SClib, a hack for straightforward embedded C functions in Python

Esteban Fuentes*,†, Hector E. Martinez‡

1 1NTRODUCTION

Embedding code written in other languages is a common theme in the Python context, the main motivation being speed boosting. Several alternatives exist to achieve this, such as Cython [Cython], CFFI [CFFI], SWIG [SWIG], weave [weave], among others. We present yet another alternative, which may be quite close to CFFI than to the others. The motivation to write SClib grew out of the urge to integrate C code, which was already written, into the Python environment, minimizing the intervention of the code. Part of the resulting work is briefly introduced later, in the engineering application section.

Nevertheless, embedding compiled code in Python will naturally have an impact in performance, for instance, when the compiled code takes care of computer intensive numerics. The first application we introduce (in particle physics), leverages SClib in this sense, outsourcing the numerics to the compiled code and using the Python environment for visualization.

2 SCLIB

At the core of SClib† is ctypes [Hell], which actually does the whole work: it maps Python data to C compatible data and provides a way to call functions in DLLs or shared libraries. SClib acts as glue: it puts things together for the user, to provide him with an easy to use interface.

The requirements for SClib are very simple: call a function on an array of numbers of arbitrary type and size and return the output of the function, again of arbitrary type and size.

The resulting interface is also very simple: A library is initialized in the Python side with the path to the DLL (or shared library) and a list with the names of the functions to be called:

In [1]: import SCLib as sc
In [2]: lib = sc.Clib('test.so', ['fun'])

The functions are then available as members of the library and can be called with the appropriate number of arguments, which are one dimensional arrays of numbers. The function returns a list containing the output arrays of the function:

In [3]: out, = lib.fun([0])

In the C counterpart, the function declaration must be accompanied with specifications of the inputs and outputs lengths and types. This is accomplished with the helper macros defined in sclib.h:

An arbitrary number of inputs or outputs can be specified, for example:

1. The code for SClib and example use are available at <https://github.com/drestebon/SClib>
SCL_IT(fun, 2, INT, FLOAT);
void fun(int * out0, float * out1,
    int * in0, float * in1) {
    out0 = 42 + in0[0];
    out1[0] = in1[0] + in1[1];
    out1[1] = powf(in1[0], in1[1]);
}

In the function declaration, all the outputs must precede the inputs and must be placed in the same order as in the SCL macros.

These specifications are processed during compilation time, but only the number of inputs and outputs is static, the lengths of each component can be overridden at run time:
In [4]: lib.INPUT_LEN[‘fun’] = [10, 1]
In [5]: lib.retype()

In these use cases the length of the arguments should be given to the function through an extra integer argument.

In the function body, both inputs and outputs should be treated as one dimensional arrays.

3 Application in Quarkonium Physics

3.1 Motivation

The Schrödinger equation is the fundamental equation for describing non-relativistic quantum mechanical dynamics. For the applications we will present in this section we will focus on the time-independent version which, in natural units, is given by

\[
\left( -\nabla^2 + V(r) \right) \psi(r) = E \psi(r). \tag{1}
\]

It corresponds to an eigenvalue equation where the term inside the parenthesis in l.h.s. is called the Hamiltonian operator, the value \( E \), its eigenvalue, is the measurable quantity (the energy) associated with it, \( \mu \) is the reduced mass of the system (it correspond the mass of the particle in one-particle systems) and the wavefunction, \( \psi(r) \) is the entity containing all the information about the system, since its modulus squared \( r \) corresponds to the probability density of a given measurement, it has to be normalized to unity. The term \( V(r) \) in the Hamiltonian is called the potential.

Since its discovery, the Schrödinger equation has played an important role in our understanding of nature and it is present in almost every aspect of modern physics. In this section we will review some cases where SClib has been used to implement solutions of the computing problems associated with eq. (1) that arise in the study of heavy quarkonia\(^2\).

Quarkonium is a bound-state composed by a quark and its corresponding antiquark. By heavy we mean states composed by charm and bottom quarks, called charmonium and bottomonium respectively. Due to its large mass, the top quark decays before forming a bound state. In heavy quarkonium the relative velocity between the quark and antiquark inside of the bound-system is believed to be small enough for the system to be considered, at least in a first approximation, non-relativistic, making it suitable for being described by eq. (1). Considering the equal mass case with a spherically symmetric potential the angular part can be neglected (it correspond to the spherical harmonics) and the relevant part of eq. (1) reduces to the one-dimensional equation given by

\[
\left[ -\frac{1}{m^2} \frac{d^2}{dr^2} + \frac{l(l+1)}{mr^2} + V(r) \right] y_{n,l}(r) = E_{n,l} y_{n,l}(r), \tag{2}
\]

where \( r \) is the relative distance between the quark and the antiquark, \( l \) is the angular momentum quantum number, \( m \) is the (anti)quark mass (for equal mass \( 2 \mu = m \)), \( y_{n,l} \) is called the reduced wavefunction and the eigenvalue \( E_{n,l} \) is interpreted as the binding energy of the bound-system, where \( n = 0, 1, 2, \ldots \) accounts for the number of nodes (radial excitations) of the wavefunction. The mass of the quarkonium is then given by

\[
M = 2m + E_{n,l}. \tag{3}
\]

The potential \( V(r) \) describes the quark-antiquark interaction, it is a function of \( r \) and \( \Lambda_{QCD} \), the typical hadronic scale (\( \sim 200 \text{MeV} \)). For \( r \Lambda_{QCD} \ll 1 \) (short-distance regime) the potential may be evaluated perturbatively, but for \( r \Lambda_{QCD} \sim 1 \) (long-distance regime) it cannot. To overcome this issue, models based on non-relativistic reductions of phenomenological observations have been used to describe heavy quarkonia, one these being the so-called Cornell potential \([Eich74, Eich78, Eich79]\)

\[
V(r) = \frac{a}{r} + kr, \tag{4}
\]

where \( a \) and \( k \) are parameters which need to be fixed by experimental (or lattice) data of some observable. This potential incorporates two of the main observed characteristics of the quark-antiquark interaction: at short distances it exhibits a Coulombic behavior and in the long-distance regime the interaction is dominated by a confinement phase.

Since the beginning of the last decade, non-relativistic effective field theories (EFT), in particular non-relativistic QCD (NRQCD) \([Cas85, Bod94]\) and potential NRQCD (pNRQCD) \([Bra99]\), have become the state-of-the-art tools for the study of heavy quarkonia (for review see \[Bra04]\). NRQCD is obtained from QCD integrating out modes that scale like \( m \), while pNRQCD is obtained from NRQCD integrating out modes that scale like the quark momentum\(^3\).

The physics of the modes that have been integrated out is encoded in Wilson coefficients that must be calculated comparing at the same scale the results (observables, Green functions) of the EFT, with the ones of QCD (for NRQCD) or NRQCD (for pNRQCD). A key feature of pNRQCD is that it allows the relativistic corrections to the quark-antiquark potential to be organized as an expansion in powers of \( 1/m \).

Up to second order \( V(r) \) can be written as

\[
V(r) = V^{(0)}(r) + \frac{V^{(1/m)}(r)}{m} + \frac{V^{(1/m^2)}(r)}{m^2}, \tag{5}
\]

3. These EFT exploit the hierarchy of energy scales present in the bound-system. If the relative velocity of the (anti)quark, \( v \), is small, we have that \( mv^2(\sim E) \ll mv(\sim p) \ll m \), where \( p \) is the momentum of the particles and \( E \) its kinetic energy. If one is interested in studying a phenomena that happens at the scale \( E \) (like the binding) it is more suitable to integrate out degrees of freedoms with energies that scale like the other two higher scales, this is the motivation behind pNRQCD. For a detailed analysis of the scales present in heavy quarkonia we refer the reader to \[Bra04]\.

---

\(^2\) For a comprehensive review of the status and perspectives of the research in heavy quarkonia we refer the reader to chapter four of \[Bra14]\.
where \( V^{(1/m)} \) and \( V^{(1/m^2)} \) are derived from QCD through the matching procedure with NRQCD. The details about \( V^{(1/m)} \) and \( V^{(1/m^2)} \) and how they are obtained are beyond the scope of this document, however, we can list some of their features:

- They correspond to correlators that in the short-distance regime can be computed in perturbation theory.
- In the long-distance regime they can be computed in lattice QCD, however only some of these correlators have been calculated.
- Eq. (4) correspond, at least qualitatively, to the leading order \( V^{(0)} \) in eq. (5).

For the details about the derivation of the terms present in eq. (5) we refer the reader to [Bra00] and [Pin00]. It is important to recall that, although it can not be evaluated analytically in the whole range of \( r \), eq. (5) represents a model-independent expression for the quark-antiquark potential, contrary to models like the one presented in eq. (4).

Using perturbation theory to include the relativistic corrections to the potential, the expression for the bound-state mass reads

\[
M = 2m + E_{n,l}^{(0)} + \frac{\langle nl|V^{(1/m)}(r)|nl\rangle}{m} + \frac{\langle nl|V^{(1/m^2)}(r)|nl\rangle}{m^2} + \frac{1}{m^2} \sum_{m \neq n} \left| \frac{\langle nl\rangle|V^{(1/m^2)}|ml\rangle}{E_{n,l}^{(0)} - E_{m,l}^{(0)}} \right|^2,
\]

where \( E_{n,l}^{(0)} \) comes from solving eq. (2) with \( V(r) = V^{(0)}(r) \) and

\[
\langle nl|f(r)|nl'\rangle \propto \int_0^\infty dr y_{n,l}(r)f(r)y_{n',l'}(r),
\]

where the proportionality factor will depended on the corresponding quantum numbers of the operators appearing in \( V^{(1/m)} \) and \( V^{(1/m^2)} \).

### 3.2 Applications of SClib

The simplest computational problem related to eq. (2) is to find \( E_{n,l} \) for a given \( n \) and \( l \). Methods to solve this problem have been implemented since long ago (see for instance [Fall85]), in a nutshell, the standard method consist of applying two known constraints to the reduced wavefunction \( y_{n,l} \):

- The number of nodes of \( y_{n,l}(r) \) must be equal to \( n \).
- \( y_{n,l}(r) \) has to be normalizable

\[
\int_0^\infty dr y_{n,l}^2(r) = 1.
\]

In general \( y_{n,l}(r) \) will diverge except when \( E_{n,l} \) corresponds to an eigenvalue. The procedure to find the eigenvalue consists in to perform a scan of values of \( E_{n,l} \) until \( y_{n,l}(r) \) has \( n \) nodes and converges for a large enough value of \( r \) (see Fig. 1). This implies that for each test value of \( E_{n,l} \) eq. (2) must be (numerically) solved. A popular Mathematica [Mat9] implementation of this method has been available in [Luc98]. This script has the advantage that the user can profit from the Mathematica built-in functions to plot, integrate or store the resulting wavefunctions, however, it has a very poor performance. With the goal of mimicking some of the advantages of this script, but without compromising speed, we ported the algorithm in [Luc98] to Python. The resulting script, SChroe.py, uses SClib to implement the speed-critical parts of the algorithm. In SChroe.py the wavefunctions are stored as NumPy arrays [NumPy] so when the script is run within IPython [IPy] together with SciPy [SciPy], NumPy and Matplotlib [Mplot] the user can profit of the same or more flexibility as with the Mathematica script plus a boosted speed. In table 1 we compare the performance of SChroe.py against

![Fig. 1: Reduced wavefunctions \( y_{n,l}(r) \) for two steps in the search of the eigenvalue \( E_{n=1,l=1} \). For each step in the process to find the eigenvalue the nodes of the \( y_{n,l} \) are counted, only when the value of \( E_{n,l} \) corresponds to an eigenvalue \( y_{n,l} \) is not divergent. In the plot (a) \( E_{n,l} = 3.1 m \) and \( y_{n,l} \) fulfills the condition of having one node, however, the accuracy in the value of \( E_{n,l} \) is too low and the function diverges. In the plot (b) \( E_{n,l} = 3.10952 m \) so \( y_{n,l} \to 0 \) for larger values of \( r \). We have used the Cornell potential eq. (4) with parameters \( m = 1 \, a = 0.1, \, k = 0.5 m^2 \), all dimensions defined in terms of the mass.](image)
other implementations of the same algorithm.\(^6\)

In [Bra14] SChroe.py has been used to evaluate the relativistic corrections to the mass spectrum of quarkonium in the long-distance regime. In that paper the relativistic corrections \(V_{1/m}\) and \(V_{1/m^2}\) appearing in (6) were evaluated assuming the hypothesis that in the long-distance regime the interaction between the quark and the antiquark can be described by a string. In Fig. 2 we show some of the energy levels (masses) corresponding to the string spectrum. It is noteworthy to mention that all the numerical calculations and plots of that paper were done with IPython using the SciPy library.

An application in which the speed of SChroe.py plays an important role is fixing the parameters of the potential given some experimental input. For instance, consider the problem of finding the parameters \(a\) and \(k\) of eq. (4) together with \(m\), given the experimental values of the masses of three different quarkonium states. If relativistic corrections are included, in order to find the parameters we must solve a system of three equations like eq. (6). For each probe value of \((a,k,m)\) we have to find the eigenvalues and reduced wavefunctions of eq. (2) and then with these values evaluate the sums and integrals in (6). A parameter fixing of this type was necessary to implement in [Bra14b]. The implementation has been carried out using SChroe.py together with a mixture of C and SciPy functions using SClib to link both environments.\(^7\)

Another related computational problem that arises from the study of heavy quarkonium hybrids, bound-states composed by a quark-antiquark pair plus an exited gluon, is to solve a system of \(N\) S\(\text{chrödinger-like}\) coupled equations. Explicitly the system to solve reads

\[
\left(-\frac{\delta_{ij}}{m} \frac{d^2}{dr^2} + V_{ij}(r,l)\right)u_{j,n,l}(r) = E_{n,l}u_{i,n,l}(r), \tag{8}
\]

4. The paper describing the script ranks fifth among the most cited papers (91 citations) of the International Journal of Modern Physics C with the last citation from July 2014.

5. Code available at <https://github.com/heedmane/schroepy/>.

6. Although the aim of this section is not to compare performance of S\(\text{chrödinger equation solvers, but to present an application in which SClib can improve the speed of a known algorithm, we must mention that there are solvers that offer better performance than the current version of SChroe.py. For instance, the solver presented in [dftatom] implements a more sophisticated integration method and allows refinements in the radial mesh. With these improvements the dftatom solver can reach a speed-up of at least two orders of magnitude compared to the current version of SChroe.py.

7. Some of the code will be available once the paper appears online.

where \(i = 1, 2, \ldots N\) and the angular momentum dependence has been included in the potential matrix. A method to solve this equation for the case \(N = 2\) has been implemented in [Ber14]. The method relies on an extension of the nodal theorem [Ama95] and convergence conditions for the components of the vector wavefunction \(u_{j,(n,l)}(r)\). The extension of the nodal theorem states that the number of nodes of the determinant of the matrix \(U_{n,l}(r)\), whose columns are \(N\) lineal-independent solutions of eq. (8), is equal to \(n\). The procedure then consist in a scan of values \(E_{n,l}\); in each step the set of equations (8) is solved and the nodes of \([U_{n,l}(r)]\) are counted for a large enough interval of \(r\). As in the one-dimensional case, if \(E_{n,l}\) approached to an eigenvalue the components of \(u_{j,(n,l)}(r)\) converge for large \(r\). In the solution presented in [Ber14] the performance-intensive parts of the implementation rely on C functions linked to the IPython interface trough SClib.

As an example of the application of the method implemented in [Ber14], in Fig. 3 we show the results for the search of the first two eigenvalues and wavefunctions with the matrix potential given by

\[
V_{ij}(r,l) = \begin{pmatrix}
\frac{l(l+1)+2}{mr^2} + F_0(r) & -2\sqrt{l(l+1)} \frac{m}{mr^2} \\
-2\sqrt{l(l+1)} \frac{m}{mr^2} & \frac{l(l+1)+2}{mr^2} + F_1(r)
\end{pmatrix}
\tag{9}
\]

where

\[
F_i(r) = \ln(a_i + b_ir).
\]

In all the applications described in this section the combination of SClib and the SciPy library within an interactive environment like IPython provided a powerful framework based entirely on open source software for solving problems that require a high performance and visualization tools.
measurements and the reference $r$:

$$C : u = \pi(y, r),$$

but it also may comprise internal states. They are commonly used to reconstruct the state $x$ out of the history of $y$ and $u$. The latter systems are called state observers and the whole is called feedback control.

We use SClib to put together a simulator for these kind of systems. Both the system derivatives $f(\cdot)$ and the control $\pi(\cdot)$ are written in C and are evaluated using SClib. As stated before, the system state represents a natural phenomena, therefore it is natural to describe it as a continuous time variable, as eq. (10) suggests. To calculate the system state we have to solve this equation. In our simulator this is achieved using numerical methods, namely the integration routines available in scipy.integrate. On the other hand, the controller is usually implemented in a real-time computer, which can only sample $y$ at a fixed interval (called $h$): it is a discrete-time system. This means, that the simulator only needs to evaluate $\pi(\cdot)$ at given times.

Traditional controllers took the form of linear filters, which could even be implemented using analog circuitry. As control techniques and requirements advance, more complex controllers are devised. Many modern control techniques are based on optimization methods. Time-optimal controllers, for example, require the solution of an usually very complex optimization problem, to find a control $u$ that leads the system state $x$ towards its target $r$ in minimum time [Gru11]:

$$u^* = \pi^*(x) = \underset{\pi \in U, x \in X}{\text{argmin}} \{T_x(u)\}.$$  

(11)

Here $T_x(u)$ is the time required to lead $x$ towards its target and $X$ and $U$ are the regions where we want $x$ and $u$ to be confined, they constitute the constraints for the control problem. These kind of controllers require exhaustive computation and it is natural to implement them in C.

For motivation, we present the results for a minimum-time control strategy for a relatively simple and well known problem, the double integrator [Fu13]:

$$\frac{d}{dt} \begin{bmatrix} x_0 \\ x_1 \end{bmatrix} = \begin{bmatrix} u/\tau_0 \\ u/\tau_1 \end{bmatrix}.$$  

(12)

The relevance of this system lays in that it models many mechanical systems: $u$, $x_0$ and $x_1$ may represent acceleration, speed and position, for example.

Fig. 5 presents a minimum time control strategy for this system.

The form of $\pi(x)$ for this case reveals its non-linear nature.

Fig. 6 presents the trajectory developed by the state using this control strategy and random initial conditions.

These results were obtained using SClib and the devised simulator. The example code provided with SClib is ready to reproduce them.

The main advantage we obtained from this work was that, since we were using a Linux based real time system in our test rig, we could use exactly the same code for the simulations and the experimental tests. Another feature of this work is that

\[\text{Fig. 3: Solutions for the components of the vector wavefunction } u_{nl}(r) \text{ for the first two eigenvalues } (l = 1) \text{ of eq. (8) with the matrix potential given in (9). We have used } m = 1, a_0 = 1, b_0 = 0.5, a_1 = 2 \text{ and } b_1 = 0.1. \text{ The eigenvalues are } E_{m=0,n=1} = 1.01727 m \text{ for Fig. (a) and } E_{m=1,n=1} = 1.18789 m \text{ for Fig. (b).} \]

\[\text{Fig. 4: General scheme of a control system.} \]
Weave provides tools for including C/C++ code
[Ralf Gommers.](http://github.com/scipy/weave)

**SWIG** SWIG developers. [SWIG: Simplified Wrapper and Interface Generator](http://swig.org/)

Weave effectively replaces Simulink in all of our use cases using only free software.

### 5 Final Remark

We hope the applications of SClib scope beyond the ones listed in this paper since we believe it provides a simple but powerful way to boost Python performance.

### 6 Acknowledgments

H.M. acknowledges financial support from DAAD and the TUM Graduate School during the realization of this work.

### References

[Bra14] N. Brambilla, S. Eidelman, P. Foka, S. Gardner, A. S. Kronfeld, M. G. Alford, R. Alkofer and M. Butenschoen et al., QCD and Strongly Coupled Gauge Theories: Challenges and Perspectives arXiv:1404.3723.

[Eich74] E. Eichten, K. Gottfried, T. Kinoshita, J. B. Kogut, K. D. Lane and T. M. Yan, The Spectrum of Charmonium, Phys. Rev. Lett. 34 , 369 (1975) Erratum-ibid. D 36 , 1276 (1976)

[Eich78] E. Eichten, K. Gottfried, T. Kinoshita, K. D. Lane and T. M. Yan, Charmonium: The Model, Phys. Rev. D 17 , 3090 (1978) Erratum-ibid. D 21 , 313 (1980)

[Eich79] E. Eichten, K. Gottfried, T. Kinoshita, K. D. Lane and T. M. Yan, Charmonium: Comparison with Experiment, Phys. Rev. D 21 , 203 (1980).

[Cas85] W. E. Caswell and G. P. Lepage, Effective Lagrangians for Bound State Problems in QED, QCD, and Other Field Theories, Phys. Lett. B 167 , 437 (1986).

[Bod94] G. T. Bodwin, E. Braaten and G. P. Lepage, Rigorous QCD analysis of inclusive annihilation and production of heavy quarkonium, Phys. Rev. D 51 , 1125 (1995) Erratum-ibid. D 55 , 5855 (1997)

[Bra04] N. Brambilla, A. Pineda, J. Soto and A. Vairo, Effective field theories for heavy quarkonium, Rev. Mod. Phys. 77 , 1423 (2005)

[Bra99] N. Brambilla, A. Pineda, J. Soto and A. Vairo, Potential NRQCD: An Effective theory for heavy quarkonium, Nucl. Phys. B 566 , 275 (2000)

[Bra00] N. Brambilla, A. Pineda, J. Soto and A. Vairo, The QCD potential at O(1/m), Phys. Rev. D 63 , 014023 (2001)

[Pin00] A. Pineda and A. Vairo, The QCD potential at O(1/m^2); Complete spin dependent and spin independent result, Phys. Rev. D 63 , 054007 (2001) Erratum-ibid. D 64 , 039902 (2001)

[Fal85] P. Falkensteiner and H. Grosse and F. Schoeberl and P. Hertel Comput. Phys. Comm. 34 , 287 (1985)

[Luc98] W. Lucha and F. F. Schoberl, Solving the Schrödinger equation for bound states with Mathematica 3.0, Int. J. Mod. Phys. C 10 , 607 (1999)

[dftatom] Čertík, O., Pask, J. E., Vackář, J. (2013). dftatom: A robust and general Schrödinger and Dirac solver for atomic structure calculations, Computer Physics Communications, 184(7), 1777–1791.

[Math] Wolfram Research, Inc. Mathematica Version 9.0 (2012)

[Cython] Stefan Behnel, Robert Bradshaw, Lisandro Dalcín, Mark Florisson, Vitja Makarov, Dag Sverre Seljebøt. Cython is an optimising static compiler for both the Python programming language and the extended Cython programming language (based on Pyrex). [http://www.cython.org/](http://www.cython.org/)

[CFII] Armin Rigo and Maciej Fijalkowski C Foreign Function Interface for Python. [http://cfi.readthedocs.org/en/release-0.8/](http://cfi.readthedocs.org/en/release-0.8/)

[SWIG] SWIG developers. [SWIG: Simplified Wrapper and Interface Generator](http://swig.org/)

[weave] Ralf Gommers. [Weave provides tools for including C/C++ code within Python code.](https://github.com/scipy/weave)

[Heller] The ctypes module, [https://docs.python.org/3.4/library/ctypes.html#module-ctypes](https://docs.python.org/3.4/library/ctypes.html#module-ctypes)