Continuous method of thermodynamic optimization using first-derivative matrices for large multicomponent systems

E. Nekhoroshev[1], *D. Shishin*[2], *M. Shevchenko*[3], *E. Jak*[4]

1.Theme Leader in Computational Thermodynamics, Pyrosearch group, University of Queensland, Brisbane, Qld, 4850. Email: e.nekhoroshev@uq.edu.au
2.Theme Leader in Thermodynamic Modelling, Pyrosearch group, University of Queensland, Brisbane, Qld, 4850. Email: d.shishin@uq.edu.au
3.Theme Leader in Experiments and Thermodynamic Optimization, Pyrosearch group, University of Queensland, Brisbane, Qld, 4850. Email: m.shevchenko@uq.edu.au
4. Professor of Chemical Engineering, Pyrosearch group, University of Queensland, Brisbane, Qld, 4850. Email: e.jak@uq.edu.au

Keywords: thermodynamic optimization, multicomponent systems, FactSage, ChemApp, CALPHAD, automation, machine learning, Python

ABSTRACT

Real-life applications require thermodynamic databases consisting of many elements. Currently, Pyrosearch supports a 20-component thermodynamic FactSage database for Cu-, Pb-, and Fe-production provided to a group of industrial sponsors. The database development follows an agile cycle of experiment planning, modelling/update of database subsystems, quality tests and release. A unique challenge is development and maintenance of a database self-consistent across all binary/ternary and multicomponent systems. Practice shows that binary and ternary system data do not provide the necessary accuracy for predictions within multicomponent systems. Commercial databases are often developed using only available literature data and it is hard to estimate the prediction quality for industry-relevant multicomponent compositions. Therefore, to sustain a high reliability, there is a need of targeted experiments, including multicomponent data.

To support the continuous development process, a new optimization methodology has been developed. After long years of optimization experience, it is an authors' belief that for optimization, it is essential to pinpoint a few trusted target points instead of using a whole cloud of available experimental data. In case if the optimization procedure is automated, using all data, even with trust weights, still leads to inferior optimization quality. This is caused by the fact that available optimization algorithms usually assume a normal error distribution of experimental data which can be false in cases when experimental results are affected by sample contamination, evaporation, incomplete reactions, etc.

Typically, iterative optimization involves multiple re-calculations of targets values which can take significant time and the whole procedure is not interactive. It is hard to find an exact combination of target weights to achieve desired outcomes. This problem has been solved by the authors by replacing slow FactSage/ChemApp calculations by parameter value/target value relations using first-order derivative matrices. Precalculated derivative matrices allow fast analysis of parameter influence across multiple systems, selection of interaction combinations with superior performance, and analytical matrix solution to the optimization problems using weighted linear least squares. The current optimization system supports in-sheet ChemApp calculations for Excel using custom Python binding code, graphical and statistical representation of experimental data and target points, and real-time first-derivative matrix analysis of parameter sensitivity.