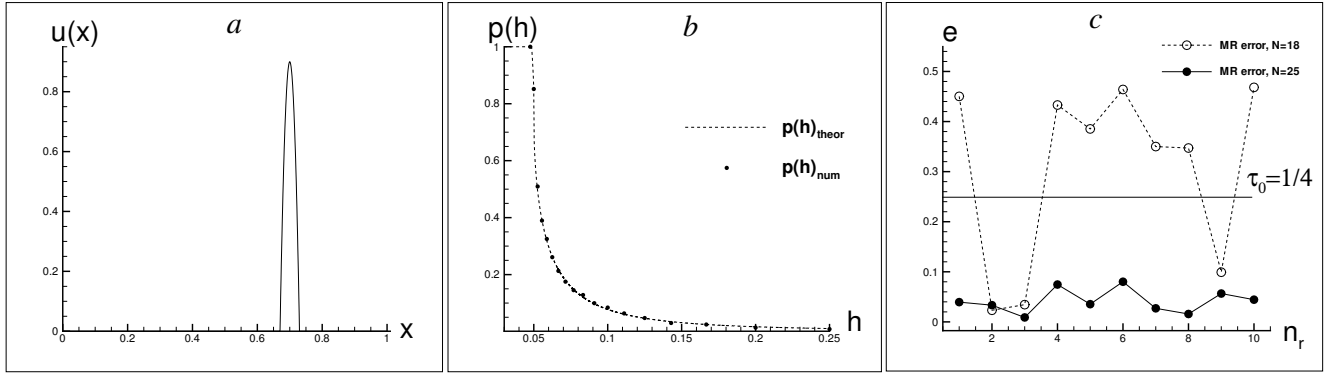
10

Figure 5: *(a) A quadratic function (7). The peak width is $\delta = 0.06$ and the parameters are $A = 1000$, $B = 0.9$. (b) The probability $p(h)_{num}$ obtained by direct computation agrees with the theoretical results $p(h)_{theor}$ for the function (7) of Fig. 6a. (c) The integration error for the function (7) on an ultra-coarse grid and a coarse grid with the fixed number of nodes. The error (4) is shown for the ten realizations n_r of the random variable x^*, where x^* is uniformly distributed. The probability of achieving an accurate answer (5) is $p \approx 0.3$ on an ultra-coarse grid of $N = 18$ nodes (dashed line), while $p = 1$ on a coarse grid of $N = 25$ nodes (solid line).*

is a linear function of the peak width δ, that is $h_t = \alpha_t \delta$, where α_t depends on the chosen tolerance τ only. The coefficient α has been computed for the tolerance $\tau = 0.25$ as $\alpha \approx 0.8333$.

The results above have been obtained under the assumption of a quadratic approximation of the integrand function. Hence, numerical investigation of the problem is required when a peak funciton has the shape different from quadratic in order to verify our findings. That will be done in the next section.

4 Numerical examples

In this section we first consider several standard test cases to illustrate our approach to numerical integration on ultra-coarse grids. We then turn our attention to ecologically meaningful density distributions and discuss how the theoretical predictions of Section 3 work for them.

4.1 Standard test cases

We begin our consideration with a quadratic function (7), as our first test case is to verify the probability estimate $p(h)_{theor}$ derived in Section 3. Let us fix the peak width δ and consider the peak location x^* as a random variable that is uniformly distributed over the interval $[\delta, 1 - \delta]$, as we require that the entire peak is stationed within the unit interval $[0, 1]$. In our test we provide $n_r = 10^4$ realizations of the random variable x^* on a grid with the fixed number N_l of nodes and compute the integral error (4) where we integrate the function (7) by the midpoint rule for each realization of x^*. The probability $p(h_l)_{num}$ of the accurate numerical integration is computed as

$$p(h_l)_{num} = \frac{\hat{n}_r}{n_r}, \tag{22}$$

where $h_l = 1/(N_l - 1)$ is the grid step size on a grid of N_l nodes and $\hat{n}_r$ is the number of realizations for which the integration error is $e \leq \tau_0$, $\tau_0 = 1/4$. We then increase the number of grid nodes as $N_{l+1} = N_l + 1$ and repeat computation (22). We stop when the number N_L of grid nodes results in a grid with the grid step size $h_L \leq \delta/2$.
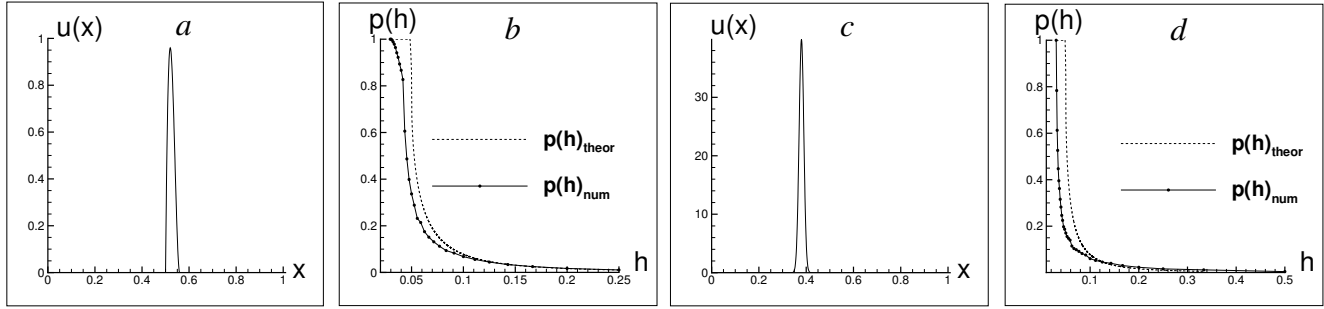
11

Figure 6: *Numerical test cases: (a) the cubic function (23), (c) the normal distribution (1). For the functions (a) and (c) the peak width is chosen as $\delta = 0.06$. (d)-(f) The probability (22) (solid line) of an accurate answer (5) and its comparison with the theoretical curve $p(h)_{theor}$ (dashed line) obtained for the quadratic function: (b) the probability graph $p(h)_{num}$ computed for the for the cubic function (23) (d) the probability $p(h)_{num}$ for the normal distribution (1).*

The quadratic function (7) is shown in Fig. 5a for the peak width $\delta = 0.06$. The probability $p(h)_{num}$ of the accurate numerical integration (5) of the function (7) is shown in Fig. 5b. We start from the grid of $N_1 = 5$ nodes and end our computations on the grid of $N_{17} = 22$ nodes where the condition (17) still holds. It can be seen from the figure that all values of the probability $p(h_l)$, $l = 1, \ldots, 17$, computed by direct evaluation (22) lie very close to the theoretical curve $p(h)_{theor}$ shown as a dashed curve in the figure.

The results of Sections 2 and 3 are further illustrated in Fig. 5c where the integration error (4) is computed for the function (7) on an ultra-coarse grid and a coarse grid. The computation resulting in the graphs shown in Fig. 5c is similar to the test case discussed in Section 2 (cf. Fig. 1c). Namely, we randomly move the peak (7) ten times ($n_r = 10$) on a grid with a fixed number of nodes and compute the integration error every time that the peak is moved. It can be seen from the figure that the error (4) depends on the peak location when we integrate the function (7) on an ultra-coarse grid with the number of grid nodes $N = 18$ (i.e., the grid step size is $h = 1/17$). The probability of achieving an accurate answer (5) is $p \approx 0.3$ (see the graph in Fig. 5b). Meanwhile, the error is deterministic on a grid of $N = 25$ nodes ($h = 1/24$) and the error is $e \leq \tau_0$, no matter where the peak is located.

We now consider several peak functions different from the quadratic function (7) in order to understand how the probability estimate obtained for (7) will work for them.

A cubic function

$$u(x) = \begin{cases} A(x - x^* + (\delta/3))(x - x^* - (2\delta/3))^2, & x \in [x^* - \delta/3, x^* + \delta/3], \\ \\ 0, & otherwise, \end{cases} \quad (23)$$

presents an interesting test case because the peak now is asymmetric and we therefore use the notation x^* for the location of the first root rather than for the maximum point. The function (23) is shown in Fig. 6a, where we still keep the peak width $\delta = 0.06$ for $A = 3 \cdot 10^4$. The probability graph for the function (23) is shown in Fig. 6b. The graph $p(h)_{num}$ is in a good agreement with the curve $p(h)_{theor}$ on very coarse grids ($\alpha > \alpha_3$) where the probability of accurate integration is small. At the same time, the actual probability in the transition layer is smaller than the probability estimate based on the quadratic approximation. One important observation about the graph $p(h)_{num}$ of Fig. 6b is that the grid step size h_t^{num} for which the error (4) becomes deterministic (i.e. $p(h_t^{num}) = 1$ and $p(h) < 1$ for any $h > h_t^{num}$), is smaller than the theoretical estimate

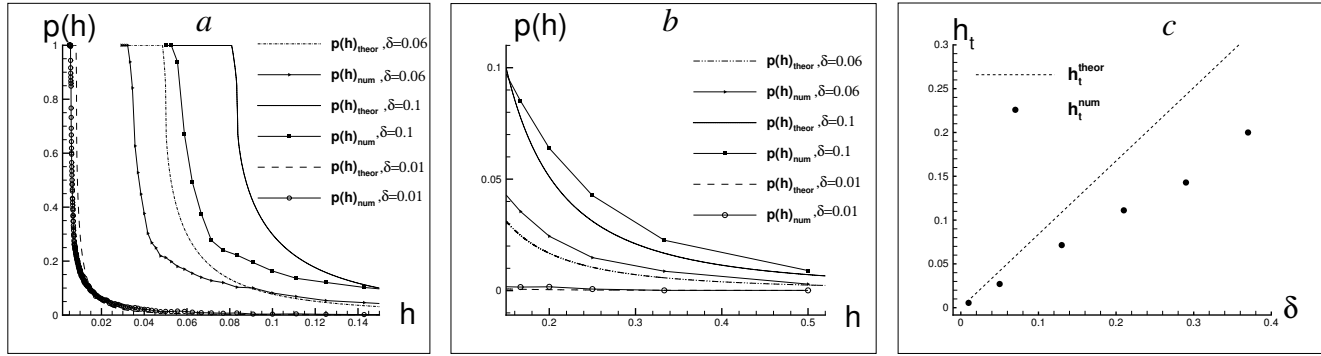$$h_t = \alpha_t \delta, \quad (24)$$

12

Figure 7: *Numerical test cases: (a) the probability graphs $p(h)_{num}$ computed for the function (25) when the peak width δ varies. $\delta = 0.06$: solid line, closed right triangle; circle; $\delta = 0.1$: solid line, closed square; $\delta = 0.01$: solid line, open circle; The computed probability $p(h)_{num}$ is compared with the theoretical estimate $p(h)_{theor}$ made for a quadratic function. $\delta = 0.06$: dash-dot-dotted line; $\delta = 0.1$: solid line; $\delta = 0.01$: dashed line; (b) the "tail" of the probability graphs of Fig. 7a. The legend is the same as in Fig. 7a. (c) the threshold values h_t^{num}, computed for the function (25) when the peak width δ varies, are compared with the theoretical curve (24) (dashed line) obtained for a quadratic function.*

obtained for the quadratic approximation of the integrand function.

The normal distribution (1) already discussed in Section 1 gives us an example of a peak function that is different from zero everywhere in the domain $x \in [0, 1]$ (see Fig. 6c). However, the peak of width $\delta = 6\sigma$ is a dominant feature of the function (1), and therefore we can expect a random integration error when we integrate the normal distribution on ultra-coarse grids. The graph $p(h)_{num}$ computed for the function (1) with peak width $\delta = 0.06$ is shown in Fig. 6d. It can be seen from the figure that the entire probability graph $p(h)_{num}$ is now shifted with respect to the curve $p(h)_{theor}$. The maximum deviation is $d_p = 0.8129$ and the threshold value of the grid step size is $h_t^{num} \approx 0.5\delta$ instead of $h_t \approx 0.8\delta$ obtained for the quadratic function.

A smaller threshold grid step size h_t^{num} is a consequence of the interpolation error $e_{int} = O(\delta^3)$ that we introduce when we approximate the integrand function by a quadratic polynomial. As the interpolation error depends on the peak width δ, our probability estimate $p(h)_{theor}$ should be accurate enough when δ is small (a narrow peak) and it will diverge from the actual probability $p(h)_{num}$ when we increase the peak width δ. Below we study this dependence in more detail by considering the Lorentz distribution - another peak function that often appears in various fields of physics and interdisciplinary research (e.g., see [35]),

$$
u(x) = \begin{cases} \dfrac{\delta^2}{4} \dfrac{1}{4(x - x^*) + \delta^2/4}, & x \in [x^* - \delta/2, x^* + \delta/2]. \\[2em] 0, & otherwise, \end{cases} \tag{25}
$$

Let us vary the peak width in (25) and compute the probability graph $p(h)_{num}$ for each fixed value δ by employing the procedure (22). A family of probability curves is shown in Fig. 7 for the function (25) with the peak width $\delta = 0.06$ (a baseline peak), the peak width $\delta = 0.1$ (a wide peak) and the peak width $\delta = 0.01$ (a narrow peak). For the sake of convenience we show the 'transition layer' in Fig. 7a, while the 'tails' of the probability curves are shown in in Fig. 7b. It can be seen from the figure that the computed probability curve $p(h)_{num}$ gets closer to the probability graph $p(h)_{theor}$ when we decrease the peak width δ. From the ecological viewpoint, the most important conclusion we derive from the consideration of the probability graphs shown in Fig. 7 is that the difference between the predicted value (24) and the actual threshold value h_t^{num} obtained by direct computation

13

decreases when we decrease δ. In other words, we can rely upon the estimate (24) to determine how many grid nodes are required to guarantee the accuracy (5) when a narrow peak is integrated. Thus our next test is to compute the function $h_t^{num}(\delta)$ and to compare it with the theoretical estimate (24).

The results of the computation of h_t^{num} for the function (25) where the peak width δ varies are shown in Fig. 7c. The value h_t^{num} is computed for each fixed peak width δ from the condition that $p(h_t^{num}) = 1$ and $p(h) < 1$ if $h > h_t^{num}$. The computed threshold grid step size is in a very good agreement with the estimate (24) for narrow peaks ($\delta < 0.1$). At the same time, the theoretical curve (24) lies above the actual values when wide peaks are considered and therefore the estimate (24) can be used in ecological problems as the upper bound for the threshold value h_t.

The standard numerical test cases discussed in this section confirmed that

- Replacing the integrand function by a quadratic polynomial can be considered a reliable approximation when the probability of accurate integration is computed.

- The formula (24) gives us a good estimate of the threshold grid step size h_t when a narrow peak is integrated.

- The same formula (24) can be used as the upper bound for h_t if a wide peak is considered.

Once the the estimate (24) of the threshold grid step size h_t has been verified for standard numerical test cases, our next goal is to check how formula (24) will work when ecologically meaningful peak functions are considered. That will be done in the next sub-section.

4.2 Ecological test cases

We use the pest density distributions $u(x)$ generated from the spatially explicit predator-prey model describing spatiotemporal dynamics of a pest insect population [22, 23],

$$\frac{\partial u(x,t)}{\partial t} = d\frac{\partial^2 u}{\partial x^2} + u(1-u) - \frac{uv}{u+p} ,$$

$$\frac{\partial v(x,t)}{\partial t} = d\frac{\partial^2 v}{\partial x^2} + k\frac{uv}{u+p} - mv . \tag{26}$$

The model (26) presents a system of coupled diffusion-reaction equations written in dimensionless variables. The functions $u(x,t)$ and $v(x,t)$ are the densities of the pest insect (considered as the prey) and its consumer (the predator), respectively, at time $t > 0$ and position x. Coefficient d describes species diffusivity due to the movement of the individuals, p is the half-saturation prey density, k is the food assimilation efficiency coefficient and m is the predator mortality. Also, numerical solution of the system (26) requires us to augment the equations by some initial conditions $u(x,0)$, $v(x,0)$.

Ideally, our results on the accuracy of integration on ultra-coarse grids should be checked against appropriate ecological data. However, one serious obstacle in the problem is that our further discussion will require the handling of data on a sequence of refined grids and it is not possible to fulfil this requirement when field data are considered. Meanwhile the model (26) and its two-dimensional counterpart (e.g., see [22]) have been thoroughly validated against experimental data to demonstrate that the model is in good agreement with the results of field measurements [5]. Hence we believe the model (26) can be used for the generation of ecologically realistic data.

The equations (26) provide a rich variety of spatiotemporal distributions of the pest density function $u(x,t)$. In particular, it is well known (e.g., see [5]) that the properties of the spatial distribution $u(x)$ considered at a given time t are determined by the diffusion d in the problem and that the initial conditions $u(x,0)$, $v(x,0)$ can evolve
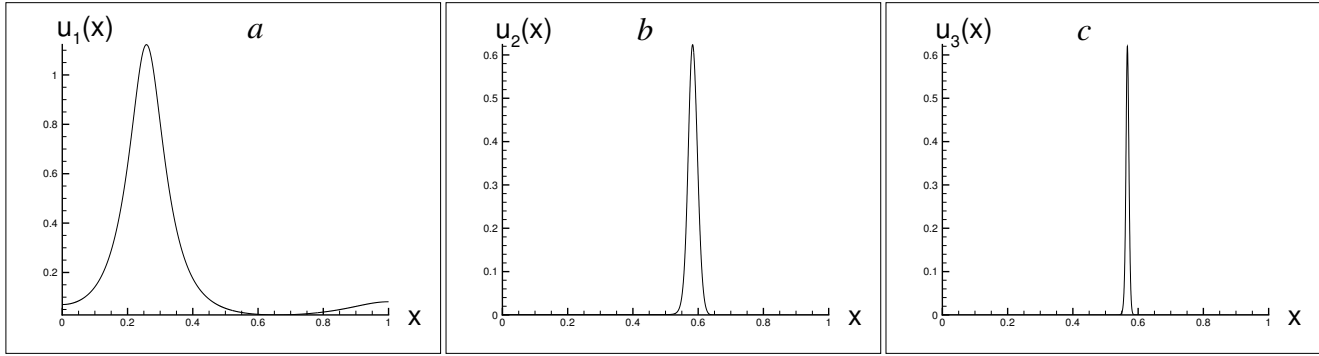
14

Figure 8: *Ecological test cases. The spatial distribution of the pest population density $u(x)$ as predicted by the model (26) for different values of the diffusivity d; (a) the density distribution $u_1(x)$ has been obtained for for $d = 10^{-4}$ (b) the pest population density $u_2(x)$ obtained for $d = 10^{-5}$ (c) the density $u_3(x)$ for $d = 10^{-6}$. An example of the system's parameters (the density distribution $u_2(x)$): $t = 50$, $k = 0.5$, $p = 2.0$, $m = 0.42$. The initial conditions are $u(x,0) = u_0, 0 < x < x_u$, $v(x,0) = v_0, 0 \leq x \leq x_v$, where $u_0 = 0.8$, $v_0 = 0.5$, $x_u = 0.6$, $x_v = 0.55$, and $u(x,0) = 0, x > x_u$, $v(x,0) = 0, x > x_v$.*

into the one-peak spatial pattern if the diffusion is $d \ll 1$ (see [25] for a more detailed discussion on the pattern formation). Several examples of one-peak density distributions are shown in Fig. 8, where the functions $u_1(x)$, $u_2(x)$ and $u_3(x)$ have been obtained from the numerical solution of the equations (26) for the diffusion coefficient $d = 10^{-4}$, $d = 10^{-5}$ and $d = 10^{-6}$ respectively. One important observation that can be made from Fig. 8 is that the peak width δ depends on the diffusion coefficient d. A simple estimate of the function $\delta(d)$ discussed in [20] can be written as

$$\delta = \omega\sqrt{d}, \tag{27}$$

where the coefficient ω in the expression (27) depends on the system's parameters. An extensive numerical study performed in [24, 25] revealed that, in the predator-prey system (26), the value ω is relatively robust to changes in the parameter values, and can typically be considered as $\omega \approx 25$. Hence we can evaluate the upper bound for the threshold grid step size h_t as

$$h_t = \alpha_t\delta \approx C\sqrt{d}, \tag{28}$$

where the coefficient $C = \alpha_t\omega$.

Consider the function $u_1(x)$ shown in Fig. 8a. The diffusion coefficient is $d = 10^{-4}$, hence the estimate (28) gives us the grid step size as $h_t \approx 0.2$. The corresponding number of grid nodes is $N_t \approx 6$.

Let us compute the integration error (4) by the midpoint rule on a sequence of refined grids, where we add just one node to the grid at each consecutive refinement step. The results of numerical integration of the function $u_1(x)$ are shown in Table 1.

N	3	4	5	6	7	8	9	10
h	0.5	0.3333	0.25	0.20	0.1667	0.1429	0.125	0.1111
e	0.6948	0.1119	0.5459	0.07983	0.1699	0.02305	0.08228	0.001918

Table 1: The integration error (4) for the density distribution $u_1(x)$ on a sequence of refined grids with grid step size h. N is the number of grid nodes on a uniform grid.

It can be seen from the table that our estimate of the threshold grid step size is in surprisingly good agreement

15

with the results of computation. The actual threshold number obtained in computations is $N_t = 6$ and, while the integration error keeps oscillating on grids with $N > N_t$ nodes, the condition (5) always holds on those grids.

Let us repeat the previous computational procedure for the function $u_2(x)$ shown in Fig. 8b. The diffusion coefficient $d = 10^{-5}$, for which the density distribution $u_2(x)$ has been generated, gives us the upper bound for the threshold number of nodes estimated as $N_t \approx 16$. The results of numerical integration of the function $u_2(x)$ are shown in Table 2 where the threshold number obtained in computations is $N_t = 21$. That number of traps is not realistic in real-life computations and other factors should be taken into account when a decision is made on the number of traps to be installed. One such factor that will be considered in future work is that the integral value depends obviously on the peak width and the true value of the integral can be small if a narrow peak is considered. A small value of the integral (i.e., a small number of pest insects) means, in turn, that the risk of making a wrong decision about pesticide application is low. Hence we can increase the tolerance τ_0 in the problem (e.g., $\tau_0 = 0.5$ or even $\tau_0 = 1$ in some cases) and use the technique discussed in Section 2 and 3 in order to re-compute our estimate of the threshold number N_t for the new tolerance.

N	17	18	19	20	21	22	23	24	25
h	0.0625	0.0588	0.0556	0.0526	0.05	0.0476	0.0455	0.0435	0.0416
e	0.4127	0.5412	0.5101	0.4124	0.1960	0.0028	0.1454	0.2234	0.2137

Table 2: The integration error (4) for the density distribution $u_2(x)$ on a sequence of refined grids with grid step size h. N is the number of grid nodes on a uniform grid.

Finally, let us have a look at a very narrow peak $u_3(x)$ shown in in Fig. 8c. The diffusion coefficient $d = 10^{-6}$ gives us the estimate $N_t \approx 50$. Clearly this number of traps is beyond any reasonable range and cannot be used in real-life applications. However, as the integral is very small (for the density distribution $u_3(x)$ we have $I = \int_0^1 u_3(x)dx = 0.007161$) our recommendation to ecologists would probably be to not take any action until the time evolves and the peak function gets wider. On the other hand, an important conclusion derived from the extreme case of the distribution $u_3(x)$ is that installing a reasonable number of traps (e.g., $N \sim 10$) with the aim to try to evaluate the total number of insects at very early stage of the biological invasion is senseless, as the probability of a correct answer is very small (see the discussion of the case $h > \delta$ in Section 2).

5 Conclusions

We have considered a problem of numerical integration for high aggregation density distributions (peak functions) on coarse uniform grids. Our study has originally been prompted by the needs of ecological monitoring and control where minimization of the number of measurements made in order to accurately reconstruct a density distribution remains an important problem.

The main results of the paper are as follows:

1. We have introduced the concept of ultra-coarse grids. An ultra-coarse grid is defined as a grid where the integration error is a random variable because of the insufficient information (uncertainty) about the integrand function.

2. The definition of a coarse grid implies that we have to evaluate the probability p of achieving an integration error smaller than the given tolerance rather than to evaluate the error itself when computations on ultra-

16

coarse grids are performed. In our paper we have obtained a probability estimate based on a quadratic approximation of the integrand function. We then carried out a number of numerical test cases to demonstrate that our probability estimate gives us a reliable upper bound for the threshold grid step size h_t for which the probability $p(h_t) = 1$. Thus the approach suggested in the paper can be used to evaluate the minimum number N_t, such that the desirable accuracy of integration is guaranteed on a grid of N_t nodes.

3. It has been shown in the paper that conventional methods of the error control and analysis do not work on ultra-coarse grids. It has been discussed in Section 1 of the paper that the uniform refinement of an ultra-coarse grid does not reduce the integration error, as the integration error on a refined grid can be even bigger than the error on the original grid unless the number N_t of nodes is reached. However, increasing the number N of grid nodes does increase the probability of an accurate integral estimate and is therefore desirable. This conclusion corrects the results of paper [32] where it was suggested that the refinement of a coarse grid is not efficient as it does not improve the accuracy of integration.

One important direction of future research is to investigate the transition from ultra-coarse to coarse grids for other classes of functions. Peak functions present just one class of density distributions and a similar problem of accurate integral evaluation arises when other spatially heterogeneous functions are considered. In particular, we are interested in rapidly oscillating functions that often appear in ecological applications. Also, our future work will be to extend our results to the two-dimensional case as we want to apply our approach to real-life problems. Since field measurements are carried out on a regular basis in ecological applications, heuristic estimates are available for many common pest species distributions and practical recommendations with regard to the number N_t of traps have already been obtained based on these heuristic estimates. We intend to check those recommendations against a theoretical probability estimate of the threshold number N_t after our evaluation technique is extended to two-dimensional problems.

The examples of integration of ecologically meaningful distributions we discussed in the paper can only be considered as the very first steps on our way to design a robust procedure for risk evaluation in agricultural applications. At the same time our results demonstrate that a conventional approach that ecologists use in their problems has to be revisited. A standard procedure of the risk evaluation in pest management is to compare an estimate of the total number of pest insects with a certain critical number and to make a decision about the use of pesticides based on that comparison. Ecologists readily admit that there is uncertainty surrounding the estimation of the pest abundance, which may become worse as the number of samples decreases [33]. However, to the best of our knowledge, the error in the pest abundance estimation has never been considered in the ecological literature as a random variable in order to take into account the risk factor related to the uncertainty in integral evaluation when the number N of traps is small. Taking this risk factor into account when designing an appropriate methodology for decision making in pest insects management should constitute a topic of our future work.

References

[1] D. Smallwood, Integration of equally spaced sampled data using a generalized Whittaker's reconstruction formula, IEEE Transactions on Acoustics, Speech and Signal Processing, 28 (1980) 341–343.

[2] L. Yaroslavsky, A. Moreno, J.Campos, Frequency responses and resolving power of numerical integration of sampled data. Optics Express, 13 (2005) 2892–2905.

[3] S.X. Liao, M. Pawlak, On image analysis by moments. IEEE Transactions on Pattern Analysis and Machine Intelligence, 18 (1996) 254–266.

17

[4] S.M. Dunn, A. Constantinides, P.V. Moghe, Numerical Methods in Biomedical Engineering, Academic Press, USA, 2006.

[5] H. Malchow, S.V. Petrovskii, E. Venturino, Spatiotemporal Patterns in Ecology and Epidemiology: Theory, Models, and Simulations, Chapman & Hall / CRC Press, London, 2008.

[6] H. Engels, Numerical Quadrature and Cubature, Academic Press, USA, 1980.

[7] G. Evans, Practical Numerical Integration, Wiley, NY, 1993.

[8] R.Piessens, E. deDoncker-Kapenga, E. Uberhuber, D. Kahaner, QUADPACK: A Subroutine Package for Automatic Integration, Springer-Verlag, Berlin, 1983.

[9] W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, Numerical Recipes in C, Cambridge University Press, 1988.

[10] N. Shigesada, K. Kawasaki, Biological Invasions: Theory and Practice, Oxford University Press, 2001.

[11] V.M. Stern, Economic thresholds, Ann. Rev. Entomol., 18 (1973) 259–280.

[12] J.A. Byers, O. Anderbrant, J. Lofqvist, Effective attraction radius: A method for comparing species attractants and determining densities of flying insects, J. Chem. Ecol., 15 (1989) 749–765.

[13] S. Petrovskii, D. Bearup, D.A. Ahmed, R. Blackshaw, Estimating insect population density from trap counts. Ecological Complexity. 10 (2012), 69–82.

[14] N.B. Petrovskaya, S.V. Petrovskii, A.K. Murchie, Challenges of ecological monitoring: estimating population abundance from sparse trap counts, J. R. Soc. Interface, 68 (2012), 420–435.

[15] P.J.S.G. Ferreira, Iterative and noniterative recovery of missing samples for 1-D band-limited signals, in: F.A. Marvasti (Ed.), Sampling Theory and Practice, Plenum Publishing Corporation, 2001, pp.235–282.

[16] P. Sava, Micro-earthquake monitoring with sparsely-sampled data, Journal of Petroleum Exploration and Production Technology, 1 (2011) 43–49.

[17] L.P. Yaroslavsky, G. Shabat, B.G. Salomon, I.A. Ideses, B. Fishbain, Nonuniform sampling, image recovery from sparse data and the discrete sampling theorem, J. Opt. Soc. Am. A, 26 (2009) 566–575.

[18] J.Z. Wilcox, T.J. Wilcox, Algorithm for extraction of periodic signals from sparse, irregularly sampled data, Astron. Astrophys. Suppl. Ser., 112 (1995) 395–405.

[19] N.B. Petrovskaya, S.V. Petrovskii, The coarse mesh problem in ecological monitoring, Proc. R. Soc. A, 466 (2010) 2933–2953.

[20] N.B. Petrovskaya, N.L. Embleton, S.V. Petrovskii, Numerical study of pest population size at various diffusion rates, in: M.Lewis, P.Maini, S.Petrovskii (Eds.), Dispersal, Individual Movement and Spatial Ecology: a Mathematical Perspective, Lecture Notes in Mathematics, Springer, 2012, in press.

[21] S.V. Petrovskii, A.Y. Morozov, E. Venturino, Allee effect makes possible patchy invasion in a predator-prey system, Ecol. Lett., 5 (2002) 345–352.

[22] J.D. Murray, Mathematical Biology, Springer, Germany, 1989.

18

[23] P. Turchin, Complex Population Dynamics: a Theoretical / Empirical Synthesis, Princeton University Press, USA, 2003.

[24] S.V. Petrovskii, H. Malchow, Spatio-temporal chaos in an ecological community as a response to unfavorable environmental changes, Advances in Complex Systems, 4 (2001) 227–250.

[25] S.V. Petrovskii, B.-L. Li, H. Malchow, Quantification of the spatial aspect of chaotic dynamics in biological and chemical systems, Bull. Math. Biol., 65 (2003) 425–446.

[26] P.J. Davis, P. Rabinowitz, Methods of Numerical Integration, Academic Press, NY, 1975.

[27] M.A. Pascual, P. Kareiva, Predicting the outcome of competition using experimental data: Maximum likelihood and Bayesian approaches, Ecology, 77 (1996) 337–349.

[28] J.A. Sherratt, M. Smith, Periodic travelling waves in cyclic populations: field studies and reaction diffusion models, J. R. Soc. Interface, 5 (2008) 483–505.

[29] J.G. Mayor, M.H. Davies, A survey of leatherjacket populations in south-west England, 19631974, Plant Pathol., 25 (1976) 121–128.

[30] P. Northing, Extensive field based aphid monitoring as an information tool for the UK seed potato industry, Asp. Appl. Biol., 94 (2009) 31–34.

[31] S.V. Petrovskii, H. Malchow, F.M. Hilker, E. Venturino, Patterns of patchy spread in deterministic and stochastic models of biological invasion and biological control, Biol. Invas., 7 (2005) 771–793.

[32] N.B. Petrovskaya, E. Venturino, Numerical integration of sparsely sampled data, Simulat. Modell. Pract. Theory, 19 (2011) 1860–1872.

[33] M.R. Binns, J.P. Nyrop, W. Werf, Sampling and Monitoring in Crop Protection: the Theoretical Basis for Developing Practical Decision Guides, CABI, 2000.

[34] P.M. Davis, Statistics for describing populations, in: L.P. Pedigo, G.D. Buntin, (Eds.), Handbook of Sampling Methods for Arthropods in Agriculture, CRC Press (1994), pp.33–54.

[35] D. Belkić, K. Belkić, Signal Processing in Magnetic Resonance Spectroscopy with Biomedical Applications, CRC Press, 2010.

19