The software design of Gridap: a Finite Element package based on the Julia JIT compiler

FRANCESC VERDUGO1,* AND SANTIAGO BADIA1,2

ABSTRACT. We present the software design of Gridap, a novel finite element library written exclusively in the Julia programming language, which is being used by several research groups world-wide to simulate complex physical phenomena such as magnetohydrodynamics, photonics, weather modeling, non-linear solid mechanics, and fluid-structure interaction problems. The library provides a feature-rich set of discretization techniques for the numerical approximation of a wide range of mathematical models governed by Partial Differential Equations (PDEs), including linear, nonlinear, single-field, and multi-field equations. An expressive API allows users to define PDEs in weak form by a syntax close to the mathematical notation. While this is also available in previous frameworks, the main novelty of Gridap is that it implements this API without introducing a domain-specific language plus a compiler of variational forms. Instead, it leverages the Julia just-in-time compiler to build efficient code, specialized for the concrete problem at hand. As a result, there is no need to use different languages for the computational back-end and the user front-end anymore, thus eliminating the so-called two-language problem. Gridap also provides a low-level API that is modular and extensible via the multiple-dispatch paradigm of Julia and provides easy access to the main building blocks of the library if required. The main contribution of this paper is the detailed presentation of the novel software abstractions behind the Gridap design that leverages the new software possibilities provided by the Julia language. The second main contribution of the article is a performance comparison against FEniCS. We measure CPU times needed to assemble discrete systems of linear equations for different problem types and show that the performance of Gridap is comparable to FEniCS, demonstrating that the new software design does not compromise performance. Gridap is freely available at Github (github.com/gridap/Gridap.jl) and distributed under an MIT license.

Keywords: Mathematical Software, Finite Elements, Partial Differential Equations, Julia Programming Language

Program summary

• Program title: Gridap.jl (version 0.16)
• Developer’s repository link: https://github.com/gridap/Gridap.jl
• Licensing provisions: MIT license (MIT)
• Programming language: Julia
• Supplementary material: Source code of the Listings presented in this paper. Each Listing below indicates the name of its corresponding source file.
• Nature of problem: Computational simulation of a broad range of application problems governed by partial differential equations including linear, nonlinear, single field, and multi-physics problems. Gridap is currently being used by several research groups world-wide to simulate complex physical phenomena such as magnetohydrodynamics, photonics, weather modeling, non-linear solid mechanics, and fluid-structure interaction problems.
• Solution method: Arbitrary-order grad-, curl-, and div-conforming finite elements on n-cube and n-simplex meshes. Continuous and Discontinuous Galerkin methods. Newton–Raphson linearization. Krylov subspace iterative solvers. Sparse direct solvers.

1. INTRODUCTION

Mathematical models based on Partial Differential Equations (PDEs) have gained popularity thanks to the computational power increase unlocked by modern computers, which has pushed forward the complexity of the problems engineers and scientists can simulate. However, computational power alone is not enough. Recent advances in numerical discretizations, e.g., Finite Element (FE) methods, and efficient implementation
are required to efficiently and accurately solve PDEs in the most complex scenarios. Besides, domain-specific researchers and practitioners that need to simulate challenging applications in their fields are not necessarily programming and numerical experts. Thus, there is a need for software packages that provide high-performance state-of-the-art simulations tools with simple user interfaces. As a consequence, a myriad of FE frameworks have been proposed, from libraries designed to solve a narrow set of problems with specific computational techniques to general-purpose FE packages that provide a rich set of discretization methods and can solve a wide spectrum of PDEs.

The development of numerical frameworks is a huge endeavor. Since numerical techniques share many software components, it is wise to leverage this effort by broadening the scope without incurring performance hits. However, the design and development of flexible and extensible general-purpose FE packages is not an obvious task. Several software designs have been proposed with this goal in mind. On the one hand, FE libraries like Deal.II [1], FEMPAR [2], and MoFEM [3] provide the building blocks that are common in different types of FE solvers, allowing users to combine them for their particular needs. These libraries are usually modular and extensible via an object-oriented design that provides flexibility to accommodate new scenarios. However, working with these libraries involves a steep learning curve and can be tedious. These libraries are written in static languages like C/C++ or Fortran to provide performance. The user needs to write a considerable amount of glue code to put all pieces together (e.g., one typically must write the FE assembly loop by hand). The Application Programming Interface (API) is usually also in a compiled language to eliminate inter-language boundaries when combining the underlying tools, which further worsens the user experience.

An alternative approach that aims at solving these pitfalls are FE frameworks like FEniCS/DOLFIN [4], Firedrake [5], and FreeFEM [6]. These libraries are written in compiled languages to be efficient but allow the user to define the FE problem with a high-level Domain-Specific Language (DSL) that is compact and mimics the underlying mathematical definition of the weak form of the PDE. The code takes the high-level definition of the problem and automatically generates an efficient FE code, e.g., releasing the user from writing for loops in the FE assembly. Moreover, this high-level DSL is usually available in a dynamic language, typically Python, facilitating the usage of the library and its coupling with other software components from, e.g., the Python ecosystem. Unfortunately, this approach has also its drawbacks. Performance is achieved by a compiler of variational forms, e.g., FFC [7]. It generates efficient low-level code in C/C++ that is specialized for the concrete problem defined via the high-level API. This approach is affected by a two-language (or even three-language, counting the DSL) barrier. The user workspace is limited, and its extension is not obvious. This introduces rigidity when it comes to extending the library since form compilers are sophisticated systems that are not usually designed to be extended by average users.

In this paper, we present the software design of Gridap [8], a novel, free, and open-source FE library distributed under a MIT license that aims at solving the drawbacks of previous packages while keeping their strengths. That is, we aim at combining the extensibility and composability of libraries like Deal.II with the ease of use and compact high-level syntax of frameworks like FEniCS without compromising run-time performance. The novel Gridap software design leverages the capabilities of the emerging Julia programming language [9]. Julia provides the performance of compiled computer languages like C/C++ or Fortran and the productivity of interpreted ones like Python or MATLAB. As a result, there is no need to use different languages for the computational back-end and the user front-end anymore, thus eliminating the so-called two-language problem. This is accomplished by the Julia just-in-time (JIT) compiler, which parses high-level code (similar to Python and MATLAB) and generates efficient native machine code that is specialized for the particular types encountered at run-time, thus being as efficient as statically compiled code implemented in C/C++ or Fortran. The main design goal of Gridap is to exploit the Julia JIT compiler to construct a modern FE library that is easy to use and maintain and is also efficient. In particular, we use the JIT compiler to generate problem-specific native machine code as an alternative to developing and maintaining a compiler of variational forms like FFC at the package level. The result is a library with a compact high-level API similar to FEniCS but much easier to maintain and extend. Gridap also provides a modular low-level API that gives access to the main building blocks of the library if required, thus providing expressive high-level and a low-level interface at the same time. This low-level API is also coded in Julia. It is accessible to users, which can easily extend the library via the multiple-dispatching paradigm in Julia. In contrast, the extension of deal.II, FEMPAR or FEniCS requires modifying the core of the library, written in C/C++ or Fortran. Gridap can also be easily coupled with other packages from the Julia ecosystem, e.g., for optimization [10], ordinary differential equations [11], machine learning [12], automatic differentiation [13], or scientific visualization [14]. Being completely written in Julia, Gridap solvers can be automatically differentiated using automatic differentiation tools to compute, e.g., sensitivities of solvers to parameters.
In any case, Gridap is not a simple Julia translation of an existing FE code. There are other more standard FE libraries written in Julia, see, e.g., FinEtools [15], JuliaFEM [16], or Ferrite [17], whose interface is inspired by Deal.II. In contrast, Gridap relies on a novel representation of data structures associated with the cells in the computational mesh. The user can manipulate cell-wise data without the need to explicitly iterate over the cells as required in other FE packages. It leads to a much more compact and expressive interface while keeping memory requirements similar to traditional codes. Quantities are not stored for all mesh cells simultaneously. Instead, Gridap relies on so-called lazy arrays that compute their entries on demand. Combining these ideas with the Julia JIT compilation, Gridap provides a high-level API that resembles the whiteboard mathematical statement of weak forms of PDEs.

The main contribution of this paper is the detailed presentation of the novel software abstractions behind the Gridap design (as of version 0.16). Due to length constraints, we focus on the low-level and high-level tools used to build elemental matrices and vectors, which is the part of the library that follows a more unconventional software approach and is the cornerstone of the high-level API close to the mathematical notation. We present both single-field and multi-field problems. Note however that Gridap is a feature-rich FE library that supports a large set of discretization techniques that go beyond the ones covered in this paper (including different types of continuous and Discontinuous Galerkin (DG) methods) and supports a wide range of PDEs types, including, e.g., nonlinear and complex-variable equations. The second main contribution of the article is a performance comparison against FEniCS. We measure CPU times needed to assemble discrete systems of linear equations for different problem types and show that the performance of Gridap is comparable to FEniCS, demonstrating that the new software design of Gridap does not compromise performance.

The next sections of the paper are structured as follows. Sect. 2 presents the low-level software components of the Gridap project. At the end of this section, we show how one can combine these low-level tools to compute elemental stiffness matrices for all cells of the FE mesh without explicitly writing for-loops. Then, in Sect. 3, we show how the high-level interface wraps low-level tools, which allows one to manipulate quantities in the FE computation in an even more convenient way. Sect. 4 is about the extension to multi-field PDEs. Sect. 5 contains numerical experiments that show the performance of the library. We draw some conclusions in Sect. 6.

2. THE COMPUTATIONAL BACK-END

2.1. Overview. We present the architecture of the computational back-end of Gridap with a bottom-up approach, starting from the lowest-level functionality and building on top of it. The back-end is structured in terms of several abstractions represented by abstract types, which are the Julia counterparts of classes with virtual methods in C++. These abstractions make the code modular and extensible since users and developers can implement new specializations of these abstract interfaces if needed. The main abstract types are listed in Table 1. Note that this part of the code relies on very few abstractions. The complexity of a FE solver is achieved by combining these basic building blocks. For instance, the local shape functions in a reference cell are represented as vectors of fields, thus combining the interfaces of the abstract types AbstractArray and Field, instead of introducing a new abstraction for this particular case. Note that the abstract types in Table 1 stem naturally from fundamental mathematical concepts and are not introduced due to implementation artifacts. A minor exception to this rule is the abstract type Map, which is introduced mainly for performance reasons. Map is the fundamental abstraction that allows us to reuse computations and avoid spurious dynamic allocations when iterating over the cells of the mesh. This type could only be replaced by Function if performance was not a concern. In the following subsections, we describe how we use and extend the AbstractArray interface and how we introduce important performance optimizations via the Map abstract type. We also detail the other abstract types listed in Table 1. Finally, we show how to build core components of a FE solver combining these basic ingredients, e.g., to implement the local interpolation spaces or elemental matrices associated with a PDE weak form and the cells of a FE mesh.

2.2. Julia arrays. Julia provides an excellent framework to work with multi-dimensional arrays similar to NumPy in Python. The main difference is that the user can write arbitrary for-loops over these arrays without performance limitations. Explicit loops are usually slightly faster than using broadcast operations, which might be surprising to Python programmers. Julia provides the built-in array type Array{T,N}, which is a parametric type with type parameters T and N, being the type of the elements stored in the array and the dimension of the array, respectively. In particular, vectors are represented by the type Array{T,1} and matrices by the type Array{T,2}, or by the aliases Vector{T} and Matrix{T} respectively. If the element type T has values contained in a chunk of bits (i.e., an isbits type in Julia notation), e.g., Int32 or Float64, the internal memory layout of the built-in array type is a one-dimensional contiguous section of memory. The length is equal to the number of bits needed to represent one instance of type T times the number of elements in the array. Julia sweeps the array entries in row-major order to represent multi-dimensional arrays with a 1D section of memory, like Fortran. The default Julia array type is a reference to a heap-allocated section of memory similar to a dynamically allocated array in C/C++ or Fortran. In
particular, this means that a Julia array of Julia arrays is a reference to a vector of references pointing to the different memory locations storing the local arrays, similar to an array of pointers in C/C++.

In addition to the built-in array type \texttt{Array(T,N)}, Julia provides the abstract type \texttt{AbstractArray{T,N}} that represents types that behave like an array but are not necessarily the built-in array type. It allows one to implement new array types optimized for particular scenarios to achieve better performance. These new concrete types of \texttt{AbstractArray{T,N}} specialize the abstract type by implementing its interface. To illustrate this, Listing 1 shows how to implement a new vector type for the particular case of vectors with all entries having the same value. Formally, we define the element $v_i$ of the constant vector as $v_i = \hat{v}$ for all $i = 1, \ldots, |v|$, where $\hat{v}$ is the unique value stored in the array and $|v|$ denotes the array length. This specialized vector implementation is less memory consuming than the default one.

### Listing 1. Implementing a custom array type in Julia.

```
# Type definition
struct ConstantVector{T} <: AbstractArray{T,1}
    value::T
end

# Define length at each direction
length::Int

# Define indexing
function Base.getindex(a::ConstantVector,i::Integer)
    if !(1<=i && i<=a.length )
        throw(BoundsError(a,i))
    end
    return a.value
end

@boundscheck function sum(a::AbstractArray{T}) where T
    r = zero(T)
    for i in eachindex(a)
        r += a[i]
    end
    return r
end

# Specialized sum
sum(a::ConstantVector) = a.value*a.length

# Run the code
a = rand(100)
b = ConstantVector(3.14,100)
sum(a) # This calls the generic sum
sum(b) # This calls the specialized sum
```

One of the most powerful features of Julia is that one can write \texttt{generic} code for any sub-type implementing the interface of an abstract type without suffering from the performance hit associated with polymorphic methods in object-oriented languages like C++, Fortran 2003, or Python. It is possible thanks to the Julia JIT compiler, which converts just before run-time the generic code into efficient machine code specialized for the concrete array types used in the computation. For instance, we implement a vanilla version of the built-in function \texttt{sum} that sums all the entries in the given array in line 21 of Listing 1. When this function is called for a concrete array type, the Julia JIT compiler will compile a specialized version of this function for the concrete types it encounters in the argument list, resulting in efficient code. In addition, it is possible to help the JIT compiler to generate even more efficient machine code by providing specialized function definitions using so-called \textit{multiple dispatch}. For instance, the sum of vector entries does
not require looping over all entries when they are all the same. Thus, the generic method is not as efficient as it could be even if it is specialized by the compiler. The developer can help the compiler by defining an optimized version that is dispatched for this particular case as shown in line 30 of Listing 1. This strategy is called multi-dispatch since one can specialize any number of arguments simultaneously in the function definition.

2.3. Main custom array types used in the Finite Element code. Apart from the built-in Julia arrays, Gridap considers several other array types to leverage optimizations that apply in FE computations. These arrays represent cell-wise data efficiently and conveniently. Different algorithms in the code are implemented in terms of these arrays. Custom array types allow one to manipulate cell-wise data independently of how it lays in memory. For instance, one can write a generic code in terms of an array representing the quadrature points for all mesh cells without exposing how these points are stored. The same generic code can work in different scenarios, e.g., for meshes with a single reference element, where one can use a constant array similar to the one presented in Listing 1 to store the quadrature points, and also for meshes with more than one element type that will require the usage of a more sophisticated array type, i.e., the so-called \texttt{CompressedArray} introduced later. We avoid code duplication by using generic code and custom array types while keeping performance by using the Julia JIT compiler to specialize the emitted machine code for each case.

Gridap uses different array types in its implementation being \texttt{Fill}, \texttt{CompressedArray}, \texttt{Table}, \texttt{CachedArray}, and \texttt{LazyArray} the most important ones. The type \texttt{Fill} is a curated implementation of a constant multi-dimensional array provided in the Julia package FillArrays\cite{18}. \texttt{Fill} is similar to the constant vector presented in Listing 1, but it comes with some improvements like support for the multi-dimensional case. \texttt{CompressedArray} is a type implementing an array \( v \), whose entry \( v_j \) is defined as \( v_j = \hat{v}_j \) with \( j = t_i \). The quantity \( \hat{v}_j \) represents a small vector (small with respect to \( v \)) containing the (unique) values of \( v \) and \( t_i \) is an indirection array that for each index \( i \) returns the index \( j \) in the small vector \( \hat{v}_j \). E.g., this array type represents quantities defined in reference elements for FE spaces with multiple element types. In this context, \( \hat{v}_j \) can be interpreted as the reference value associated with element type \( j \). \( t_i \) can be interpreted as the element type for a cell id \( i \). Note that the type \texttt{CompressedArray} only stores in memory the vector of reference values \( \hat{v}_j \) and the vector of cell types \( t_j \). The next array type is \texttt{Table}, which represents a vector of vectors in a memory-efficient way. E.g., this array stores cell connectivities (a vector of node ids for each cell id) in a general unstructured grid with potentially different element types (i.e., different number of local nodes per cell). This type adopts the usual representation of a jagged array, i.e., vector of vectors, using a contiguous vector of values \( a \) and an auxiliary vector of pointers \( p \). The entry \( v_i, i = 1, \ldots, |v| \), of a \texttt{Table} is a small vector consisting in a consecutive portion of the vector \( a \) for indices delimited by the vector of pointers \( p \), namely \( v_i = (a_r, a_{r+1}, \ldots, a_s) \) with \( r \leq p_i \) and \( s = p_{i+1} - 1 \). We use this data layout instead of a built-in Julia vector of Julia vectors since the data of this latter type is not stored consecutively in memory, thus not leveraging memory locality. In contrast, the data stored in the custom array type \texttt{Table} is stored in memory consecutively, which contributes to more efficient memory access. Another important array type is \texttt{CachedArray} which implements a multi-dimensional array that can change its size efficiently. E.g., this array type represents elemental matrices and vectors whose size can change between cells for FE spaces with multiple element types. The size of an instance of \texttt{CachedArray} changes by calling \texttt{setsize!(a,new_size)}; it sets the size of \( a \) to match the tuple of integers \texttt{new_size}. The implementation of \texttt{setsize!} allocates a new array under-the-hood each time a new size is requested. The allocated arrays are cached in this process so that it is not needed to allocate a new array if one with the right size was already cached. This caching strategy is designed having FE computations in mind since the number of different array sizes one needs for an instance of \texttt{CachedArray} is usually the same as the number of different element types in the mesh. Thus, the number of dynamic allocations when iterating over a mesh is proportional to the number of element types but independent of the total number of cells. The last array type, \texttt{LazyArray}, deserves a more in-depth discussion and will be presented in Sect. 2.5.

2.4. Avoiding heap allocations when indexing arrays of arrays. When implementing custom array types, like \texttt{Table}, reducing the number of objects dynamically allocated in the heap is challenging. If the value returned when indexing an array \( a \) at index \( i \), namely \( a[i] \), is an isbits object (e.g., an instance of Int or Float64), the result will be efficiently allocated in the stack of the calling function and the programmer does not need to take any extra action related with memory allocation. However, if \( a[i] \) is not an isbits object, e.g., if \( a[i] \) is an instance of a Julia array\(\langle T, N, \rangle\), the result is usually allocated in the heap by the Julia compiler. In this case, dynamically allocating objects when fetching entries \( a[i] \) while iterating over large arrays is not efficient. One needs to adopt a mechanism to cut down the number of dynamic allocations. A naive solution is to pre-allocate the result of \( a[i] \) and store it in a global variable. Then, when indexing \( a[i] \) for a particular \( i \), one can fill the pre-allocated object and return it without allocating new memory dynamically. However, this naive solution is not the one adopted in Gridap because it is not thread-safe. Several threads cannot use the same pre-allocated output without incurring in race conditions.

To solve the problem of memory allocations in a thread-safe manner, Gridap introduces its own API to iterate over arrays, see Listing 2. This API defines two new functions, namely \texttt{array\_cache} and \texttt{getindex!}, to iterate arrays efficiently. Before starting the loop, one calls \texttt{array\_cache} to generate a cache object holding arbitrary data. This cache is reused each time the array is indexed. Function \texttt{getindex!} is a generalization of the built-in function \texttt{Base.getindex}, which accepts a cache object in its first argument that can speed up the computation of the result. Custom array types overload \texttt{array\_cache} and \texttt{getindex!} to establish which data is reused and how it is reused at each iteration. For instance, the cache object can store a pre-allocated object that is filled and returned when calling \texttt{getindex!}, thus avoiding allocating new objects within the loop. Note that this strategy is thread-safe; one
can generate an independent cache for each thread before starting the loop (see, e.g., the last loop in Listing 2), thus avoiding race conditions.

### Listing 2. Demonstration of the Gridap array API that allows to iterate over arrays without allocating memory in the heap at each iteration.

The new Gridap API for looping over arrays is general. It can be used to iterate over any array type specializing AbstractArray, including (but not limited to) the array types defined in Gridap. By default, array_cache does nothing and getindex! calls the built-in function Base.getindex. Thus, using the Gridap API to iterate arrays that do not define the functions array_cache and getindex! is equivalent to iterating them with the standard Julia API. It means that one does not need to define array_cache and getindex! for array types that do not need to pre-allocate any output or reuse any data.

2.5. Mapping large arrays under memory constraints. Many PDE discretization techniques, e.g., FE methods, can be expressed as complex cell-wise operations between fields, linear functionals, scalars and tensors. Here, we provide expressive and efficient machinery for these cell-wise operations. We use the FE method as a driving application. However, the software abstractions are general in scope and apply to other scenarios with similar requirements.

Like other programming languages, Julia provides the function map, which transforms arrays and other collections in a very convenient way via a high-level API without the need of explicitly writing for-loops. When applied to arrays, the map function returns an array \( r \), whose entries are computed by applying a given function \( f \) entry-by-entry to a number \( n \) of input arrays \( a^1, \ldots, a^n \), namely \( r_i = f(a^1_i, \ldots, a^n_i) \). The syntax provided by the map function is especially convenient to implement FE codes since it allows one to represent quantities defined on the cells of a FE mesh using arrays and transform them without for-loops. However, the default implementation of map provided by Julia has some limitations related to memory requirements. The array \( r \) returned by the map function is an intrinsic Julia array by default, which implies that all entries \( r_i \) are stored in memory for all indices \( i \) simultaneously. It becomes a serious bottleneck when manipulating large arrays, e.g., cell-wise arrays for a FE mesh with millions of cells.

Gridap provides the alternative function lazy_map, which solves this memory bottleneck while exploiting the powerful high-level syntax provided by the map function. Both map and lazy_map have the same syntax. The main benefit of using lazy_map over the default map function is that the former returns instances of a custom array type called LazyArray instead of a built-in Julia Array; see a simplified implementation of lazy_map and LazyArray in Listing 3. LazyArray does not store the result \( r_i \) for all indices \( i \) simultaneously. Instead, it stores references to function \( f \) and the input arrays \( a^1, \ldots, a^n \), as shown in Listing 3. With this information, it computes on-the-fly the value \( r_i \) when the index \( i \) is requested by overloading the Base.getindex function. We say that this array type is lazy because it delays the computation of \( r_i \) till its consumption. One potential drawback of this approach is that \( r_i \) must be computed each time one gets a given index. This approach is better than pre-computing all entries of large arrays in terms of memory requirements and allocations (see Sect. 2.6 for more details about the latter). In addition, fetching pre-computed value from main memory comes with a cost and computing the value on-demand can be faster for some operations. Anyway, one can still use the default Julia map when it can potentially be faster than lazy_map, e.g., for small arrays. In Gridap, lazy_map is used with arrays of length proportional to the number of the cells in the mesh. map is used otherwise, e.g., in operations at the reference cell level. Further optimizations can be considered when array entries are going to be accessed multiple times when positioned in a cell but are not covered here due to length restrictions.

To illustrate the usage of lazy_map, we consider the example in Listing 4. Here, \( \mathbf{a} \) and \( \mathbf{b} \) are two standard Julia arrays of random values. We generate a new array \( \mathbf{c} \) by summing the entries of arrays \( \mathbf{a} \) and \( \mathbf{b} \) via the lazy_map function. Array \( \mathbf{c} \) is an instance of the LazyArray type. It contains a reference to the operation + and the input arrays \( \mathbf{a} \) and \( \mathbf{b} \). This information can be represented graphically as a small operation tree (see Fig. 1(b)). The operation tree can be visualized using the helper function print_op_tree, which prints a representation of the tree in plain text format into the standard output. See the output of this function on array \( \mathbf{c} \) in Fig. 1(A).
From the output, it is clear that \( c \) is of type \texttt{LazyArray}. It has two inputs of type \texttt{Vector\{Float64\}} combined via the operator \(+\). A new lazy array \( d \) with a more complex operation tree is generated by multiplying the entries of the lazy array \( c \) with the entries of the array \( a \). In this case, \( d \) stores a reference to operation \(*\) and a reference to arrays \( c \) and \( b \). Thus, the operation three associated with \( d \) contains part of the operation tree of array \( c \), see Fig. 1(c) and Fig. 1(b). It illustrates how one can create arbitrarily complex \texttt{LazyArray} objects by nesting calls to the \texttt{lazy_map} function. In some sense, the type \texttt{LazyArray} is related with types encoding symbolic operation trees in FE codes based on symbolic DSLs. However, \texttt{LazyArray} is not just a \textit{symbolic} type, but also a \textit{numeric} one, since it can be indexed to recover the numerical values represented by the array. This ambivalent nature of \texttt{LazyArray} is one of the most powerful features of Gridap. It allows one to implement a high-level user API that resembles symbolic FE frameworks and, at the same time, the objects involved have a numerical value that can be directly accessed without calling a C/C++ code generator.

**LISTING 3.** Simplified implementation of \texttt{lazy_map} and \texttt{LazyArray}.

**LISTING 4.** Basic usage of the \texttt{lazy_map} function provided by Gridap.

![Diagram of operation trees](image)

**Figure 1.** Output of function \texttt{print_op_tree} and graphical representation of the corresponding operation trees for arrays \( c \) and \( d \) in Listing 4.

2.6. **Efficient elemental operations.** Let us take a deeper look into \texttt{lazy_map}. The performance of \texttt{lazy_map} is compromised when the output or intermediate computations of \( f \) require dynamic allocations. Elemental operations in \texttt{lazy_map} can potentially be called many times and should be performant. Gridap introduces the abstract type
Map to solve this problem. Map represents functions or any other callable object able to pre-compute and reuse a cache to avoid dynamic allocation each time they are called. Specializations of the Map type have to overload two new functions, namely return_cache and evaluate!, which generate the cache and evaluate the object by reusing this cache, respectively.

Let us consider the example in Listing 5 to illustrate the usage of these two functions. We start by creating two random Julia arrays \( a \) and \( b \). Then, we create \( f \) (an instance of the Broadcasting type, a specialization of Map) that represents a broadcast operation over some input arguments. Like in Python, a broadcast operation in Julia is an operation that is performed element-wise on some input arrays with implicit expansion of singleton axes (see the Julia documentation of further details). Calling an instance of this type is equivalent to broadcast the given operation on the input arrays. However, the main difference is that type Broadcasting implements the interface of the abstract type Map and thus, one can use function return_cache and evaluate! to pre-allocate and reuse the array resulting from the broadcast operation. Listing 5 shows the basic usage of these functions. Function return_cache computes a cache object containing a pre-allocated array storing the result of the broadcast operation. Then, evaluate! uses this cache to compute the broadcast operation in-place without allocating the result.

```
using Gridap
using Gridap.Arrays

# Build two random arrays
a = rand(4)
b = rand(4)

# Build two other input arrays
a2 = 2*a
b2 = -5*b

# Call the Map object by pre-allocating a cache
f = Broadcasting("*")

@assert f(a,b) == a .* b

# Call the Map object by pre-allocating a cache

# Build two random arrays
C = lazy_map(f,A,B)

# Build two arrays of arrays
A = [ rand(4) for k in 1:10 ]
B = [ rand(4) for k in 1:10 ]

# Build two other input arrays
c = evaluate!(cache,f,a2,b2) # no allocations here
c2 = evaluate!(cache,f,a,b) # no allocations here

# Loop over C by pre-allocating a cache
for i in 1:length(C)
    Ci = getindex!(cache,C,i) # no allocations here
end
```

### Listing 5. Usage example of functions return_cache and evaluate!.

One of the main applications of functions return_cache and evaluate! is to efficiently implement the LazyArray type. By using these methods, it is possible to iterate an instance of LazyArray without allocating memory at each step (see, e.g., the last loop in Listing 5). The implementation of array_cache for the LazyArray type (see Listing 3) builds all the cache objects needed to iterate an instance of LazyArray, i.e., the cache objects needed to index the input arrays, and the cache needed to evaluate the elemental operation \( f \). Then, the specialization of getindex! for the LazyArray type (see also Listing 3) uses these cache objects to compute a particular element of the array without allocating memory. See how the input arrays are indexed by using their cache objects and how the elemental operation \( f \) is called by using its cache within the definition of getindex! in Listing 3. In summary, the LazyArray type does not only avoid allocating large arrays when mapping elemental operations on input arrays but also allows one to efficiently iterate on the result by enabling to reuse data at each iteration. This feature is at the heart of the efficient FE implementation in Gridap.

#### 2.7. Representing fields.

The actual implementation of FE methods in Gridap starts with the definition of an interface representing fields. In Gridap, fields are understood as functions \( f : D \subset \mathbb{R}^n \rightarrow R \) that take vectors in a physical domain \( D \subset \mathbb{R}^n \), with \( 0 < n \leq d \), and return values in a range \( R \). In this context, we usually refer to elements of \( D \subset \mathbb{R}^n \) as evaluation points since these are the points in the physical domain where fields are evaluated. Values in \( R \) can be real numbers \( \mathbb{R} \), complex numbers \( \mathbb{C} \) or arbitrary order tensors with real or complex components, e.g., \( R \subset \mathbb{C}^d \). In the code, the value of a field can be represented with any type specializing the abstract type Number defined in Julia. Note that even though vector and tensor values are not numbers in a strict mathematical sense, we implement them as sub-types of Number for convenience. In particular, one can use the Julia types Float64 and ComplexF64 to represent the value of real-valued and complex-valued fields respectively, or the Gridap types VectorValue and TensorValue for vector-valued or second-order tensor-valued fields respectively (further details on the implementation of these tensor types will be provided later in Sect. 2.8). One can even consider more sophisticated number types, like dual numbers to forward-mode automatic differentiation [19] to automatically differentiate Gridap solvers in forward-mode [19] using Julia packages like ForwardDiff [13].

Different types of fields involved in a FE computation, such as shape functions, FE interpolations or manufactured solutions, are implemented as specializations of the abstract type Field. The API of this type is a concretization of the API of the abstract type Map. The Map type represents a general transformation between any input and
output arguments, whereas Field stands for the particular case in which the input is a point in \( \mathbb{R}^n \) (an instance of VectorValue) and the output is a number or a vector/tensor (an instance of some specialization of Number in general). The abstract type Field is defined as a sub-type of Map and, in consequence, all the generic code implemented for Map objects can also be used with Field instances.

In order to implement a new Field type one needs to specialize functions return_cache and evaluate!, just as previously explained for Map. With these methods, it is possible to use a set of generic methods implemented for Field such as algebraic operations, function composition, differential operators, etc. (See Sect. 2.9 for further details). In order to illustrate how a new field type is implemented, Listing 6 contains a simplified implementation of the Gridap type ConstantField that represents a field with a constant value, i.e. \( f(x) \) is defined in this case as \( f(x) = \hat{f} \) for some number \( \hat{f} \) and for any evaluation point \( x \). This specialization optimizes several computations in the code. E.g., the Jacobian of the geometrical map in linear simplicial meshes is a ConstantField in Gridap, which makes the evaluation of the Jacobian at the integration points very efficient. ConstantField is a specialization of Field that stores the value of the field in the variable value, returned each time the field is evaluated. The last argument in the definition of evaluate! is the point \( x \) where the field is to be evaluated. Object \( x \) is an instance of Point, which is an alias of VectorValue used to emphasize that vector \( x \) is interpreted as an evaluation point in this context. The implementation of this function is trivial for this particular case since it just needs to return the value stored within the given instance of ConstantField. However, other field specializations will consider the evaluation point to compute the result. Note also that the cache object is not used in this simple example since the result is already pre-computed, but other field implementations can also overload function return_cache and make use of the cache object if needed as explained before for the Map abstract type. Once the evaluate! function is defined, one can, e.g., evaluate the resulting field objects using function notation or evaluating the field at an array of points, as demonstrated in the last part of Listing 6. In the latter case, the result is a Julia vector containing the field values at all points that needs to be heap-allocated. Thus, it is recommended to generate a cache object and then evaluate the field without memory allocations. In order to implement a new Field type one needs to specialize functions return_cache and evaluate!, just as previously explained for Map. Once these methods are implemented, it is possible to use a set of generic methods implemented for Field such as algebraic operations, function composition, differential operators, etc (see Sect. 2.9 for further details). In order to illustrate how a new field type is implemented, Listing 6 contains a simplified implementation of the Gridap type ConstantField that represents a field with a constant value, i.e. \( f(x) \) is defined in this case as \( f(x) = \hat{f} \) for some number \( \hat{f} \) and for any evaluation point \( x \). This specialization is used to optimize several computations in the code, e.g., the Jacobian of the geometrical map in meshes of linear simplices is represented in Gridap with a ConstantField, which makes the evaluation of the Jacobian at the integration points very efficient. The struct ConstantField is defined as a specialization of Field and it stores the value of the field in the variable value, which can be returned each time the field is evaluated. In the definition of function evaluate!, the last argument is the point \( x \), where the field is to be evaluated. Object \( x \) is an instance of Point, which is an alias of VectorValue used to emphasize that vector \( x \) is interpreted as an evaluation point in this context. The implementation of this function is trivial for this particular case, since it just needs to return the value stored within the given instance of ConstantField. However, other field specializations will consider the evaluation point to compute the result. Note also that the cache object is not used in this simple example since the result is already pre-computed, but other field implementations can also overload function return_cache and make use of the cache object if needed as explained before for the Map abstract type. Once the evaluate! function is defined, one can e.g., evaluate the resulting field objects using function notation or evaluating the field at an array of points, as demonstrated in the last part of Listing 6. In the latter case, the result is a Julia vector containing the value of the field at all evaluation points. Since the result is a vector that needs to be heap-allocated, it is recommended to generate a cache object and then evaluating the field at a given vector of points as many times as needed without allocating extra memory.

2.8. Tensor values and efficient tensor algebra. Gridap provides a family of concrete types representing tensors of different orders. First order tensors (i.e., vectors) are represented with instances of the type VectorValue{D,T}, where the type parameters D and T are the vector length and component type respectively. Similarly, second-order tensors are represented with instances of the type TensorValue{D1,D2,T}, where the size of the tensor is D1 times D2 and the component type is T. Other related types include symmetric second-order tensors (e.g., stress and strain in mechanical computations), third-order tensors (e.g., the gradient of vector-valued fields) and symmetric fourth-order tensors (e.g., constitutive relations), but they are not detailed here for the sake of brevity. All these concrete tensor types inherit from the abstract type MultiValue, which serves as the umbrella that covers all these cases. Gridap defines several tensor algebra operations for tensor-values, like scaling, addition, subtraction, dot product, single and double contractions, inner products, and outer products. As an example, the first part of Listing 7 shows how to perform some operations with first and second-order tensors.\(^1\)

\(^1\)LaTeX symbols can be easily written in the Julia REPL by typing the corresponding LaTeX macro and pressing the tabulator. E.g., \( \alpha \) is written as \texttt{\textbackslash alpha + Tab}. Several text editors and IDEs (like Vim, Emacs and VSCode) provide Julia extensions to write latex symbols this way.
Listing 6. Implementation and usage of a new field type in Gridap.

the tensors in the type parameters of the corresponding types. By doing so, the size of the tensors is known at compile
time instead of at run time, which makes it possible to introduce some important optimizations. In particular, it allows
one to specialize operations for particular tensor sizes and unroll for-loops to achieve performance. For instance, the
dot product of two instances of `VectorValue{2,T}` (i.e., vectors of 2 components) can be implemented efficiently for
this particular size by explicitly writing the result by loop unrolling (see line 20 of Listing 7). Even though the number
different tensor sizes in a FE simulation is bounded (the number of space dimensions is usually $d \leq 4$), writing all
specializations in this way can be tedious. To solve this problem, Julia provides an advanced feature called generated
functions, which are functions with a special syntax. In a conventional Julia function, the function body includes
operations that compute and return a value at run-time. In contrast, the function body of a generated function
contains commands to generate and return a piece of code that, once compiled, computes and returns the value of
the function at run time. A generated function can be interpreted as a pre-compilation step in which the code to be
compiled is generated using information about the types of the input arguments. For instance, in line 26 of Listing 7
we show how to use a generated function to define the dot product of two instances of `VectorValue{D,T}` for a generic
$D$. The body of the function generates the code to be compiled for a particular $D$. Here, this is done by generating a
string containing the code and parsing it, which returns an `Expr` object that is the format used in Julia to represent
Julia code. Note that the routines that generate the code are invoked once, right before the function compilation,
whereas the routines within the generated code are called in run-time. As a result, the for-loop that generates the
code is executed once in the compilation stage and the code that is called at run-time has no for-loops, thus achieving
performance.

Listing 7. Overview of tensor operations available in Gridap and their efficient implementation.

2.9. Algebraic and differential operations on fields. Gridap does not only allow one to operate on values but
also to operate on fields directly. Let us assume that $\odot$ denotes some operation, e.g., $+$ or any other binary operation
over fields $f$ and $g$ defined on the same domain. In Gridap, it is possible to create a new field $f \odot g$, that represents
the result of applying operation ◦ to f and g, namely \((f \circ g)(x) = f(x) \circ g(x)\). The result field is an instance of a concrete field type called `OperationField`, which contains references to the given operation and the input fields. When an `OperationField` is evaluated, e.g., with `evaluate!`, the input fields are evaluated first. Next, the operation is applied to the result. One of the main advantages of operating over fields (instead of evaluating the fields first and then operating over the values) is that the result is an instance of a sub-type of `Field`. Thus, one can use it in any generic function implemented for `Field` objects. This feature is heavily used in the library, e.g., in functions performing numerical integration or visualization implemented in terms of the abstract type `Field`, thus reducing code duplication.

### Listing 8. Overview of algebraic and differential operations on fields.

The syntax for operating fields is illustrated in Listing 8. We start the code snippet by creating three instances of `GenericField`. This type is a specialization of the abstract `Field` type used to generate fields from objects that behave like fields, but their type is not a specialization of `Field`. A typical application is to build fields from analytical functions, as shown in this piece of code. Note that the generated fields, \(f\) and \(g\) are scalar-valued, whereas \(h\) is vector-valued (we specifically need a vector-valued field in some of the operations displayed in the example). For basic operations defined in Julia like + one can write operations on fields as they were operations on values. However, user-defined operations must follow an alternative syntax because they cannot be anticipated. In this case, one uses the auxiliary constructor `Operation` that returns an object that can be evaluated to perform the desired operation over the input fields. In addition, the constructor `Operation` allows us to implement the function composition \(f \circ h\) of two fields \(f\) and \(h\), where the range of \(h\) is equal to the domain of \(f\), namely \((f \circ h)(x) = f(h(x))\). The call \(f \circ h\) is simply implemented as `Operation(f)(h)` since \(f\) can be interpreted as a user-defined unary operation on the (vector) value of \(h\). It is also possible to operate fields with instances of other types like numbers or functions. In this case, the objects that are not fields are converted into fields internally before doing the operation. Numbers are converted into `ConstantField` objects and functions into `GenericField` objects. In addition to algebraic operations, Gridap also defines several differential operators like gradient, divergence, and curl. See last lines in Listing 8 for some particular examples. The result of a differential operation on a `Field` object is also a `Field` object that can be used in other parts of the code. When differentiating fields defined from user-defined functions, like the fields \(f\) and \(h\) in Listing 8, Gridap considers the automatic differentiation package ForwardDiff [13] to automatically compute the derivatives under-the-hood. However, it is also possible to bypass the automatic differentiation step if needed by explicitly defining the gradient, as shown in line 32 of Listing 8.

### 2.10. Representing linear functionals.

The final low-level ingredient in the computational back-end is the abstract type `Dof` that represents a linear functional on some space of fields. The name of this type is motivated by the Ciarlet definition of a FE element [20], where the linear functionals defining the shape functions are referred to as DOFs. The abstract type `Dof` represents callable objects that take a `Field` object and return a real number via a linear transformation. The abstract type `Dof` inherits from `Map` and, thus, a `Dof` object can be used in any generic function implemented for `Map`. All the functionality previously presented for the abstract types `AbstractArray`, `Field` and `Dof` is modular and composable and can be combined to construct a myriad of different objects involved in a FE computation. To illustrate this, let us consider the following two practical examples.

### 2.11. Example: reference shape functions.

In this example, we demonstrate a possible way to build the shape functions of a 3-node reference triangle via the low-level API of Gridap and how to use these functions to interpolate...
some given nodal values. To this end, we start from a monomial basis spanning the 1st order \((k = 1)\) 2-variate polynomial space \(P\) on the vector variable \(x\), namely \(M \doteq \{x_1^i x_2^j : e_1 + e_2 \leq k\}\). From these monomials, we obtain the shape functions with the change of basis \(\hat{s}_i \doteq \sum_j A_{ij} m_j\), where \(\hat{s}_i\) is the \(i\)-th shape function, \(m_j\) is the \(j\)-th monomial in \(M\), and \(A\) is the inverse of the change of basis matrix. We compute \(A_{ij} = \hat{l}_i(m_j)\) by introducing a basis \(\{\hat{l}_i : i = 1, \ldots |M|\}\) of the dual space of \(M\), i.e., the basis of DOFs. In this case, we consider the Lagrangian DOF basis associated with the 3-noded reference triangle and \(\hat{l}_i(m_j)\) is simply the evaluation of the monomial \(m_j\) at the coordinate vector of the \(i\)-th node in the reference triangle. We build Lagrangian shape functions for simplicity, but other FE spaces (like Raviart-Thomas or Nédélec) can be constructed similarly.

```
1 using Gridap
2 using Gridap.Fields
3 using Gridap.Polynomials
4 using LinearAlgebra
5 # Evaluation at a vector of points
6 u_at_nodes = evaluate(u,nodes)
7 @assert u_at_nodes == vals
8
9 # Build a vector of monomials
10 filter(e,k) = sum(e) <= k
11 m = MonomialBasis(2)(Float64,1,filter)
12 @assert isa(m,AbstractVector{<:Field})
13 ndofs = length(m)
14 # Lagrangian DoF basis
15 nodes = Point(1/3,1/3)
16 l = LagrangianDofBasis(Float64,nodes)
17 @assert isa(l,AbstractVector{<:Dof})
18 # Interpolation
19 A = evaluate(l,m)
20 change = inv(A)
21 # Change of basis
22 s = linear_combination(change,m)
23 @assert isa(s,AbstractVector{<:Field})
24 # Evaluation at a vector of points
25 s_at_nodes = evaluate(s,nodes)
26 @assert s_at_nodes == Id
27 # Evaluation at a single point
28 x = Point(1/3,1/3)
29 s_at_x = evaluate(s,x)
30 @assert s_at_x == Id
31 # Evaluation at a single point
32 u_at_x = evaluate(u,x)
33 @assert u_at_x == Id
34 # Evaluation at a single point
35 u_at_nodes = evaluate(u,nodes)
36 @assert u_at_nodes == vals
```

**Listing 9.** Construction of the shape functions for a 1st order \(P\) element in 2D and interpolation of some user-defined nodal values via the low-level API of Gridap.

The actual computation of the shape functions is given in Listing 9. We start by building the monomial basis via the constructor `MonomialBasis`, which accepts several inputs. In this case, we chose 1st order 2-variate monomials with values represented by a `Float64` number. The function object passed in the last argument is a filter to select the subset of the tensor-product multivariate space \(Q\) to be included. We define a function `filter` to select only monomials in the \(P\) space; it returns `true` for \(e_1 + e_2 \leq k\) and false otherwise. Note that the inputs \(e\) and \(k\) of function `filter` correspond to the tuple \((e_1, e_2)\) and the value \(k\) respectively. The resulting object \(m\), containing the monomial basis, is a vector of `Field` objects (as checked after its construction) with a field per monomial in the basis. This object can be consumed by the methods defined in the abstract interfaces of `AbstractArray` and `Field` generically. The subsequent lines compute the change of basis that transform the monomials into the shape functions. To this end, we build the Lagrangian DOF basis associated with the nodes of the reference triangle. The resulting object \(l\) is as a vector of `Dof` objects. Each one of those DOFs corresponds to the evaluation at a node in the reference triangle. Then, we compute matrix \(A\) by evaluating the DOF basis \(l\) on the monomials \(m\). The shape functions are constructed with a linear combination of the coefficients in the inverse of matrix \(A\) and the monomial basis \(m\) (see, e.g., [2] for more details). Note that function `linear_combination(x,y)` is just a short-hand for `transpose(x)*y`. The resulting value \(s\) containing the shape functions is also a vector of `Field` objects. It can be consumed using the generic methods associated with `AbstractArray` and `Field`, just like for the monomials. E.g., we evaluate the shape functions at the reference nodal coordinates to check that the resulting matrix is indeed the identity as expected. Note that shape functions (and any `Field` object) can be evaluated at arbitrary points, not only nodes of the reference element or integration points. This is useful for interpolation purposes. E.g., high-order FE fields can be visualized by interpolating them in linear spaces on finer grids. The final lines in the Listing 9 show how one can interpolate some nodal values using the previously computed shape functions. The interpolation is simply built as a linear combination of the nodal values and the shape functions, namely \(u(x) = \sum u_i \hat{s}_i(x)\), where the vector `vals` stores the coefficients \(u_i\). The result \(u\) is a `Field` object that can also be evaluated at arbitrary points. We check that \(u\) is correct by evaluating it at the nodal coordinates and checking that the resulting vector coincides with the vector of nodal values. Note that the code giving support to the low-level API in this example works for any space dimensions. As a result,
this code snippet can be easily modified to compute the shape functions of a simplex in any other number of spatial dimensions.

2.12. Example: cell-wise elemental stiffness matrices. In this second example, we go a step further and illustrate how one can use the low-level API of Gridap to build a vector containing the elemental stiffness matrix for each cell in a computational mesh. In this example, we use a simple 2x2 Cartesian mesh of the domain (0,2)^2 as an easy way to initialize the vector of nodal coordinates and cell connectives of the mesh by hand. These values can also come from a mesh generated via a mesh generator (further details later in Sect. 3). The goal of this example is to compute component \(ij\) of the stiffness matrix \(K^e\) for each cell \(e\) in the mesh, namely

\[
[K^e]_{ij} \equiv \int_{\hat{\Omega}} (J^{-1} \cdot \nabla \hat{s}_i) \cdot (J^{-1} \cdot \nabla \hat{s}_j) |\det(J^e)| \, d\hat{\Omega}.
\]

Here, \(\hat{s}_i\) is the \(i\)-th Lagrangian shape function in the reference quadrilateral \(\hat{\Omega} \equiv (0,1)^2\), \(J^e \equiv \nabla \varphi_e\) is the transposed Jacobian, and \(\varphi_e : \hat{\Omega} \to \Omega_e\) is the isoparametric map, \(\varphi_e(x) = \sum_i x^e_i \hat{s}_i(x)\), being \(x^e_i\) the coordinate vector of the \(i\)-th local node in the \(e\)-th cell of the mesh. Note that the gradient operator \(\nabla\) is the transpose of the derivative.

The code of this example is in Listing 10. The main goal of the code snippet is to build a vector `cell_mat` that represents the elemental stiffness matrices for all cells of the mesh; the stiffness matrix for cell id \(e\) is fetched as `cell_mat[e]`. The code in Listing 10 is intentionally very low-level to illustrate the fundamental building blocks of the library but an average user rarely needs to build all these quantities manually. (The high-level API is discussed in Sect. 3 and used in Listing 13 to build `cell_mat`.) We start by defining the computational mesh via the plain vector of nodal coordinates and the vector of cell connectivity. Then, we built the reference shape functions for this mesh, which is done via a small variation of the code previously discussed in Listing 9, since now we need to build Lagrangian shape functions for the reference quadrilateral instead of the reference triangle. Then, we define the vectors of integration points and weights to integrate quantities in the reference element. We use a single integration point (enough to integrate the elemental matrix exactly in this case), but one can use an arbitrary number. The next step is to prepare the integrand of the elemental matrices, representing the cell-wise geometrical map with a cell-by-cell linear combination of the nodal coordinates and the shape functions in \(\hat{\Omega}\). On can further apply the gradient operator cell-by-cell to this array to end up with a matrix of size \(n \times n\) representing the components of the stiffness matrix before integration at each one of the underlying cells. The cell-wise stiffness matrix `cell_mat` is computed by integrating these objects cell-by-cell using the quadrature rule and the Jacobian of the geometrical map.

3. The high-level API

3.1. From low-level to high-level. All the low-level machinery presented in previous sections is sufficient to implement the computation of elemental matrices and vectors for a wide range of weak forms. In any case, this machinery is general and reaches FE applications. However, these low-level functions require an in-depth knowledge of the library and are not convenient for non-expert users willing to use the library at a higher level. To solve this issue, Gridap introduces higher-level objects that wrap the low-level arrays and related functions, providing a more convenient API that is safer and simpler to use in the majority of cases. A Gridap user rarely needs to deal with the low-level
arrays directly. Only library developers and advanced users need to be aware of the low-level API. In contrast, the high-level API of Gridap described below is specific to FE implementations. It relies on the abstract type CellDatum, which represents a generic object containing an array describing a quantity defined on the cells of a mesh (i.e., each entry in the array is associated with a cell id), plus some meta-data that gives useful information not available in the plain array itself. This metadata includes the mesh on top of which the cell-wise array is defined and a trait indicating whether the cell-wise data is defined in the reference or physical domain of the corresponding mesh. Concrete types implementing the CellDatum interface need to implement three main functions, namely get_data, get_triangulation, and DomainStyle. Function get_data returns the plain array wrapped in the CellDatum object, whereas the other two functions provide the extra meta-data. On the one hand, get_triangulation returns the underlying mesh, encoded via the abstract type Triangulation. There are two main types for representing computational grids in Gridap, namely DiscreteModel and Triangulation. The former contains all the data provided by mesh generators like Gmsh, namely node coordinates, cell and lower-dimensional objects (like vertices, edges and faces) connectivities, and face labels for imposing boundary conditions. On the other hand, the Triangulation type represents a discretization of an integration domain. Note that one can extract different Triangulation objects from the same DiscreteModel since a FE mesh can represent different integration domains (e.g., the cells in the bulk of the meshes and the faces on the boundary would define two different integration domains). On the other hand, DomainStyle returns either ReferenceDomain() or PhysicalDomain() indicating whether the underlying data is defined in the reference or the physical domain respectively. This trait allows the user to handle high-level objects without knowing whether the underlying data is in the reference or physical domain. It allows us to accommodate a wide spectrum of FE methods in the same framework since the library is not limited to interpolations defined in the reference space only. In summary, working with CellDatum objects is more convenient than working with the plain array directly because the stored meta-data allows us to implement safety checks and implicit conversions when manipulating the object.

To illustrate this, let us consider the code in Listing 11. In this snippet, we build a CellField object \( f_\Omega \) and evaluate it at two different CellPoint objects \( x_\Omega \) and \( x_\Gamma \). The types CellField and CellPoint are among the more important ones within the CellDatum type-hierarchy. The CellPoint type represents evaluation points defined on the cells of a mesh (including, but not restricted to, integration points). In this case, \( x_\Omega \) represents integration points in the bulk of a Cartesian FE mesh, whereas \( x_\Gamma \) contains integration points on the boundary faces of the mesh. A CellPoint object is composed of three ingredients. First, the plain vector that contains the evaluation points for all cells of the mesh. In particular, \( x_\Omega \) is built from the array cell_q that, for each cell, contains a small vector of evaluation points (a single point in this case). The second ingredient is the triangulation object on top of which the integration points are defined. In this case, we build \( \Omega \) (representing the bulk of the mesh) and \( \Gamma \) (representing its boundary). The final ingredient of a CellPoint is a trait that indicates whether the given coordinates are in the reference or the physical domain.
We define the CellField object $f_\Omega$ (which is finally evaluated at $x_\Omega$ and $x_\Gamma$) in the remaining part of Listing 11. The abstract type CellField represents a piece-wise function (or function basis) defined on the cells of a given mesh. Here, we show how to manually build a CellField from the user-defined function $\text{ffun}$. We use the GenericCellField constructor that takes the plain array containing a Field object for each cell of the mesh. (In this case, all cells share the same field, but one could assign each cell with a different function to have a truly piece-wise function.) The other arguments are the triangulation on top of which the CellField is defined and the trait indicating the domain type. In this case, we define the cell field on top of the bulk triangulation $\Omega$ and we use the value PhysicalDomain() since the user-defined function $\text{ffun}(x)$ is defined in terms of a physical point $x$. The resulting object $f_\Omega$ can be easily evaluated at the points in $x_\Omega$ and $x_\Gamma$ using the function notation, namely $f_\Omega(x_\Omega)$ and $f_\Omega(x_\Gamma)$. Safety checks and conversions happen under the hood when evaluating $f_\Omega$ using the metadata stored in these objects. On the one hand, when calling $f_\Omega(x_\Omega)$, the code infers that $f_\Omega$ is defined in the physical space and the evaluation points $x_\Omega$ in the reference one. The coordinates in $x_\Omega$ are internally transformed to the physical space using the geometrical map of $\Omega$ before their actual evaluation. On the other hand, when calling $f_\Omega(x_\Gamma)$, the code infers that $f_\Omega$ and $x_\Gamma$ are defined on different triangulations. The restriction of $f_\Omega$ to the boundary triangulation $\Gamma$ happens implicitly before the evaluation on the points in $x_\Gamma$. This example clearly shows that it is more convenient to work with CellDatum objects than handling the plain low-level arrays directly. Writing the operations that happen under the hood would be much more tedious and less generic. In addition, this high-level API does not rely on any code-generation step. It simply leverages the fact that we can efficiently represent cell-wise data using (lazy) arrays.

### 3.2. Example: cell-wise elemental matrices (revisited)

Now, we revisit the example introduced in Listing 10 and will re-write it using the more flexible syntax provided by CellDatum objects (see Listing 12). Recall that the objective is to build an array $\text{cell\_mat}$ that represents the elemental stiffness matrix for all cells of a FE mesh. In Listing 12, we start by building the discrete model using the built-in Cartesian mesh generator, instead of writing the node coordinates and cell connectivity by hand as we previously did in Listing 10. The next step is to define the shape functions and the quadrature rule for the reference cell. We build these objects from Gridap functions instead of explicitly doing all the underlying operations. On the one hand, the shape functions are extracted from a ReferenceFE object. The abstract type ReferenceFE represents a reference FE in the sense of Ciarlet [20]. In particular, the reference FE provides the cell topology, a basis of the polynomial space defined on the reference cell (e.g., the shape functions) and the corresponding basis of DOFs. In this example, we build a scalar first-order Lagrangian reference FE on top of the reference quadrilateral (represented by the constant QUAD) and extract the reference shape functions. Note that the returned object $s$ is equivalent to the one computed more manually in Listing 10. On the other hand, the quadrature points and weights are computed from a quadrature rule object constructed with the Quadrature.
constructor. Here, we build a quadrature rule of degree 1 for the reference quadrilateral. The resulting arrays \( q \) and \( w \) of reference quadrature points and weights are also the same as the ones constructed manually in Listing 10.

### Listing 12. Creating a lazy vector of elemental stiffness matrices via CellDatum objects.

The next part of the code consists in wrapping all these quantities in CellDatum objects to compute the elemental matrices. The array of cell-wise shape functions are wrapped in an instance of the concrete type SingleFieldFEBasis. This type is a specialization of CellField for the particular case of FE basis in single-field computations. Even though not exposed in this example, functionality for multi-field problems is also provided by Gridap and will be presented later in Sect. 4. The SingleFieldFEBasis constructor takes the plain array of cell-wise shape functions, the triangulation object \( \Omega \) on top of which the shape functions are defined, a trait indicating if the shape functions are test or trial bases and a trait indicating if the shape functions are defined in the reference or the physical space (as commented above). Test shape functions are constructed from \( s \) (a column vector), whereas trial shape functions are built from the transpose of \( s \) (a row vector). In this way, we end up having a matrix when broadcasting a product over the row and column vectors as previously shown in Listing 10. On the other hand, we build a Measure object to numerically compute the integrals in the cells of the triangulation \( \mathbf{x}_\Omega \). The Measure object is built from the arrays containing the quadrature rule for all cells in the triangulation, the corresponding triangulation object, and the trait indicating that the quadrature rule is in the reference space.

Using the resulting FE bases \( u \) and \( v \) and the measure object \( d\Omega \), computing the elemental stiffness matrix is trivial by using the high-level API associated with them. Note that the resulting syntax is almost identical to the corresponding mathematical notation used to define the bilinear form of the Poisson equation, namely \( a(u,v) = \int_{\Omega} \nabla u \cdot \nabla v \ d\Omega \). Function \( \nabla \) returns a new CellField object representing the (physical) gradient of the input. In this case, the code detects that the given shape functions \( u \) and \( v \) are in the reference space (thanks to the trait ReferenceDomain()). The reference shape function gradients are internally pushed to the physical space by using the inverse of the Jacobian of the geometrical map associated with the underlying triangulation. The resulting objects are multiplied using the dot product, which results in a new CellField object representing the result of the operation. Finally, we integrate the result using the measure \( d\Omega \). Note that \( \int f d\Omega \) or \( \int f d\Omega \) is just a short-hand for integrate(\( f, d\Omega \)), which is the function in charge to compute the cell-wise integral of \( f \) using the measure \( d\Omega \). The object \( a \) resulting from the integral is an instance of the type DomainContribution. This type is a dictionary-like structure that stores the contributions of several domains to an integral. In this case, the elemental matrices have contributions from a single domain (the triangulation \( \Omega \)), but other examples might have contributions from more than one mesh (e.g., the right-hand-side of a Poisson equation with body forces and Neumann boundary conditions would lead to contributions from two separate domains). To extract the contribution from the cells in the triangulation \( \Omega \), we simply index \( a[\Omega] \), which returns a lazy array of length the number of cells in \( \Omega \) containing the contribution to the computed integral for each cell. It is indeed the array containing the elemental stiffness matrices. Even though the code in Listing 12 is fully functional, it can be further simplified by building the objects \( u \), \( v \), and \( d\Omega \) using even higher-level functions instead, see Listing 13.
### 3.3. Finite Element assembly

We finally show how to assemble global matrices and vectors from the elemental contributions computed in previous examples. The FE assembly routines in Gridap are defined employing the abstract type `Assembler`. This type defines the interface of the assembly routines, which can be specialized for particular cases of interest. Currently, `SparseMatrixAssembler`, which implements an assembler that builds global sparse matrices from elemental contributions, is the only specialization, but others can be added in the future. E.g., one could easily implement a new `Assembler` type that builds matrix-free operators [21] instead of sparse matrices.

Conceptually, an `Assembler` object takes some elemental matrices or vectors and returns a globally assembled object (e.g., a sparse matrix or a vector). To illustrate this, let us consider the code in Listing 13. In the first lines, we build the array `cell_mat` containing the elemental stiffness matrices of all the cells in the underlying mesh. This array is equivalent to the one previously computed in Listing 12, but in this case, we have used higher-level functions of the Gridap API. In particular, the FE basis functions `u` and `v` are extracted from the `FESpace` object `V`. In short, `FESpace` is an abstract type that represents an interpolation space defined on top of the cells of a FE mesh. This data structure contains detailed information about the underlying interpolation, e.g., one can recover the test and trial shape functions `v` and `u` from it. The objective of this snippet is to assemble the local matrices in `cell_mat` into a global sparse matrix `A`. First, we need to define the local-to-global DOF map. This data is available in the array `cell_dofs`, which can be extracted from the FE space `V`. Note that `cell_dofs` is a vector of vectors such that `cell_dofs[e]` returns a vector of global DOF ids corresponding to cell id `e`. The final assembly is straightforward once the arrays `cell_mat` and `cell_dofs` are available. One needs to build an assembler object and then call the `assemble_matrix` function. This function implements the assembly loop and, thus, the user does not need to write it by hand. The assembly of the right-hand-side vector is handled analogously, but it is not included here for the sake of brevity. Note that the assembly interface is flexible enough to potentially use different local-to-global DOF maps for the rows and columns (useful, e.g., for multi-physics problems). It can also build the assembled operator from several arrays of local contributions (useful, e.g., to assemble operators defined via contributions associated with different integration domains). In addition, the assembly routines are generic and completely independent of the way the local contributions and the local-to-global DOF maps are constructed. These quantities only need to be instances of sub-types of the Julia `AbstractArray` interface. Thus, one could even assemble local contributions computed with other Julia packages or directly in the user code if needed, which shows the high modularity of the library.

#### Listing 13. Creating a lazy vector of elemental stiffness matrices via the high-level API of Gridap and assembling it into a global sparse matrix.

```julia
using Gridap
using SparseArrays

# Mesh
domain = (0,2,0,2)
cells = ...
data = ([cell_mat],[cell_dofs],[cell_dofs])
A = assemble_matrix(assem,data)
@assert isa(A,SparseMatrixCSC)
```

#### 3.4. Putting all pieces together: A Poisson FEM solver explained

At this point, we are ready to combine all pieces discussed in previous sections to build an actual FE solver. To illustrate the underlying steps, we consider a Poisson equation as the model problem. This choice is just for simplicity and brevity. Gridap can readily handle much more complex PDEs, including multi-field, nonlinear and time-dependent ones. E.g., we refer to the Gridap tutorial web page [22] for a wide range of examples solved with the library. Here, the goal is to find a numerical approximation of the solution `u` of a Poisson equation with both Dirichlet and Neumann boundary conditions, namely

\[
\begin{align*}
-\Delta u &= f_D \text{ in } \Omega, \\
\frac{n \cdot \nabla u}{\nu} &= f_N \text{ on } \Gamma_N.
\end{align*}
\]

The problem data are the computational domain `\Omega`, the Dirichlet and Neumann boundaries, namely `\Gamma_D` and `\Gamma_N`, respectively, and their corresponding boundary values, described by functions `f_D`, `f_D` and `f_N`. We define `\Omega = \Omega_D \cup \Omega_N` as the Boolean difference of the box `\Omega_D = (0, .5) \times (0, 1) \times (0, 1.5)` and the cylinder `\Omega_N = \{ x \in \mathbb{R}^3 \ : \ (x_2-.5)^2 + (x_3-.75)^2 \leq .3^2 \}`. The loading functions `f_D`, `f_D` and `f_N` are defined so that the manufactured function `u(x) = (x_1 + x_2 + x_3)^k`, `k` being
the interpolation order, solves the problem. This manufactured solution will be used to evaluate the discretization error and to assess the correctness of the solver. We consider a standard FE method with conforming second-order \((k = 2)\) Lagrangian elements to compute the numerical solution (see, e.g., [23] for more details). The corresponding weak form of the (discrete) problem is: find: \(u_h \in U_h\) such that \(a(u_h, v_h) = \ell(v_h)\) for all \(v_h \in V_h\), where

\[
a(u, v) = \int \Omega \nabla u \cdot \nabla v \, d\Omega \quad \text{and} \quad \ell(v) = \int \Omega f \, v \, d\Omega + \int \Gamma_N f_N \, v \, d\Gamma_N
\]

are the forms of the problem, and \(U_h\) and \(V_h\) are the Lagrangian FE spaces that fulfill the Dirichlet conditions \(u = f_D\) and \(u = 0\), respectively. This weak problem can be easily built and solved using the high-level user API of Gridap as shown in Listing 14.

```julia
1 @using Gridap
2 @using GridapGmsh
3
4 # Geometry
5 model = GmshDiscreteModel("model.msh")
6 ... )dΩ)) < tol
40
41 # Visualization
42 writevtk(Ω, "results_poisson",
43 order=k, cellfields=["eh"=>eh, "uh"=>uh])
```

**LISTING 14.** Solving a Poisson equation with the Gridap API.

We start by generating the FE mesh, a.k.a. the discrete model. So far, we have considered only the built-in Cartesian mesh generator for the sake of simplicity. In this example, we show how to deal with more complex domains represented by unstructured grids. The Gridap library outsources the generation of unstructured meshes to well-known mesh generators like Gmsh [24]. In this example, we consider the plugging GridapGmsh to read a mesh file generated by Gmsh and convert it into a discrete model usable within a Gridap computation. The mesh file is generated using Listing 15 via the official Julia API of the Gmsh project. As a result, it is possible to both generate the mesh (with Gmsh) and solve the PDE (with Gridap) in the same Julia environment, which is very convenient. Note that we have defined some physical groups in the mesh file called "domain", "dirichlet", and "neumann", which identify the domains \(Ω\), \(Γ_D\) and \(Γ_N\) respectively. These labels are used to build the Triangulation objects \(Ω\) and \(Γ\), representing \(Ω\) and \(Γ\), respectively.

The next step in the FE solver is the definition of the manufactured solution \(u\) and the corresponding loading functions \(f_D\), \(f_D\) and \(f_N\). Note that the manufactured solution is defined as a Julia function \(u\) by using the standard Julia syntax to define function objects. It is not the case for FE codes based on symbolic DSLs, which usually force the user to define functions via strings containing the function definition expressed in C/C++ instead of the programming language of the front-end. However, Gridap allows the user to write arbitrarily complex functions directly in the front-end without compromising performance by leveraging the Julia JIT compiler, which will be in charge of compiling efficient machine code. From the manufactured solution \(u\), once can easily define the loading functions \(f_D\), \(f_D\) and \(f_N\) thanks to the build-in automatic differentiation capabilities of Gridap. In particular, \(\Delta(u)\) and \(\nabla(u)\) compute the Laplacian and the gradient of function \(u\) via automatic differentiation. The next lines in the code snippet are similar to the ones seen in previous sections. We build the interpolation space \(V\) and \(U\) and the measure objects \(dΩ\) and \(dΓ\) associated with the triangulations \(Ω\) and \(Γ\) respectively. Note that Dirichlet boundary conditions are specified by the name of the physical group "dirichlet" describing the Dirichlet boundary in the FESpace constructor. The DOFs located on faces included in the selected physical group will be removed from the system (i.e., they are constrained to zero in the resulting test FE space \(V\)). Spaces with non-homogeneous Dirichlet values, e.g., the trial space \(U\), are defined from a test space plus the user-defined function describing the Dirichlet values. Internally, the passed function is interpolated on the faces of the Dirichlet boundary. A vector of Dirichlet DOF values is created and stored within \(U\), which is used later in the FE assembly to modify the right-hand-side accordingly.
The software design of Gridap

The resulting cell contributions are assembled into the global matrix and vector, as seen in Listing 13. The returned object \( \text{op} \) represents the FE element problem to be solved, which is a data structure that holds the matrix and vector of the underlying linear system. The FE problem is solved with a \text{FESolver}. It is built from an object implementing the linear solver interface defined in Gridap. This generic interface allows one to use different linear solvers in the same framework. For instance, one can solve the linear system via a \text{LU} decomposition from UMFPACK \[25\] using \text{LUSolver}, MKL-Pardiso with \text{PardisoSolver} or PETSc \[26\] with \text{PetscSolver}. The two latter solver wrappers are available via the plugins GridapPardiso \[27\] and GridapPETSc \[28\], respectively. The solution of the FE problem, namely \( \text{uh} \), is an instance of the type \text{FEFunction}, which is a specialization of \text{CellField} representing a function in a FE space. Thus, one can manipulate \( \text{uh} \) via the high-level API implemented for \text{CellField} objects as we have seen in previous sections. We use this API to compute the interpolation error \( \text{eh} \) and integrate the \( H^1 \) error norm. The last line of the code snippet writes the discrete solution \( \text{uh} \) and the discretization error \( \text{eh} \) into a VTK file to be visualized in Paraview to produce the plots in Fig. 2.
4. Multi-field equations

4.1. From single-field to multi-field equations. In this section, we discuss the main software abstractions that enable the solution of multi-field PDEs with Gridap. This functionality is of high interest since complex applications in science and engineering involve the solution of coupled multi-physics problems and, thus, the handling of single-field equations is insufficient in many cases. We say that a weak form, find \( u \in U \) such that \( a(u, v) = \ell(v) \) for all \( v \in V \), is multi-field when the functional spaces \( U \) and \( V \) are defined as the Cartesian product of single-field ones, namely \( U = U_1 \times \ldots \times U_B \) and \( V = V_1 \times \ldots \times V_B \). Here, \( B \) denotes the number of fields (or blocks) and \( U_I \) and \( V_I \) for \( I = 1, \ldots, B \), are the underlying single-field spaces, i.e., functional spaces of scalar, vector, or tensor-valued functions.

Note that we will use upper-case letters like \( I \) to refer to the blocks/fields and lower-case ones like \( i \) to refer to particular entries in an array. Some systems of PDEs, e.g., linear elasticity, are naturally handled by defining vector/tensor-valued single-field spaces. Others, e.g., the Stokes problem, involve several separate fields within a multi-field method. In any case, Gridap is general enough to support both approaches; the former was explained in previous sections, whereas the latter is detailed now. The Gridap implementation of multi-field PDEs builds on top of the machinery already explained for the single-field case. Thus, it requires the introduction of very few new ingredients. Essentially, we need a new type of FE space called MultiFieldFESpace that represents the Cartesian product of some single-field spaces and a new container type that preserves the block structure of elemental matrices and vectors in the multi-field case.

4.2. Recursive block representation of elemental matrices and vectors. The weak form in multi-field equations has a well-defined block structure. It is associated with the decomposition of \( a \) and \( \ell \),

\[
a(u, v) = \sum_{I,J=1}^{B} a_{IJ}(u_I, v_J) \quad \text{and} \quad \ell(v) = \sum_{I=1}^{B} \ell_I(v_I),
\]

in terms of the underlying single-field quantities. Here \( v_I \) represents the component of \( v \in V = V_1 \times \ldots \times V_B \) belonging to the single-field space \( V_I \) (analogously for \( u_I \)) and \( a_{IJ} \) and \( \ell_I \) are the individual blocks of the weak form. The diagonal terms \( a_{I\ell} \) and \( \ell_I \) represent the operators and external loads of the underlying single-field equations. The off-diagonal blocks \( a_{IJ} \neq \ell_I \), are the coupling between them. In practice, a significant number of blocks \( a_{IJ} \) and \( \ell_I \) are potentially equal to zero. (All fields are not necessarily directly coupled, e.g., saddle point systems lead to some zero diagonal blocks.) This block sparsity needs to be taken into account to achieve performance. Motivated by this requirement, we adopt the following strategy to represent the elemental contributions to the weak form in multi-field equations.

In Gridap, the elemental matrix and vector \( A^e \) and \( b^e \) (corresponding to the contribution of a cell \( e \) to the multi-field forms \( a \) and \( \ell \)) are never built explicitly. Instead, we build the elemental matrices and vectors \( A^e_{IJ} \) and \( b^e_I \) for each of the underlying non-zero single-field forms \( a_{IJ} \) and \( \ell_I \). By doing so, we reuse the functionality for single-field PDEs, we preserve the block structure of the multi-field problem, we take into account the block sparsity, and we do not need to generate a local DOF numbering for the multi-field spaces (since the local DOF numbering of the single-field ones is enough to build the arrays \( A^e_{IJ} \) and \( b^e_I \)). These arrays are stored in a suitable container type called \texttt{ArrayBlock}(T,N); \( T \) is the type of the blocks within the object, typically some array type, and \( N \) is the number of dimensions. For convenience, we define the aliases \texttt{VectorBlock(T)} and \texttt{MatrixBlock(T)} for \( N=1 \) and \( N=2 \), respectively. The \texttt{ArrayBlock} type is very simple. It just contains two variables, namely \texttt{array} and \texttt{touched}. The first one is an array containing the blocks, e.g., the elemental matrices \( A^e_{IJ} \) for \( I,J = 1, \ldots, B \), and the latter is a Boolean mask that indicates which of the entries in \texttt{array} are zero and thus have not been computed.

Let us consider Listing 16 to illustrate the basic usage of the \texttt{ArrayBlock} type. In the first part of the snippet, we build an \texttt{ArrayBlock} object that represents three blocks \( b_I, I = 1,2,3 \), where \( b_2 \) and \( b_3 \) are conceptually zero, and the first block is the array \( b_1 = (11,12)^t \). You can interpret this particular case as the local vector in a multi-field PDE with three fields, where only the external loads of the first field are different from zero. The \texttt{ArrayBlock} object is build from the underlying variables \texttt{array} and \texttt{touched}. On the one hand, \texttt{array} is built as a vector of vectors. Note that we only set entry \texttt{array}[1], whereas the other components of \texttt{array} remain uninitialized (they are like dangling pointers in C/C++). On the other hand, we define the Boolean mask \texttt{touched} indicating that only the first block has been initialized. Note that one does not need to know the sizes of the zero blocks to build an \texttt{ArrayBlock} object. This fact enormously simplifies the implementation of multi-field equations in Gridap. E.g., many contributions of one of the single-fields (e.g. the external loads) can be computed without knowing anything from the other fields, thus allowing to reuse a lot of functionality for single-field problems. Once an \texttt{ArrayBlock} is available one can easily inspect its block structure with the display function. In this case, the output indicates that the first block is a vector of two entries, whereas the second and third blocks are empty. Several operations are defined over \texttt{ArrayBlock} objects. Most of them are optimized to take into account the underlying block sparsity. See for instance the object \texttt{c} in Listing 16, which is defined as the difference of two \texttt{ArrayBlock} objects. The resulting block sparsity is computed from the given inputs. In this case, the second block is empty since it was empty in both input arrays. A final feature of the \texttt{ArrayBlock} implementation worth mentioning is that it is recursive. One can build a block hierarchy by nesting \texttt{ArrayBlock} objects, see, e.g., the last lines in Listing 16. This is used in several places in the code, e.g., in DG methods for multi-field equations, where one needs two levels of blocks to represent the elemental contributions in interior faces. For instance, for the source term, one needs two blocks \( b_L \) and \( b_R \) to represent the contributions of the left and right cell around an interior face, and then one needs yet another block level for each one of the underlying fields.
LISTING 16. Construction and basic usage of *BlockArray* objects.

4.3. Putting all pieces together: A Stokes FEM solver explained. Finally, we present the high-level user API to solve multi-field PDEs in Gridap. To this end, we consider a 3D Stokes example as the model problem. The PDE we want to solve is

\[
\begin{align*}
-\Delta u + \nabla p &= f_u \text{ in } \Omega, \\
\nabla \cdot u &= f_p \text{ in } \Omega, \\
\mathbf{n} \cdot \nabla u - p n &= g_N \text{ on } \Gamma_N, \\
\end{align*}
\]

(2)

where \( u \) and \( p \) are the velocity and pressure fields, \( f_u \) and \( f_p \) are the source terms (\( f_p = 0 \) usually), and \( g_D \) and \( g_N \) are the Dirichlet and Neumann boundary conditions on the boundaries \( \Gamma_D \) and \( \Gamma_N \) respectively. In this example, the computational domain \( \Omega \) is the same as the one considered in the Poisson example in Sect. 3.4, but we define different boundaries to impose boundary conditions as detailed in Table 2. In particular, we consider a homogeneous Neumann condition on \( \Gamma_N \) and impose an inflow velocity on \( \Gamma_D \) with value

\[
g_D(x) = (0, 0, (1 - (4x_1 - 1)^2)(1 - (2x_2 - 1)^2))^t.
\]

On the other hand, we define a no-slip boundary on \( \Gamma_D \) and constrain the first component of the velocity on \( \Gamma_D \).

Table 2 contains also the labels defined in the mesh file in Listing 15, which represent each of these boundaries.

| Symbol | Label   | Definition            | Condition |
|--------|---------|-----------------------|-----------|
| \( \Gamma_N \) | "outlet" | \((0,0.5) \times (0,1) \times \{1.5\}\) | \( u \cdot \nabla u - p n = 0 \) |
| \( \Gamma_D \) | "inlet" | \((0,0.5) \times (0,1) \times \{0\}\) | \( u = g_D \) |
| \( \Gamma_D \) | "noslip" | \((0,0.5) \times \{0,1\} \times (0,1.5) \cup \partial \Omega_e\) | \( u = 0 \) |
| \( \Gamma_D \) | "ux0"   | \((0,0.5) \times (0,1) \times (0,1.5)\) | \( u_1 = 0 \) |

**Table 2.** Boundary conditions for the Stokes example in Listing 17.

The numerical approximation is based on an inf-sup stable Taylor-Hood pair [29] for velocity and pressure fields, with continuous \( P_2 \) and \( P_1 \) elements, respectively. For this formulation, stabilization is not needed and the discrete weak form for this problem is: find \((u_h, p_h)\) \( (u_h, p_h) \in U_h \times Q_h \) such that \( a((u_h, p_h), (v_h, q_h)) = l((v_h, q_h)) \) for all \((v_h, q_h) \in V_h \times Q_h \), where

\[
a((u, p), (v, q)) = \int_{\Omega} \nabla u \cdot \nabla v - p(\nabla \cdot v) + (\nabla \cdot u) q \, d\Omega \quad \text{and} \quad l(v) = \int_{\Omega} f_v v + f_p q \, d\Omega.
\]
The spaces $U_h$ and $V_h$ are the continuous Lagrangian $P_2$ spaces for the velocity fulfilling the Dirichlet conditions, whereas $Q_h$ is the continuous Lagrangian $P_1$ space for the pressure.

As for the Poisson example, the high-level API of Gridap allows one to build and solve this problem in a compact way using syntax that resembles the mathematical notation (see Listing 17). The first part of the driver builds the FE spaces for velocity and pressure. These spaces are defined using the API for single-field problems. The definition of the velocity space requires some extra work since we have three different parts of the Dirichlet boundary where we define different conditions, namely $\Gamma_{D1}$, $\Gamma_{D2}$, and $\Gamma_{D3}$. Note that one can select which components of a vector/tensor-valued space to constrain on the Dirichlet boundary by supplying a vector of masks. Once the single-field spaces are defined, we define the tensor product via the `MultiFieldFESpace` constructor. Using the tensor product spaces, we can proceed with the driver as in the single-field case. The major exception is that one needs to unpack the velocity and pressure solutions after solving the problem. The resulting objects `uh` and `ph` are single-field FE functions that can be handled as discussed in previous sections. In particular, one can export them in VTK format for visualization in Paraview, see Fig. 3.

The last part of Listing 17 is not needed to solve the problem but is included here to illustrate how one can inspect the lazy vector containing the elemental matrices. The approach is analogous to the one previously explained for the Poisson equation. One extracts the trial and test FE bases and evaluates the bilinear form with them. The vector `cell_mat` contains the local Stokes operator for all cells in the mesh. We inspect the contents of this vector for a particular cell (cell id 7 in this case). As the output of the `display` function reveals, the local operator is represented with a `MatrixBlock` object with a $2 \times 2$ block structure, as expected. Note that the last block is empty. The weak form has no term in the diagonal pressure block (we have used an inf-sup stable formulation). The code automatically infers the block sparsity from the multi-field weak form, as expected.

5. Numerical Examples

5.1. Objectives. The goal of the numerical examples below is to evaluate the performance of the FE framework introduced in this paper and available in the Gridap package. The main goal is to demonstrate that the Julia implementation in Gridap can be as efficient as other state-of-the-art FE libraries that combine a C/C++ back-end plus with a similar Python front-end. Note that the goal is not to outperform previous FE packages but to...
show that Gridap can achieve similar performance while using a simpler design based on a user-friendly programming language (Julia) without considering sophisticated compilers of variational forms like FFC [7]. To this end, we perform performance comparisons against a recent version of FEniCS [4], arguably the most popular state-of-the-art FE code, which provides a compact high-level user interface similar in spirit to the one in Gridap.

5.2. Experimental setup. The numerical experiments below are computed on a personal computer with an Intel Core i7-10510U CPU with 1.80GHz and approximately 16GiB of RAM running with Ubuntu 18.04.3 LTS. We use Gridap version 0.16.2 and Julia version 1.6.2. The Julia binary is invoked with flags --check-bounds=no and -O 3 to disable array bound checks and select the maximum optimization level of the compiler, respectively. We use FEniCS version 2019.2.0-dev0 (the latest version available when running the examples) with Python 3.6.9. In addition, we force optimization of the form compiler by setting the parameters as shown in Listing 18. All reported times are computed as the minimum of 4 identical runs to mitigate the effect of operating system jitter. Regarding mesh generation, we consider both structured and unstructured grids. The structured meshes are generated with the build-in mesh generators available in Gridap and FEniCS, whereas the unstructured meshes are generated with Gmsh version 4.4.1. The .msh file generated by Gmsh is directly read with the GridapGmsh plugin (as already shown in Listing 14). In FEniCS, the .msh file is converted to xdmf format with the meshio python package version 4.3.11 [30] before being consumed by FEniCS routines.

5.3. Poisson benchmark. We consider the Poisson equation as a model problem to assess the performance of Gridap in single-field computations. In this benchmark, the problem formulation is the same as the one in Eq. (1), except that we only consider Dirichlet boundary conditions, namely \( \Gamma_D = \partial \Omega \) and \( \Gamma_N = \emptyset \). The computational domain \( \Omega \) is defined now by two different geometries. The first one is the unit cube \((0, 1)^3\) and the second one is the perforated box geometry meshed with Listing 15. We label these geometries as geo1 and geo2, respectively. The first geometry can be meshed with structured grids. The second case requires unstructured grids. Note that we use both structured and unstructured grids to explore potential performance differences between these cases. The FE formulation used to discretize the problem is a continuous Galerkin method with Lagrangian elements of first \((k = 1)\) and second \((k = 2)\) order. We only use tetrahedral meshes for the comparison because FEniCS does not support hexahedral meshes. (Note that Gridap supports general cell topologies.) A summary of the parameters used in this example is found in Table 3. The Gridap code used in this benchmark is an adaptation of the one in Listing 14, in which we have included timing directives. The FEniCS code builds on the code examples in the FEniCS project web page.

We define two different time measures for the performance comparison. The first one is the total time needed to assemble the linear system from scratch. The second time measures the time to reassemble the linear system by reusing as much data as possible from the previous one. Note that the first timing provides information about the global performance of the library, including all simulation phases needed to build the discrete system of linear equations. It includes reading (or generating) the FE mesh and the interpolation spaces, handling the weak form,
| Parameter               | Value                                                                 |
|------------------------|----------------------------------------------------------------------|
| Model problem          | Poisson equation with Dirichlet boundary conditions                   |
| Geometry               | Unit cube (geo1) and perforated box (geo2)                            |
| Numerical scheme       | Continuous Galerkin with Lagrangian interpolation                      |
| $k$ (interpolation order) | 1 and 2                                                               |
| Cell topology          | Tetrahedron                                                           |
| Mesh type              | Structured for geo1 and unstructured for geo2                         |
| FE library             | Gridap and FEniCS                                                    |

**Table 3.** Summary of the Poisson benchmark.

allocating the system, and assembling it. This timing provides information about the overall Gridap performance and includes several parts of the library not explained in detail herein. The second timing measures the cost of the assembly loop. One can relate this measure to the time needed to assemble the linearized problem at each Newton-Raphson iteration in the solution of a nonlinear PDE since it is possible to reuse a significant amount of information from previous iterations like the sparsity pattern of the system matrix. This second measure is directly related to the low-level routines detailed in this paper. These two time measures are labeled for further reference as **assembly from scratch** and **in-place assembly** respectively.

![Graphs](https://via.placeholder.com/150)

**Figure 4.** Poisson benchmark: Scaling of CPU time for the phases **assembly from scratch** and **in-place assembly** for computations on unstructured mesh and structured meshes.

Fig. 4 shows the scaling of these two timing phases with respect to the number of cells in the mesh for both structured and unstructured meshes. We also provide Gridap times divided by FEniCS ones to facilitate the comparison between the two libraries. Gridap is slightly faster in most cases but differences are not substantial. The main differences are in the assembly from scratch phase. For $k = 1$, Gridap is faster for unstructured grids, whereas FEniCS is faster for structured ones. Note that Gridap is optimized for Cartesian grids of n-cube cells but not for structured grids of simplices. This latter case uses generic data structures for unstructured grids, which are not leveraging any particular grid topology. Optimization for this specific case would reduce the difference between the two libraries. In addition, FEniCS is tailored for simplicial elements, whereas Gridap implements general cell topologies. In any case, the more generic implementation of Gridap does not result in a significant performance hit with respect to FEniCS.
5.4. Stokes benchmark. In this second benchmark, we consider the Stokes equation to assess the performance of Gridap in multi-field computations. We consider Eq. (2) with pure Dirichlet boundary conditions, i.e., $\Gamma = \partial \Omega$ and $\Gamma_N = \emptyset$, and define the forcing functions $f_u$, $f_p$, and $g_D$ so that the manufactured velocity and pressure functions $u(x) = (x_1^2 + 2x_2^2, -x_2^2, 0)^T$ and $p(x) = x_1 + 3x_2$ are the solutions of the problem. Since we impose Dirichlet boundary conditions on the entire boundary $\partial \Omega$, the mean value of the pressure is constrained to zero to have a well-posed problem.\(^3\)

$$\int_{\Omega} p \, d\Omega = 0.$$  

We define the computational domain $\Omega$ with the same geometries (the unit cube and the perforated box labeled as geo1 and geo2 as before) and the same FE meshes (tetrahedral cells) as previously considered in the Poisson benchmark. The FE method used to discretize the problem is the one considered in the Stokes example in Listing 17. The definition of this benchmark is summarized in Table 4.

| Parameter          | Value                                      |
|--------------------|--------------------------------------------|
| Model problem      | Stokes equation with Dirichlet boundary conditions |
| Geometry           | Unit cube (geo1) and perforated box (geo2)  |
| Numerical scheme   | Taylor-Hood interpolation                   |
| $k$ (interpolation order) | 2 for velocity and 1 for pressure          |
| Cell topology      | Tetrahedron                                 |
| Mesh type          | Structured for geo1 and unstructured for geo2 |
| FE library         | Gridap and FEniCS                           |

Table 4. Summary of the Stokes benchmark.

This benchmark is run with Gridap and FEniCS. The Gridap driver is a variation of listing Listing 17 in which we include timing routines and impose different boundary conditions. Again, the FEniCS code is based on the examples on the project web page. As in the Poisson benchmark, we measure the time to assemble the linear system from scratch and to re-assemble the system reusing pre-computed information. The results are shown in Fig. 5. Note that Gridap is significantly faster for assembly from scratch whereas FEniCS is slightly faster for in-place assembly. This time difference can be attributed to the more general implementation of the assembly loop for different cell topologies available in Gridap.

6. Conclusions

In this work, we have presented the main software abstractions behind the Gridap library. They enable the implementation of general-purpose FE codes by leveraging the new possibilities provided by the emerging Julia programming language. We have detailed the main building blocks of the low-level computational back-end based on the multiple-dispatch paradigm of Julia, as well as the high-level API that allows the user to specify the weak form in a syntax almost identical to the whiteboard mathematical notation.

The low-level back-end is extensible and highly customizable, whereas the high-level API provides a convenient way to write PDE solvers in few lines of code. Both the back-end and front-end are implemented in the same programming language. We do not consider any compiler of variational forms or any code generation mechanism at the package level. This feature is not usual in previous FE frameworks with similar user APIs. It enormously simplifies the usage, development and maintenance of the library. Numerical experiments for the Poisson and Stokes problems show that this new software design and the adoption of the Julia programming language lead to run-time performance similar to existing libraries like FEniCS.

Due to length constraints, we have focused on the FE discretization of linear single-field and multi-field problems, which is also the core functionality needed to solve nonlinear and transient problems. In any case, Gridap also provides a high-level interface for nonlinear problems, and time-dependent PDEs can be conveniently solved with the GridapODEs extension package [31]. The functionality presented in this work can readily be used to integrate the local portions of distributed sparse systems of linear equations in parallel computations. A user API for distributed-memory computations based on MPI will be soon released in the extension Package GridapDistributed [32].

7. Acknowledgments

This research was partially funded by the Australian Government through the Australian Research Council (project number DP210103092), the European Commission under the FET-HPC ExaQUte project (Grant agreement ID: 800898) within the Horizon 2020 Framework Programme and the project RTI2018-096898-B-I00 from the “FEDER/Ministerio de Ciencia e Innovación – Agencia Estatal de Investigación”. F. Verdugo acknowledges support from the Spanish Ministry of Economy and Competitiveness through the “Severo Ochoa Programme for Centers of Excellence in R&D (CEX2018-000797-S)”.

\(^3\)We note that Gridap provides a dedicated finite element space that internally handles the zero mean constraint. It is useful when defining pressure spaces in confined flow problems (i.e., when the full boundary is of Dirichlet type like in this benchmark).
Figure 5. Stokes benchmark: Scaling of CPU time for the phases assembly from scratch and in-place assembly for computations on unstructured mesh and structured meshes.

References

[1] W. Bangerth, R. Hartmann, and G. Kanschat. Deal.II — A general-purpose object-oriented finite element library. ACM Transactions on Mathematical Software, 33(4), 2007. doi:10.1145/1268776.1268779.

[2] Santiago Badia, Alberto F. Martín, and Javier Principe. FEMPAR: An Object-Oriented Parallel Finite Element Framework. Archives of Computational Methods in Engineering, 25(2):195–271, 2018. doi:10.1007/s11831-017-9244-1.

[3] Łukasz Kaczmarczyk, Zahur Ullah, Karol Lewandowski, Xuan Meng, Xiao-Yi Zhou, Ignatios Athanasiadis, Hoang Nguyen, Christophe-Alexandre Chalons-Mouriesse, Euan Richardson, Euan Miur, Andrei Shvarts, Mebratu Wakeni, and Chris Pearce. MoFEM: An open source, parallel finite element library. Journal of Open Source Software, 5(45):1441, 2020. doi:10.21105/joss.01441.

[4] Anders Logg and Garth N Wells. DOLFIN: Automated finite element computing. ACM Transactions on Mathematical Software, 37(2), 2010. doi:10.1145/1731022.1731030.

[5] Florian Rathgeber, David A. Ham, Lawrence Mitchell, Michael Lange, Fabio Luporini, Andrew T.T. McRae, Georg Heinke, Teodor Bercea, Graham R Markall, and Paul H.J. Kelly. Firedrake: Automating the finite element method by composing abstractions. ACM Transactions on Mathematical Software, 43(3):24, 2016. doi:10.1145/2998441.

[6] FreeFEM++ project webpage. https://freefem.org [Accessed: 2021-08-01].

[7] Robert C Kirby and Anders Logg. A compiler for variational forms. ACM Transactions on Mathematical Software, 32(3):417–444, 2006. doi:10.1145/1163641.1163644.

[8] Santiago Badia and Francesc Verdugo. Gridap: An extensible Finite Element toolbox in Julia. Journal of Open Source Software, 5(52):2520, 2020. doi:10.21105/joss.02520.

[9] Jeff Bezanson, Alan Edelman, Stefan Karpinski, and Viral B. Shah. Julia: A fresh approach to numerical computing. SIAM Review, 59(1):65–98, 2017. doi:10.1137/141000671.

[10] Iain Dunning, Jing Huchette, and Miles Lubin. JuMP: A modeling language for mathematical optimization. SIAM Review, 59(2):295–320, 2017. doi:10.1137/15M1020575.

[11] Christopher Rackauckas and Qing Nie. DifferentialEquations.jl – A Performant and Feature-Rich Ecosystem for Solving Differential Equations in Julia. Journal of Open Research Software, 5(1), 2017. doi:10.5334/jors.151.

[12] Mike Innes. Flux: Elegant machine learning with Julia. Journal of Open Source Software, 3(25):602, 2018. doi:10.21105/joss.00602.
The software design of Gridap

[13] Jarrett Revels, Miles Lubin, and Theodore Papamarkou. Forward-Mode Automatic Differentiation in Julia. Arxiv preprints, 2016. arxiv id: 1607.07892.

[14] Simon Danisch and Julius Krumbiegel. Makie.jl: Flexible high-performance data visualization for Julia. Journal of Open Source Software, 6(65):3349, 2021. doi:10.21105/joss.03349.

[15] Petr Krysl. FinEtools git repository. https://github.com/PetrKryslUCSD/FinEtools.jl [Accessed: 2021-08-01].

[16] Tero Frondelius and Jukka Aho. JuliaFEM - open source solver for both industrial and academia usage. Rakenteiden Mekaniikka, 50(3):229–233, 2017. doi:10.23998/rm.64224.

[17] Ferrite git repository. https://github.com/Ferrite-FEM/Ferrite.jl [Accessed: 2021-08-01].

[18] FillArrays git repository. https://github.com/JuliaArrays/FillArrays.jl [Accessed: 2021-08-01].

[19] Kamil A. Khan and Paul I. Barton. A vector forward mode of automatic differentiation for generalized derivative evaluation. Optimization Methods and Software, 30(6):1185–1212, 2015. doi:10.1080/10556788.2015.1025400.

[20] Philippe G. Ciarlet. The Finite Element Method for Elliptic Problems. Classics in Applied Mathematics, 40:1–511, 2002. doi:10.1137/1.9780898719208.

[21] Utpal Kiran, Sachin Singh Gautam, and Deepak Sharma. GPU-based matrix-free finite element solver exploiting symmetry of elemental matrices. Computing, 102(9):1941–1965, 2020. doi:10.1007/s00607-020-00827-4.

[22] Gridap tutorials webpage. https://gridap.github.io/Tutorials/stable/ [Accessed: 2021-08-01].

[23] Claes Johnson. Numerical Solution of Partial Differential Equations by the Finite Element Method. Dover Publications, 2009. ISBN 048646900X.

[24] Christophe Geuzaine and Jean-François Remacle. Gmsh: A 3-D finite element mesh generator with built-in pre- and post-processing facilities. International Journal for Numerical Methods in Engineering, 79(11):1309–1331, 2009. doi:10.1002/NME.2579.

[25] Timothy A. Davis. Algorithm 832: UMFPACK V4.3 - An unsymmetric-pattern multifrontal method. ACM Transactions on Mathematical Software, 30(2):196–199, 2004. doi:10.1145/992200.992206.

[26] Satish Balay, Shrirang Abhyankar, Mark F. Adams, Jed Brown, Peter Brune, Kris Buschelman, Lisandro Dalcin, Alp Dener, Victor Eijkhout, William D. Gropp, Dinesh Kaushik, Matthew G. Knepley, Dave A. May, Lois Curfman McInnes, Richard Tran Mills, Todd Munson, Karl Rupp, Patrick Sanan, Barry F. Smith, Stefano Zampini, Hong Zhang, and Hong Zhang. PETSc Users Manual. Technical Report ANL-95/11 - Revision 3.10, Argonne National Laboratory, 2018.

[27] GridapPardiso git repository. https://github.com/gridap/GridapPardiso.jl [Accessed: 2021-08-01].

[28] GridapPETSc git repository. https://github.com/gridap/GridapPETSc.jl [Accessed: 2021-08-01].

[29] Franco Brezzi and Michel Fortin, editors. Mixed and Hybrid Finite Element Methods, volume 15 of Springer Series in Computational Mathematics. Springer New York, New York, NY, 1991. doi:10.1007/978-1-4612-3172-1.

[30] Nico Schlömer. meshio. 2021. doi:10.5281/ZENODO.5167106.

[31] Santiago Badía, Oriol Colomes, and Francesc Verdugo. GridapODEs. 2021. doi:10.5281/zenodo.5518455.

[32] GridapDistributed git repository. https://github.com/gridap/GridapDistributed.jl [Accessed: 2021-08-01].