Session 5 : accuracy assessment

We are currently in a situation where PBE is the default choice for most/all solid-state codes (in contrast to quantum chemistry, where B3LYP is the default). This is partly due to inherent properties of PBE (it works reasonably well on a reasonable number of properties for reasonably large numbers of crystals), but is also partly due to convenience (John Perdew distributed the PBE subroutines to all major code developers in the nineties, so it was easy to adopt, and that's what code developers did).

Hasn't the time come to go beyond the use of PBE as a default? It is the feeling of several experts that the answer to that question is 'yes'. This is not easy, though, as it requires a judicious assessment by a DFT user about what is the most appropriate choice of functional. By far not all users have that knowledge. Moreover, referees may object against the use of alternative functionals – they too feel often comfortable with PBE only. As all of us are referees too, we should be aware of this and try to set the right example (i.e. not objecting against papers that do not use PBE for a well-argued reason). But it might take a lot of time until the attitude of the entire community will have changed.

A bottleneck is that not all pseudopotentials are available for all exchange-correlation functionals. It is often not too complicated to make the transfer, but someone should do it. This will happen if a sufficiently large number of users asks for it, but this form a democratic demand is not the most efficient way for science to proceed.

Another issue is that many DFT codes cannot calculate all the (newest) features they provide for all functionals. LDA and PBE are usually dealt with for all features, but not always all other functionals. One could say: if users ask for it, developers will implement these features for another functional. That sounds, however, as 'scientific progress by democratic vote'. Science is not democracy.

Other fields of science might give an example about strategies our community could use. In medical science, there the phenomenon of an 'expert panel', that reviews the literature evidence on the best therapy for a given condition, and then formulates a recommendation. Medical doctors do not need to go through all literature themselves, they adopt only the recommendation. A board reviewing the practice at a hospital or by a medical doctor, checks whether the official recommendations have been followed. Translated to the electronic structure community, we could imagine an expert panel that formulates recommendations on 'what is the best functional to use for a given class of materials', or 'what is the best correction scheme for impurities in semiconductors', or 'what is the appropriate procedure to select numerical input settings', not based on their own research but based on an assessment of what has been published. Code users can follow this advice, and refer to it in publications. This will help to win over referees. On the other hand, referees can use these recommendations to check whether a manuscript uses the state of the art of choice of functional.

An effort about expert panels as described above, is difficult work to get funded and published. It can only fly if it would be a community effort. Psi-k could be the organizational level at which expert panels could be formed. The 2020 Psi-k conference could be a good occasion to brainstorm about this, and to get started.

Another way to make progress, rather than recommending the best functional, is to look at scaling relations: if the systematic deviation of two functionals for a given property/crystal class is documented, one can correct for this deviation to get closer to the experimental value. In this way, a fast and simple functional sometimes leads to results that are as good as a more expensive and more advanced functional (probably with the former having a large rescaling, and the latter a smaller rescaling). Codes could implement in their output not only the bare values, but also the rescaled

values, as 'recommended under the proper circumstances'. However, it will be a hard task to educate users to use these values in the right way, rather than cherrypicking what suits them best, perhaps for the wrong reasons. Another complication is that workflow managers may grep a value from the output (a bare or a rescaled one) without the corresponding warning. As increasingly many users do not see the code output as such any longer, but rather the parsed data by a shell or workflow manager, this may lead to delivering misleading information to the user. Alternatively, making the rescaling could be done at the level of the workflow manager, for any DFT code.

Some other aspects that were touched:

- There is often a trade-off between complexity of the functional and complexity of the physical properties. A sophisticated property can be developed on top of LDA/PBE, and that is very useful, even if LDA/PBE is not the best level for that crystal. Insisting on using a variety of better functionals (=better for easy properties) could slow down the development of more involved property modules.
- Experimentalists are used to spend most of their time to performing control experiments. But give experimentalists a DFT code, and they are happy with just the number that comes out of the box, without control calculations. We should invest more in education and tutorials.
- What do we call 'accuracy' in a broader context? It's only valid to compare with experiment if you measure exactly the same thing as experiment is measuring. For instance, if you don't include electron-phonon coupling in your band gap calculation, you cannot compare to experiment properly. The model system you examine should be sufficiently relevant before you can compare it to experiment in a meaningful way.

The bottom line: more efforts should be done to make the knowledge explicit that is implicitly known and used by some experts in the field. This should be made more visible and accessible for regular users and referees.

This is a digest of a dedicated discussion session held at PQ-DFT 2019. For other digests, videos of all talks and summarizing recommendations, please visit https://pqdft2019.abinit.org/. To access the videos directly on Youtube, visit http://bit.ly/2XFKUCl. Any comments, thoughts or items you want to discuss? Feel free to contact Stefaan Cottenier (stefaan.cottenier@ugent.be) or Kurt Lejaeghere (kurt.lejaeghere@ugent.be).