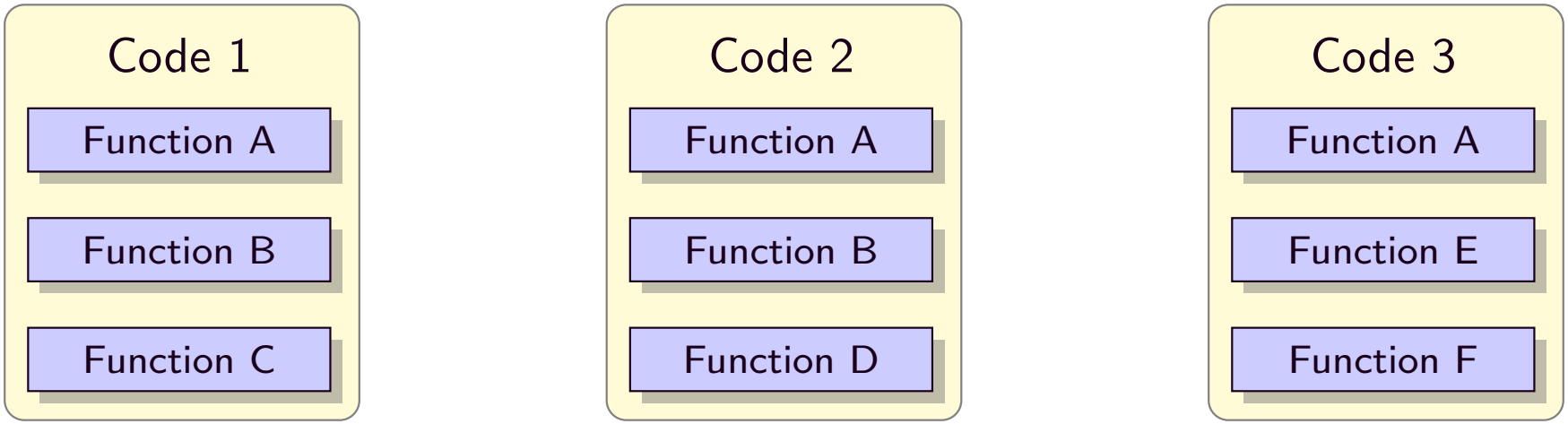
The CECAM Electronic Structure Library: An Overview

Micael Oliveira

Max Planck Institute for the Structure and Dynamics of Matter, Hamburg, Germany

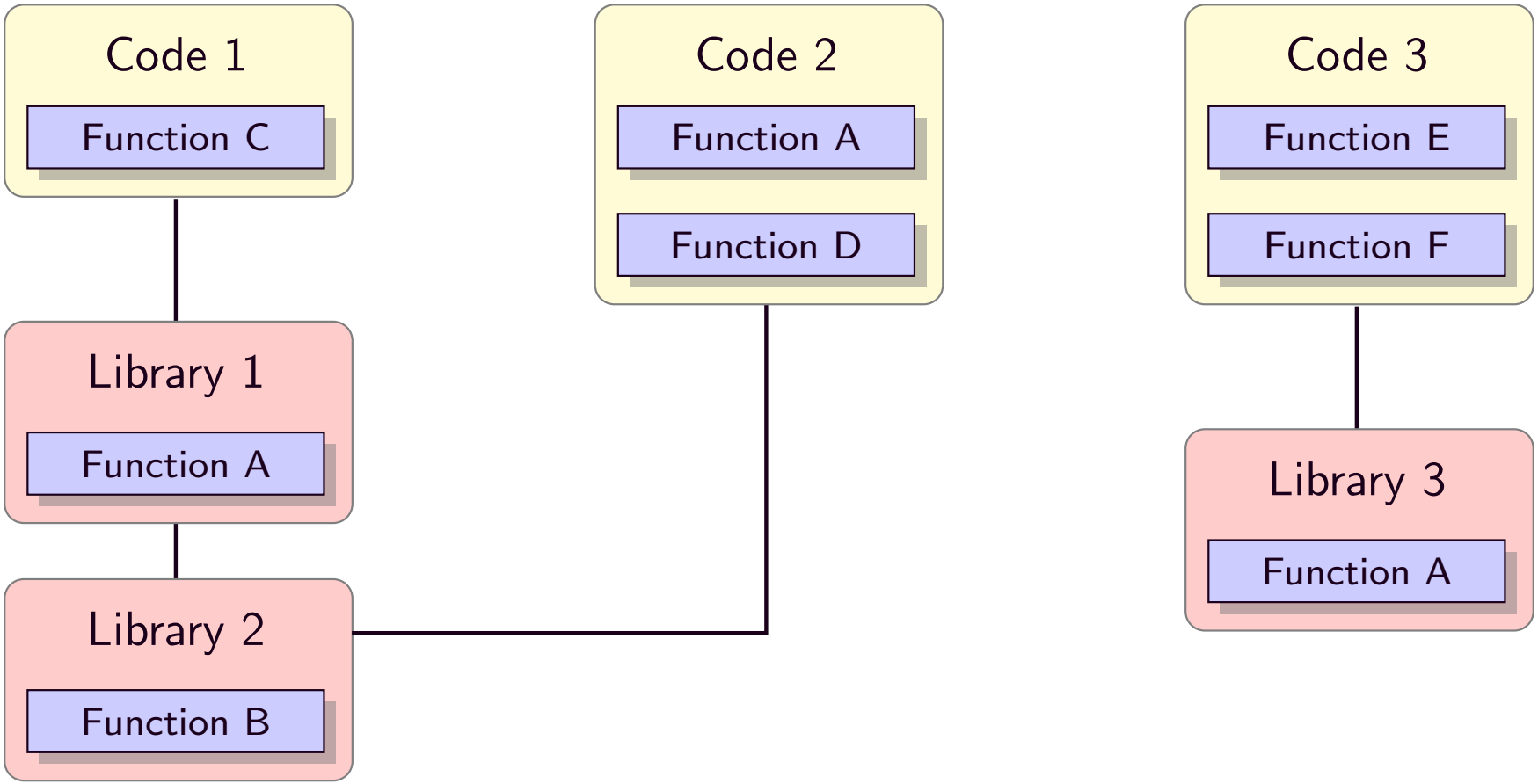
ELSI Connector Meeting, August 16, 2017

The “traditional” electronic structure code framework

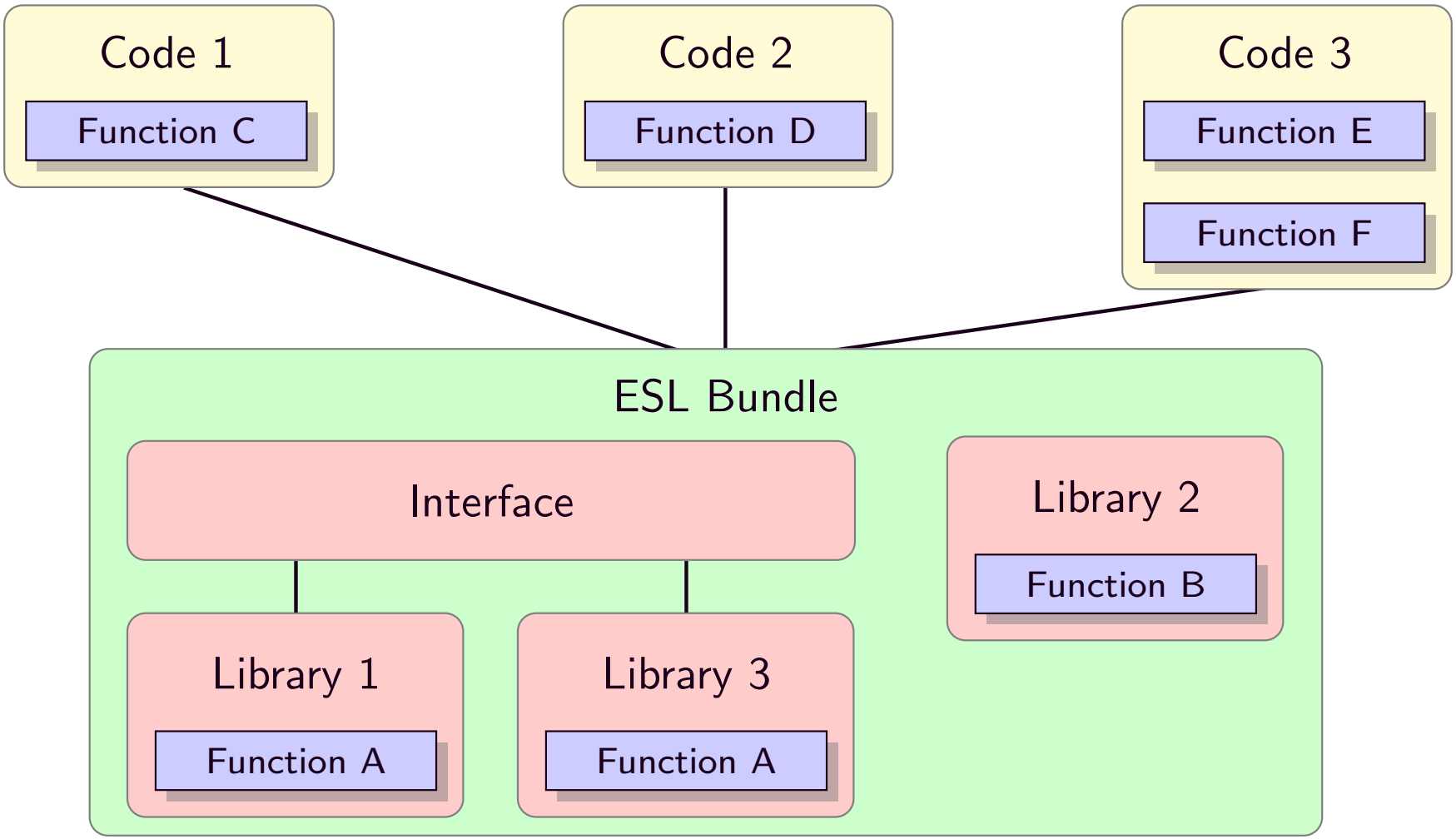


Duplication of code = Duplication of effort = Wasted time

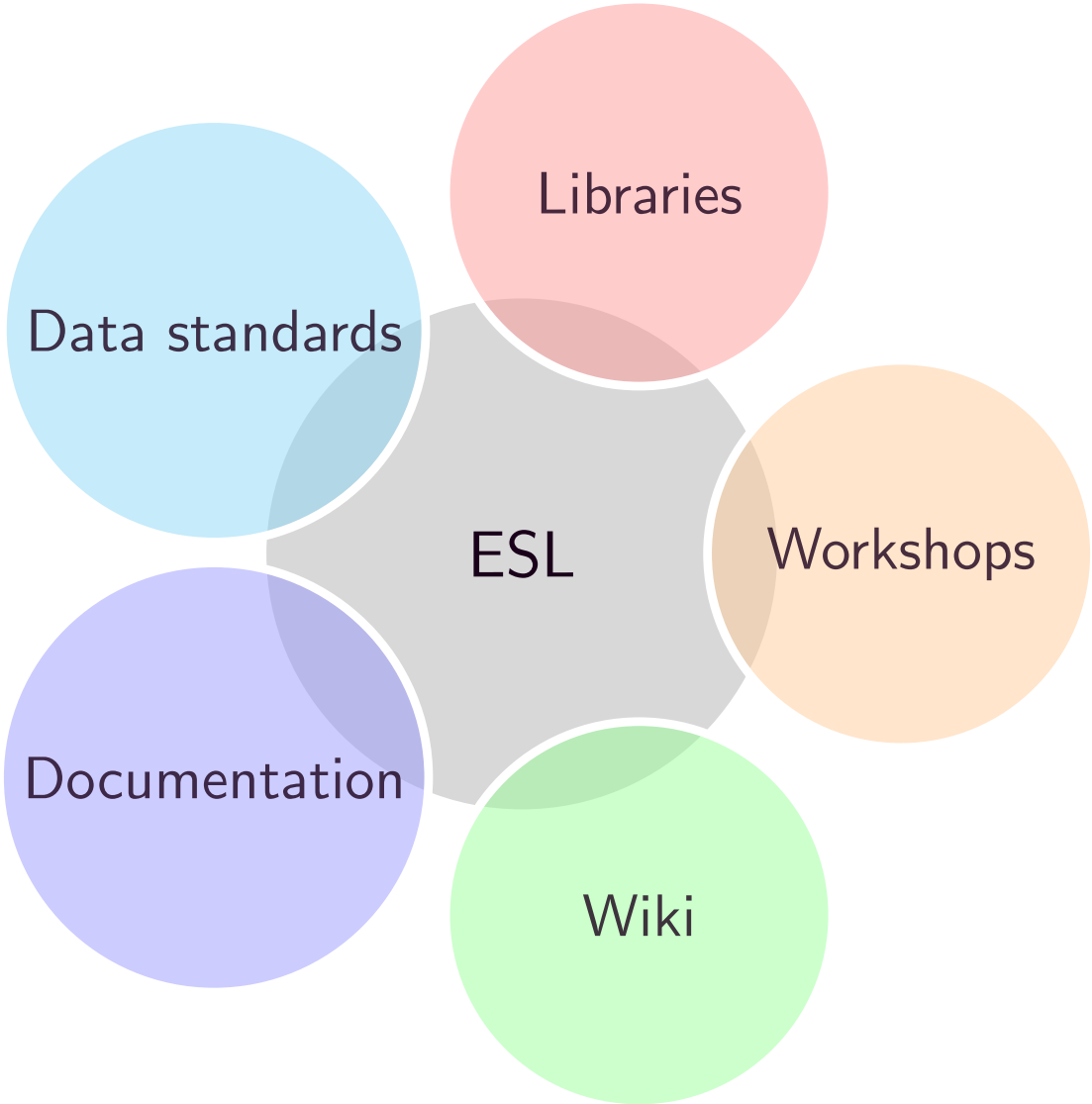
Current electronic structure code developement framework



The CECAM Electronic Structure Library



The CECAM Electronic Structure Library



esl.cecam.org

ESL Organization

- Wiki: `esl.cecam.org`
- Software repositories: `gitlab.e-cam2020.eu/esl`
- Curating team and advisory committee
- **Contributions from volunteers**

Why software libraries?

Minimize the development/maintenance effort over the whole program life:

- Implementation of new features is a time consuming task $\Rightarrow$ do not reinvent the wheel!
- In large codes most of the developers' time is dedicated to maintenance tasks $\Rightarrow$ sharing code allows to share the maintenance effort
- Larger user base $\Rightarrow$ more feedback, bug reports, etc.
- Disentangle science from software engineering
- Let the library developers worry about optimization and porting to new architectures and new languages

ESL requirements, coding standards, and best practices

ESL modules and libraries should be:

- Easy to maintain
- Easy to use

Recomendations:

- C/Fortran programming languages with bindings to other languages
- Licenses: LGPL and/or BSD
- Source code documentation (e.g., Doxygen)
- Version control system
- Use a standard build system type (e.g., Autotools, CMake)
- Provide proper error handling

E-CAM

- European Union H2020 Center of Excellence
- e-infrastructure for software, training and consultancy in simulation and modeling
- 16 CECAM nodes, 3 PRACE Centres, 12 Industrial Partners, and 1 Centre for Industrial Computing

E-CAM

Objectives:

- Create over 150 new, robust software modules, directed at industrial and academic users in:
  - electronic structure calculations
  - classical molecular dynamics
  - quantum dynamics
  - mesoscale and multi-scale modelling
- Training of current and future academic and industrial researchers
- Support industrial end-users in their use of simulation and modelling

The ESL was a used as a pilot project for E-CAM

ESL Workshops

- Unique opportunity to bring the electronic structure code developers together
- Long workshops (1 week or more)
- 10-20 participants
- Divided into discussion and coding sessions

Kick-off Workshop (July-August 2014)

- Extended software development workshop at CECAM-HQ-EPFL
- 3 days for discussion + 5 weeks of development
- $\sim$ 10 participants
- Extensive discussions on:
  - coding standards
  - documentation
  - licenses
  - ESL organization
  - etc
- Wiki design and initial content
- Start of several coding projects

Uilities Toolbox (June 2015)

- Second software development workshop at CECAM-HQ-EPFL
- 1 day for discussion + 4 days of development
- $\sim$ 20 participants (12 for the coding sessions)
- Some of the explored themes:
  - Low-level utilities (memory management, error handling, etc)
  - Further steps towards standardization of coding standards
  - Geometry tools

Towards a Common Format for Computational Materials Science Data (January 2016)

- Workshop co-organized by ESL and NoMaD
- 3 days of discussion and 8 days of coding
- Objectives:
  - Discuss relevant problems in creating a common data format for materials science
  - Set the specifications for an Electronic Structure Common Data Format (ESCDF)
  - Start the development of a library to read/write ESCDF files

ESL Coding Workshop: Solvers

- Z-CAM, Zaragoza, Spain
- 6-17 June 2016
- 1 day of discussion and 10 days of coding
- 15 participants
- Three projects:
  - Kohn-Sham solvers
  - Poisson solvers
  - Atomic solvers

The ESL bundle

- Collection of software libraries:
  - Software developed within the ESL
  - Other relevant software distributed with the agreement of the authors
- Bindings for different programming languages
- Documentation

Coming soon:

- Common installer for easy compilation and installation
- Debian and RPM packages
- Collection of easyconfigs for EasyBuild

Current ESL libraries

- ELSI
- ESLW_Solvers
- FDF
- Flook
- GridXC
- Libescdf
- LibOMM
- Libpspio
- Libvdwxc
- Libxc
- MatrixSwitch
- Poke
- SQARE

Libpspio

- Parsing and writing of DFT pseudopotentials
- Aims to support as many formats as possible
- Support for norm-conserving pseudopotentials
- PAW and ultra-soft pseudopotentials coming soon

Libescdf

The Electronic Structure Common Data Format:

- Provides a framework for saving and reading data
- Stores all metadata needed to interpret the data
- Does NOT enforce specific physical representation
- Provides means to store different types of data
- Extendable and not restrictive

Libescdf

- File format based on HDF5 with support for parallel I/O
- Use HDF5 hierarchical structure and organize data in groups:
  - system
  - basis sets
  - densities
  - states
  - etc
- Groups common to all codes supporting the format are at the root of the file structure (e.g. /system/*)
- Code specific groups are found under /extensions/code/*
- Flexible data ordering on disk for more efficient I/O

Libescdf

Flexible API:

- The host code does not have to change the way it stores the data in memory
- The user is allowed to write and/or read the data in any order
- Validation (correctness and completeness) separate from I/O

GridXC

- Evaluate XC functionals on a grid
- Returns potentials and energies
- Uses Libxc for single point evaluation
- Non-local functionals (e.g. van der Waals DF)

Solvers for Quantum Atomic Radial Equations (SQARE)

- Creation of radial grids and definition of discretized radial functions
- Solution of ordinary differential equations on a radial grid
- Solution of radial wave-equations in various flavors

Future

ESL Coding Workshop: Demonstrators

- CECAM-HQ-EPFL, Lausanne, Switzerland
- 5-16 February 2018
- Write DFT codes using the ESL libraries
- Improve libraries API, documentation, tests, etc

E-CAM Extended Software Development Workshop: Scaling Electronic Structure Applications

- CECAM-IRL, Dublin, Ireland
- Dates to be announced
- Parallelization, scalability, performance

Acknowledgements

- Dominic Tildesley and all the CECAM staff
- Fabiano Corsetti, Yann Pouillon, and Nick Papior
- NoMaD

- All the organizers of the ESL workshops:
  - Emilio Artacho
  - Volker Blum
  - Damien Caliste
  - Stefano de Gironcoli
  - Ivan Girotto
  - Georg Huhs
  - Martin Lueders
  - Arash Mostofi
  - Mike Payne

All the ESL contributors