[EURO Journal on Computational Optimization 9 (2021) 100012](https://doi.org/10.1016/j.ejco.2021.100012)

 

Contents lists available at [ScienceDirect](http://www.ScienceDirect.com/)

EURO Journal on Computational Optimization

journal homepage: [www.elsevier.com/locate/ejco](http://www.elsevier.com/locate/ejco)

(Global) Optimization: Historical notes and recent developments

Marco Locatelli [a](#_bookmark0), Fabio Schoen[b](#_bookmark1),[∗](#_bookmark3)

a *DIA - Università di Parma, Parco Area delle Scienze, 181/A, Parma - 43124, Italy*

b *DINFO - Università degli Studi di Firenze, via di S. Marta, 3, Firenze - 50139, Italy*

a r t i c l e i n f o a b s t r a c t

*Keywords:*

Global optimization Heuristics

Exact methods

Recent developments in (Global) Optimization are surveyed in this paper. We collected and commented quite a large number of recent references which, in our opinion, well represent the vivacity, deepness, and width of scope of current computational approaches and theoretical results about nonconvex optimization problems. Before the presentation of the recent developments, which are subdivided into two parts related to heuristic and exact approaches, respectively, we briefly sketch the origin of the discipline and observe what, from the initial attempts, survived, what was not considered at all as well as a few approaches which have been recently rediscovered, mostly in connection with machine learning.

# introduction

It seems worthwhile, in a special issue like this one, to devote a few words to the foundation of (global) optimization as an independent re- search topic. Looking back to the early days might prove useful, as some old ideas which, in the beginning, did not lead to much development, might prove to be interesting for current research, when considered from a modern perspective. Moreover, it is interesting to notice how some of the sub-fields which have seen a very large set of contributions in recent years, were almost totally neglected in the beginning. We refer in partic- ular to exact (global) optimization methods based on implicit enumer- ation and to the jungle of “nature inspired” population heuristics. For what concerns this kind of heuristics, we need of course to cite at least the book [Holland (1975)](#_bookmark111) where the basic ideas of genetic algorithms are nicely introduced. In what follows, we will first give a quick overview of the main approaches contained in the two classical books devoted to the subject in the 70’s ([Dixon and Szegö, 1975; 1978](#_bookmark67)). Then, we will give more details on recent developments. We also refer to our book [Locatelli and Schoen (2013)](#_bookmark112) for a detailed discussion about (global) op- timization topics, updated as of the date of publication of the book. After the historical introduction, the paper will consider two topics: recent, or re–discovered, heuristic approaches and recent developments in exact approaches. A final observation before the beginning of the paper: here and in the title we parenthesized the word (global). We will omit doing so in the paper, but we would like to observe that, although in the past the subject was considered somewhat exotic and off the main research streams, nowadays the richness of both theory as well as computational approaches gives to the subject a full recognition in the scientific com- munity. We might then propose, with a slightly provocative style, to

rename the whole subject simply as “optimization” – what else should we look for when optimizing, if not a global optimum?

The paper is structured as follows. After the brief introduction given in this section, in [Section 2](#_bookmark2) we recall the contents of the two books

edited in the70′s by Lawrence Dixon and Giorgio Szegö, which gave

a strong initial impulse to the whole discipline. In [Section 3](#_bookmark4) we re-

view some of the many heuristic computational approaches published in the literature in recent years (say, roughly after the publication of our book [Locatelli and Schoen (2013)](#_bookmark112), of which this paper might be consid- ered as a continuation). [Section 4](#_bookmark5) surveys recent literature dealing with structured optimization problems for which an exact procedure can be designed. In [Section 5](#_bookmark11) we briefly discuss some computational aspects and suggest sites where exhaustive lists of GO test problems and solvers can be found. Some concluding remarks are finally presented.

# On the origins of global optimization

Many early papers dealt with non convex optimization problems and outlined basic algorithms. We recall here, as particularly inter- [esting examples,](#_bookmark31) [Dantzig (1960)](#_bookmark45)[,](#_bookmark31) [McCormick (1972)](#_bookmark134)[, Beale and For- rest (1976),](#_bookmark31) [Falk](#_bookmark74) [and Soland (1969),](#_bookmark31) [Soland](#_bookmark145) [(1971),](#_bookmark31) [McCormick](#_bookmark135) [(1976),](#_bookmark31) [Horst (1976)](#_bookmark113). All of these papers had impact on the whole field, as well as many others we are not citing here. Despite the relevance of these as well as many other early papers, it can quite safely be assumed that the first “large scale” diffusion of the ideas of Global Optimization (denoted [by GO in what follows) can be credited to the two “orange” books Dixon and Szegö (1975, 1978). These two books, although clearly not the first](#_bookmark67) publications in GO, gave a fundamental impulse to the whole research field and since their publication, the term “Global Optimization” started to be a recognized and respected label, characterizing optimization the-

∗ Corresponding author.

*E-mail addresses:* [marco.locatelli@unipr.it](mailto:marco.locatelli@unipr.it) (M. Locatelli), [fabio.schoen@unifi.it](mailto:fabio.schoen@unifi.it) (F. Schoen).

<https://doi.org/10.1016/j.ejco.2021.100012>

Received 30 October 2020; Received in revised form 20 September 2021; Accepted 25 September 2021

2192-4406/© 2021 The Author(s). Published by Elsevier Ltd on behalf of Association of European Operational Research Societies (EURO). This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

ory and methods for nonconvex optimization problems. To insist on the fact that the field was just in its infancy at that time, it might be observed that only 9 of the 25 papers in the first volume and only 12 out of 24 in the second one were grouped in the “Global Optimization” chapters, all the others being local optimization papers, despite the title of the two books. It might be of interest to observe the topics which, at that time, were considered as the most promising ones. In this section we will briefly review some of the main ideas proposed at that time and comment on their success in the GO literature and practice.

We can group the first few papers on global optimization as follows:

Space covering. This area was related to methods aiming at implicitly exploring the whole feasible region. In particular, methods were described, mainly for 1–dimensional optimization, which, based on the knowledge of (an upper bound to) the Lipschitz constant of the objective function,

ror in the approximation of *𝑓 ⋆* = min*𝑥*∈*𝑆 𝑓* (*𝑥*). could be built with a guaranteed maximum er-

In the survey by [Dixon (1975)](#_bookmark68) the methods of [Evtushenko (1971)](#_bookmark72) and of [Shubert (1972)](#_bookmark182) (later known as Piyavskii-Shubert ([Piyavskii, 1972](#_bookmark167))) were described. In the following years those approaches generated a stream of research in Lipschitz–optimization, with some interesting ap- proaches but also severe limitations. In general, as nicely described in another milestone in the field, [Horst and Tuy (1993)](#_bookmark118), these can be seen as precursors and special cases of modern Branch & Bound algorithms. The idea of Branch & Bound for nonlinear optimization was not considered at all in the two original books.

Trajectory methods. At those times the idea of following the trajectory

of a suitable set of differential equations seemed to be very promising. Some of the proposed ap- proaches required locating saddle points in the boundary of the region of attraction of known local optima in order to be able to escape and explore new basins. A group of papers in the book, mostly from Joanna Gomulka and Gior- gio Treccani, explored this idea. In the follow- ing years this approach received less attention, although some similarities can be traced with some approaches based on stochastic differential equations and with some recent approaches in the neural network literature. A recent survey on developments in this algorithmic family can be found in [Alexandropoulos et al. (2020)](#_bookmark13).

Bayesian Optimization. At those times, GO problems with as few as ten

variables, with just box constraints, were con- sidered as the frontier in computational GO ap- proaches. Much of the effort then was devoted to very small scale problems; even at small scale, black-box problems were considered as very rel- evant. In fact, practical problems arise in which the objective function is not available in analyt- ical form, but can only be evaluated at specific points, possibly through an expensive procedure. Among the most interesting and profound ideas for small dimensional, possibly 1–dimensional, GO problems, Bayesian Optimization (BO) was particularly relevant. The main ideas of BO can be traced back to papers published in the 60’s by Harold Kushner (see, e.g., [Kushner (1964)](#_bookmark96)) and then generalized in [Mockus (1975)](#_bookmark145) and in many papers by the same author. The idea of

BO is fascinating: assume the objective function is a realization of a stochastic process, typically a Gaussian one. Then, given a prior distribu- tion on the possible sample paths of the pro- cess, after a few observations of the objective function, possibly affected by errors, have been performed, a posterior distribution can be com- puted, leading to a conditional stochastic model which is further updated as soon as the true ob- jective function is observed at new sample points. This stochastic model can be analyzed in order to answer, through suitable numerical algorithms, queries like: find the point at which the expected value of the function is minimum, or find the point at which the expected improvement over the current best observation is maximum. These queries are themselves GO problems, but, dif- ferently from the original one, possess a known analytical expression, from which gradients can be analytically computed; the global optimiza- tion required can be carried out with standard GO methods, without any necessity of excessive precision. It is worth observing that, although born in a stochastic framework, these methods are indeed deterministic. A nice and relevant gen- eralization and implementation of the BO ideas can be found in [Jones et al. (1998)](#_bookmark133), currently among the top cited papers in the whole GO lit- erature. Among the reasons for the great success of this paper we can mention the fact that, after a period in which BO almost disappeared from the literature, as a consequence of the high in- crease in complexity per iteration in the multi– dimensional case, the implementation of EGO de- scribed in [Jones et al. (1998)](#_bookmark133) recently became a standard for hyperparameter optimization in machine learning (see, e.g., [Frazier (2018)](#_bookmark84) for a recent survey). The field has considerably ex- panded in the last years and it is the subject of very active research.

Random Search. Methods based on some form of random sam-

pling were considered in the two books, as they are in general simple to implement and do not require first order information on the objective function. In [Gaviano (1975)](#_bookmark94) some general con- vergence results were proven; quite a few papers proposed random search algorithms, which, how- ever, did not find great attention in the following years. The field was quite productive for some time, and new approaches have been published in recent years. Some of those we consider as the most interesting ones are reviewed in our book [Locatelli and Schoen (2013)](#_bookmark112).

Clustering. This, at the time, seemed to be one of the bright- est and most innovative ideas. The eﬃciency of the most elementary GO method, Multistart (sample some random points and start a local search from each of them), is improved by care- fully selecting from which sample points it seems worth starting an expensive local optimization. Ideally, a single local search should be started from each basin of attraction (and, possibly, not from all of them, but just from the most “promising” ones). This idea was first proposed in [Becker and Lago (1970)](#_bookmark33) and then has been expanded and made available to a larger audi-

ence in [Törn (1978)](#_bookmark155). The idea of the proposed approach is both simple and powerful. First draw a uniform sample of feasible points. Then “con- centrate” the sample towards local optima either by temporarily discarding a fraction of highest value observations, or by performing a few de- scent steps. Finally, identify higher density re- gions (clusters), from each of which a single local search is started. The approach had immediately a great success and seemed to represent “the” GO method of choice, until a few years later, when it gradually got abandoned. There might be many reasons for the decline of clustering methods. We conjecture that this might be due to:

* the practical impossibility of dealing with large dimensional problems, say with more than 10 variables;
* the fact that the method could be used just to speed up Multistart, while more advanced algorithms appeared in the literature;
* the method became popular as it allows to save on local searches which, at the time, were very time consuming. With the advent of modern local optimization methods, al- though saving unnecessary local searches re- mains a good point, the eﬃciency of local search made those savings quite negligible in the overall method.

In recent years, some new methods have been [proposed (see](#_bookmark26) [Bagattini et al., 2018](#_bookmark25)[; Bagattini et al., 2019;](#_bookmark26) [Schoen](#_bookmark179) [and Tigli, 2021) which re-](#_bookmark26) visit these ideas trying to soften as much as pos- sible the above defects.

# Recent heuristic GO methods

In this section we would like to present a few recent papers dealing with modern GO algorithms. The papers cited in this section come either from the vast field of evolutionary approaches or have been stimulated and, in some cases, re-discovered thanks to the exponential growth in the interest around machine learning. We point out that the review of recent literature will be partial and biased towards our knowledge and preferences. It is really an impossible task to track all the papers dealing with GO heuristics, and even more diﬃcult to highlight those which are indeed relevant and truly innovative. But we tried to do our best in describing interesting new approaches in this wide field.

* 1. *Population-based, evolutionary methods*

An enormous quantity of papers dealing with variants of the basic evolutionary population schemes appeared in recent years. In the au- thors’ opinion it is quite disturbing that most of the papers in this sub- field justify themselves by some sort of an inspiration from nature, and, frequently, do not report any theoretical justification of the algorithmic choices, nor a fair and wide numerical comparison with state-of-the-art algorithms on well recognized benchmarks of test functions. A charac- teristic within this family of approaches is that they are mostly based on some variation of very basic schemes which include:

* a random generation of the initial set of solutions (initial popula- tion);
* a combination mechanism which takes parts of the components of the solutions represented in the current population and generates a new set of solutions;
* possibly a mutation, through which some solutions are randomly perturbed;
* possibly, in memetic algorithms, a local search applied to a selection of the elements in the current population;
* a substitution criterion, based on observed function values, by which from the original and the current population, a new set of solutions is built.

Some standard algorithms arose from this basic scheme. The most no- table, in GO, are Differential Evolution (DE) and Particle Swarm (PS), for whose description we address the reader to the rich literature on the sub- ject or to our chapter in [Locatelli and Schoen (2013)](#_bookmark112). Some recent sur- veys have been published on this subject (see, e.g., [Das et al. (2016)](#_bookmark46) or [Del Ser et al. (2019)](#_bookmark59)). In the recent literature some new proposals ap- peared like, e.g., [Cui et al. (2016)](#_bookmark42) or [Wu et al. (2016)](#_bookmark175), where multiple populations, with different evolution strategies, are evolved simultane- ously and compete among themselves to improve the best population. It seems that memetic variants of DE stand out as a good compromise between simplicity and quality. For DE we can even cite some theoreti- cal convergence results ([Ghosh et al., 2012; Locatelli and Vasile, 2015](#_bookmark96)) which can suggest guidelines for algorithm definition.

Among many variants of DE, those based on the exploitation of lo- cal optimization (memetic variants) are very interesting, as many GO [problems allow for fast and reliable local search tools. In Cabassi and Lo- catelli (2016) an analysis of some variants of the basic memetic DE is in-](#_bookmark31) troduced and numerically shown to be very effective on a wide range of GO test problems of varying dimension. In [Schoen and Tigli (2021)](#_bookmark179) those methods have been extended and mixed with a clustering approach in order to save unnecessary local searches; numerical results show that it is possible to significantly improve the eﬃciency of those population– based methods while keeping their good quality. On a different line of research, in [Mansueto and Schoen (2021)](#_bookmark123) a DE–based memetic ap- proach is used to build an eﬃcient GO method for optimal clustering in Euclidean spaces. In that paper a specialized local search is used, based

on the well known *𝐾*–means clustering algorithm, coupled with the ex-

ploration capabilities of DE.

A source for many evolutionary algorithms, as well as test problems, can be found in the various CEC (the IEEE Congress on Evolutionary Computation) competition websites. It appears that, in those competi- tions, a dominant role is played by population–based methods which in some way exploit the separability or partial separability of prob- lems. The idea of, at least partially, decompose a GO into sub-problems is adopted by quite a large number of successful approaches in those competitions. Of course, for separable problems it is easy to forecast their good performance, but numerical results seem to be quite inter- esting also for non separable ones. The paper [Ma et al. (2019)](#_bookmark120) con- tains a survey on decomposition–based methods and presents the ba- sic ideas of these approaches. There are many variants of the basic scheme, which typically differ in the decomposition strategy and the recombination one; many are based on a Gauss-Seidel–like strategy in which optimized subsets of variables are fixed and used while optimiz- ing different subsets. We can cite here, as an example of an eﬃcient GO approach, [Hadi et al. (2019)](#_bookmark100), where, starting from an adaptive DE scheme in which some hyperparameter is adjusted during the evolu- tion, a hybrid, decomposition–based, memetic approach is proposed. In particular, at some stage of the computation, variables are randomly grouped and different local optimization algorithms are associated to each sub-group of variables. The winner of the 2019 CEC competition [Sun et al. (2019)](#_bookmark147) uses an innovative strategy to decompose non separa- ble problems which leads to a highly eﬃcient method. Their decompo- sition scheme is based on the identification of subsets of variables which partially overlap one another.

Finally, we would like to mention the family of CMA-ES (Covari-

ance Matrix Adaptation Evolutionary Strategies) algorithms (see, among many papers on the subject, [Hansen and Ostermeier (2001)](#_bookmark102) for an early introduction to the approach, or [Hansen (2006)](#_bookmark101) for a survey on variants of the basic method). The idea behind CMA-ES is that of having a popula- tion of solutions to a GO problem which evolves through sampling from

a multinomial distribution whose mean and covariance matrix evolve during the iterations and in some sense adapts to the level sets of the objective function. The objective function is evaluated at each sample point and a new mean and covariance matrix are generated through an updating mechanism. Oversimplifying a description of this family of approaches, which is indeed quite vast, we might say that they try to adapt search directions and step sizes in order to favour changes in the current population members which are likely to contribute much to the improvement of the objective function. This is usually obtained by following directions associated to the principal components of the co- variance matrix, which is built in such a way as to adapt to observed function values. In [Diouane et al. (2015a,b)](#_bookmark65) an interesting extension of the basic scheme is presented in which, imposing a criterion of suﬃ- cient decrease in the objective function, some convergence properties are obtained.

* 1. *Basin-Hopping methods*

Multistart is one of the simplest GO methods, where a local search is started from each point randomly generated within the feasible region of a GO problem. It is usually considered a low eﬃciency algorithm for GO due to the computational waste it produces in rediscovering the same local optima more than once and due to its absence of any learn- ing mechanism (see [Section 3.4](#_bookmark7) for further comments on these topics). However a relatively simple modification of Multistart, which goes un- der the name of Basin Hopping (BH), or Iterated Local Search, is quite an interesting approach for many hard GO problems. As in Multistart, also in BH local optimization is performed starting from a random ini- tial point. However in BH, after a local optimum is found, new local searches are performed starting from a (suitably defined) neighborhood of the current one until possibly a better local optimum is found; in this case new local searches are performed starting from a neighborhood of the new local minimum, and the whole procedure is repeated until some stopping rule is satisfied. The version we sketch here is the monotonic version of the method: non monotonic ones have also been successfully implemented. BH has been used in [Vinkó and Gelle (2017)](#_bookmark164) as a tool to build a graph of neighboring local optima, much in the sense of what was described in Chapter 3 of our book [Locatelli and Schoen (2013)](#_bookmark112). Given a BH algorithm with specific parameters, a run on a specific test function generates a set of local optima which can be pairwise connected in a graph in which an oriented arc exists between two local minima if the second one has been reached through a BH step from the first one. The resulting graph associated to some classical test functions can be an- alyzed. It might be possible to exploit this information in order to build advanced BH methods, but this is still a subject of current research. A hybrid approach, mixing Differential Evolution with BH is presented in [Di Carlo et al. (2020)](#_bookmark63), where a simple criterion is presented to save un- necessary local searches.

Most recent literature on BH and its variations, including

population–based variants as discussed, e.g., in [Grosso et al. (2007)](#_bookmark99), deals with application of these methods. Many of those applications are in computational chemistry, a field where BH was born and has found a prominent space in current research. As an example, in [Zhao et al. (2017)](#_bookmark186) a modified BH is applied to the optimization of atomic clusters. Among the many variations, in this paper it is sug- gested that some variables are kept fixed at certain iterations (those corresponding to the location of specific subsets of atoms), while some others are subject to perturbation in BH steps. Moreover, some accep- tance criteria different from the improvement of the objective function are suggested, like the exploration of new geometrical configurations. In [Ferreiro-Ferreiro et al. (2019)](#_bookmark78) a simple modification of the BH scheme is introduced and tested on specific atomic clusters. The idea is to let BH explore deeper basins by reducing its “greediness”. To this aim, at each iteration, several neighboring local optima are generated and, instead of moving towards the first improving one, the best is chosen as the starting

point of the next iteration. In [Wales (2018)](#_bookmark169) variants of BH are described in the context of energy landscapes for atomic clusters. Different accep- tance rules are proposed that differ from the classical monotonic and Metropolis-like acceptance rules, generalizations to multiple objectives in the exploration phase are introduced, and further refinements based on molecular dynamics are proposed.

* 1. *Methods for expensive black-box objective functions*

This field has recently attracted enormous attention, driven by ma- chine learning research, as algorithms within this framework are con- sidered as especially suitable for hyperparameter calibration in training machine learning tools like, e.g., deep learning architectures. A survey on the subject of optimization–based hyperparameter choice in machine learning recently appeared in [Yang and Shami (2020)](#_bookmark179), where a list of available software is also provided. Also, in [Tran et al. (2020)](#_bookmark162) a dis- cussion on the use of optimization tools in hyperparameter tuning is presented. In applications for hyperparameter setting, two levels are usually present. At the lower level, given a training set, learning pro- ceeds by suitably choosing (in neural networks, as an example) con- nection weights in such a way that a loss function is minimized; at the upper level the loss function itself, the overall architecture and, some- times, the optimization algorithm used for training, depend on a rela- tively small number of hyperparameters (e.g., the parameter associated to regularization, the number of layers, the “learning rate”,...). These hyperparameters are usually optimized considering the performance of the lower level trained system on a validation set, different from the one used for training. In general, thus, the objective function to be op- timized at this level is neither known, nor cheap, and is modeled as a black box. GO methods for expensive cost functions are among the primary choices for this tuning. We will not even try to summarize the overwhelming quantity of reports recently published on this sub- ject on arXiv. Although excellent research is sometimes found in that dynamically exploding repository, here we chose to refer exclusively to papers published in high quality refereed journals. A nice survey of state-of-the-art approaches in Bayesian Optimization can be found in [Frazier (2018)](#_bookmark84), while in [Zhan and Xing (2020)](#_bookmark183) several different objec- tive functions based on variation of the originally proposed Expected Improvement are presented and discussed. In [Mathesen et al. (2020)](#_bookmark130) an interesting generalization of Bayesian Optimization models is proposed, in which the “acquisition function” takes into account the desire to im- prove the exploration capabilities, as the exploitation is delegated to a trust-region based local search which takes into account the available budget of function evaluations. It is interesting how, even in an expen- sive function setting, the idea of performing local optimization can be effectively exploited. In [Ahmed et al. (2020)](#_bookmark14) it is suggested to use an estimate of the Lipschitz constant in order to improve the search for the next observation point, an approach which, at least for moderately sized problems, might have some relevance. Instead, in [Bemporad (2020)](#_bookmark38) the proposed approach is not based on a stochastic model of the objective function, but, taking inspiration from those models, the author proposes a surrogate function composed of three elements: an interpolation of ob- served sample values, a distance from the points in the sample, distance of the objective function values. This acquisition function is optimized in order to decide where the next observation should be placed.

A recent stream of research has started exploring the possibility of

applying the idea of Bayesian Optimization to large scale GO prob- lems. The main tool used to greatly increase the dimension of problems which can be solved through this approach is the idea of *random em-* [*beddings*. The philosophy of these approaches (see Binois et al., 2020; Moriconi et al., 2020; Wang et al., 2016) relies on the assumption that](#_bookmark52) for large dimensional problems there exists a true “low effective dimen- sionality”, which can be considered as a low–dimensional linear em- bedding such that the objective function does not change when mov- ing along directions which are orthogonal to this embedding. A differ-

ent approach to large-scale adaptation of these methods is reported in [Snoek et al. (2015)](#_bookmark187), where a neural network is used to learn a set of basis functions to be used to build a Bayesian linear regressor which substitutes the computationally expensive classical one.

Among the most interesting approaches in the Bayesian Optimization literature, it seems worth citing extensions to the constrained case and, in particular, to the case in which also the constraints are expensive, black-box, functions. [Hernández-Lobato et al. (2016)](#_bookmark108) present a detailed survey of constrained Bayesian Optimization methods. A novel frame- work is also introduced for the case where at each iteration a new sample point needs to be generated at which both the objective function as well as the constraints have to be evaluated. However, in the proposed ap- proach, it is assumed that functions (objective and constraints) might be at least partially separable into parts (so called “tasks”) – as a simple example, the objective function might be evaluated on a CPU while the constraints are defined in a way that enables to exploit the parallelism of a GPU. In these cases, the merit (or “acquisition”) function which guides the search for the next evaluation point, should be defined in such a way as to support a similar decomposition. In that paper, a novel acquisi- tion function, called PESC (Predictive Entropy Search with Constraints) is proposed as particularly well suited in this framework. Constrained Bayesian Optimization is also considered in [Feliot et al. (2017)](#_bookmark76), mostly in the context of an extension of the basic approach to multi-objective optimization, an interesting topic which is, however, out from the scope of this paper.

[On a different line of research, in Garrido-Merchán and Hernández-](#_bookmark89)

[Lobato (2020) the problem is considered of adapting the approach to in-](#_bookmark89) teger and categorical variables, both ordered as well as unordered. The interest on such a kind of extension has been greatly stimulated by the application of GO to hyperparameter calibration in machine learning, where, e.g., in a neural network, one might wish to calibrate the num- ber of neurons, of layers, of filters on a convolution, which are integer– valued parameters. Otherwise hyperparametrs can be associated, e.g., to the activation function at a neuron, or to the loss function, or to the kernel function used in a Support Vector Machine; each of these can be seen as an example of categorical unordered parameter. The field has seen much expansion and interest in the scientific community and pa- pers started to appear describing novel software implementations, like, e.g., [Kandasamy et al. (2020)](#_bookmark136); [Martinez-Cantin (2014)](#_bookmark131).

We conclude this subsection by citing [Kim (2020)](#_bookmark89), a recent theo- retical paper which provides a rational explanation on how Bayesian Optimization algorithms proceed by alternating exploitation, or learn- ing, phases with exploration, or optimization. Exploitation is connected with a greedy cost optimization strategy, while exploration is associated to variance regularization. The whole process, analyzed in the context of Dynamic Programming, is analyzed and it is shown how the temporal discount factor of a Markov Decision Process plays the role of trade–off parameter between exploration and exploitation.

* 1. *Methods on the intersection between machine learning and GO*

The title of this subsection refers to methods in which the search towards a global optimum, or, at least, a good local one, is guided by some form of “learning”, much in the spirit of modern machine learn- ing approaches. In this part we would like to include the re-discovery of clustering methods. These methods, as recalled in [Section 2](#_bookmark2), were based on the idea of learning, from the sample, the shape of differ- ent basins of attractions of local optima, before starting expensive lo- cal searches. Clustering is one of the pillars of unsupervised learning and it might be of interest to look to those methods as early imple- [mentations of machine learning tools for GO. In Bagattini et al. (2018, 2019) and,](#_bookmark25) [Schoen](#_bookmark179) [and Tigli (2021) we proposed different strategies in](#_bookmark25) order to overcome the limitations of original clustering methods. In par- ticular, we showed that the idea of those algorithms can be successfully applied to much more refined algorithms than the standard Multistart,

by showing how a memetic Differential Evolution variant, inspired by [Cabassi and Locatelli (2016)](#_bookmark31), can profitably save a very large number of useless local searches while maintaining the quality of the original. Moreover, we showed how to use the same ideas with methods based on local search methods which are, in a sense, more refined and “global” than standard, gradient-based, local optimization tools. Finally, some experiments on random projections enabled us to extend the approach to high dimensional problems.

It is worth noticing also that in the field of chemical physics many approaches are proposed in which GO methods (in particular, variations of basin hopping) are guided by knowledge on the prob- lem domain obtained through machine learning. As an example, in [Meldgaard et al. (2018)](#_bookmark137), low dimensional features are extracted from the original variables of atomic clustering problems and regression is used to assign an energy contribution to each atom in order to let GO algorithms focus on the parts of the current solution which contribute most to the objective function.

We conclude this section citing a paper which does not belong to the literature on algorithms but is quite relevant to the subject. The paper [Kawaguchi (2016)](#_bookmark137) deals with a conjecture on the absence of local optima which are not global when using a quadratic loss in training a deep neural network. The results proven there are interesting, as the problem is neither convex nor concave, yet it seems that training with a local optimization algorithm can always lead to a global optimum solution, thus partially explaining the success of local methods in training neural networks.

# Exact GO methods

While in the previous section we discussed heuristic approaches, in this section we deal with exact GO methods. We will first discuss ”easy” GO problems, i.e., GO problems which can be reformulated as (tractable) convex problems or, at least, can be solved in polynomial time with re- spect to the size of the problem data and the inverse of the required pre- cision (we refer, e.g., to [Ben-Tal and Nemirovski (2001)](#_bookmark43) for a detailed discussion about computational complexity in the context of continu- ous optimization problems and, in particular, in the context of convex optimization problems). Later on, we will discuss exact GO methods (ba- sically, branch-and-cut approaches) which are usually applied to ’highly structured’ GO problems. We will often refer to Quadratic Programming (QP) problems, which arise in many different contexts. Applications of QPs include the reformulation of some combinatorial optimization prob- lems, like, e.g., max-clique and max-cut, portfolio optimization, packing problems, blending and pooling problems. More applications of QPs can be found in [Furini et al. (2019)](#_bookmark85), where a QP library is introduced.

* 1. *“Easy” GO problems*

Global optimization is, in general, a hard task. Even highly struc- tured GO problems, like, e.g., Standard Quadratic Programming (StQP), where a quadratic form is minimized over the unit simplex, have been proved to be NP-hard. All the same, there are some classes of GO prob- lems, in particular some QP problems, solvable in polynomial time. These include problems having the so called *hidden-convexity* property: though not convex, these problems can be reformulated as (tractable) convex problems. The best known of such problems is the *trust region* problem, where a quadratic function is minimized over the unit ball. It has been shown in [Rendl and Wolkowicz (1997)](#_bookmark172) that this problem [admits a semidefinite reformulation. In fact, in Ben-Tal and den Her- tog (2014) it has been shown that the trust region problem and, ac-](#_bookmark40) tually, a generalization where the unit ball constraint is replaced by a more generic quadratic constraint, can be reformulated as a sim- pler convex conic quadratic problem under the assumption that the Hessian matrices of the objective and constraint functions are simul- taneously diagonalizable. After transforming both the objective and the

tained by first introducing new variables *𝑦𝑖* and the related constraints constraint into separable functions, the conic quadratic problem is ob-

*𝑦𝑖* = *𝑥*2, and then relaxing these equalities into inequalities *𝑥*2 ≤ *𝑦𝑖* .

class of QCQPs which includes the diagonal ones, and it is shown that the condition implies the result proved in [Burer and Ye (2020)](#_bookmark30) when applied to diagonal QCQPs. In [Jeyakumar and Li (2018)](#_bookmark128) a class of min-

*𝑖 𝑖*

The work [Jiang et al. (2018)](#_bookmark129) proposes a Second Order Cone Program- ming (SOCP) reformulation for this problem without requiring the as- [sumption of simultaneous diagonalization. The paper Wang and Kilinç- Karzan (2020) proposes a convex reformulation in the original space](#_bookmark171) of variables. Further problems with the property of hidden-convexity [have been introduced in the last two decades. In Pong and Wolkow- icz (2014) it is shown that, under suitable assumptions, a semidefinite](#_bookmark170) relaxation for the problem of minimizing a quadratic function over a region defined by a two-sided quadratic constraint (i.e., a lower and upper bound are imposed over a quadratic function), is exact. For the same problem [Jiang et al. (2018)](#_bookmark129) presents a SOCP reformulation. The work [Beck and Teboulle (2009)](#_bookmark32) introduces a semidefinite reformulation for the problem of minimizing the ratio of two quadratic functions over a possibly degenerate ellipsoid. In [Burer and Anstreicher (2013)](#_bookmark79) and [Sturm and Zhang (2003)](#_bookmark146) it has been proven that problems with a quadratic objective function, a unit ball constraint and a single lin- ear cut can be reformulated as semidefinite problems with an addi- tional SOC-RLT (Second Order Cone - Reformulation Linearization Tech- [nique) constraint. The result has been extended in Burer and Anstre- icher (2013) to the case of two parallel linear cuts and later on, in](#_bookmark79) [Burer and Yang (2015)](#_bookmark88), generalized to the case of an arbitrary num- ber of linear constraints, provided that such constraints do not intersect in the interior of the unit ball. For problems with a quadratic objective function and two quadratic constraints, in [Ye and Zhang (2003)](#_bookmark180) it is proved that if all the quadratic functions are homogeneous (i.e., there are no linear terms), then an exact SDP relaxation exists.

Many papers in the literature provide conditions under which some

convex relaxation of a class of nonconvex problems turns out to be exact. For QCQP problems (problems where objective and constraint functions are all quadratic), the work [Kim and Kojima (2003)](#_bookmark91) provides some sign conditions about the data under which SDP and SOCP relax- ations are exact. The paper [Sojoudi and Lavaei (2014)](#_bookmark144) considers QCQP

problems without linear terms, i.e., with objective function *𝑥⊤𝐴*0*𝑥* and

constraints *𝑥⊤𝐴𝑘𝑥* ≤ 0, *𝑘* = 1*,* … *, 𝑚* (but it also discusses a way to in-

clude also linear terms). For these problems a graph  is built with

*𝑛* nodes (one for each variable) and an edge (*𝑖, 𝑗*) exists if and only if

*𝐴𝑘* ≠ 0 for some *𝑘* ∈ {0*,* 1*,* … *, 𝑚*}. Next, exactness of SDP and SOCP re-

*𝑖𝑗*

graph. For instance, exactness holds if for each edge (*𝑖, 𝑗*) all entries *𝐴𝑘* , laxations are related to some sign conditions and to the structure of this

*𝑖𝑗*

*𝑘* ∈ {0*,* 1*,* … *, 𝑚*}, have the same sign, and graph  is acyclic. The paper

[Jeyakumar and Li (2014)](#_bookmark126) considers the problem of minimizing quadratic

functions over a feasible region defined by a ball constraint and lin- ear constraints. A so called dimension condition is introduced under [which the SDP relaxation turns out to be exact. Following Ben-Tal and den Hertog (2014), in](#_bookmark40) [Locatelli](#_bookmark109) [(2015b, 2016a) a simpler convex conic](#_bookmark40) quadratic relaxation is considered and in [Locatelli (2016a)](#_bookmark110) it is shown that such relaxation is equivalent to the SDP one. Then, a condition for exactness of the relaxation more general than the dimension condition is derived from the KKT conditions of the convex conic quadratic relax- ation. The paper [Ho-Nguyen and Kilinç-Karzan (2017)](#_bookmark110) considers a SOCP relaxation in the original space of variables for problems with quadratic

objective function, a unit ball constraint and constraints *𝐴𝑥* − *𝑏* ∈ G,

where *𝐴* ∈ ℝ*𝑚*×*𝑛*, *𝑏* ∈ ℝ*𝑚*, and G *⊂* ℝ*𝑚* is a closed convex cone. The pa-

given the epigraph of the problem, with the additional variable *𝑡* and the per introduces conditions under which the relaxation is tight. Moreover, additional constraint *ℎ*(*𝑥*) ≤ *𝑡*, where *ℎ* is the objective function, condi-

[tions are provided for the derivation of its convex hull. In Burer and Ye (2020) diagonal QCQPs are considered, where all the Hessian ma-](#_bookmark30) trices of the quadratic functions are diagonal. In this case the work provides some conditions related to the feasibility of suitably defined polyhedral sets, which guarantee the existence of rank-one solutions of the SDP relaxations and, thus, exactness of the relaxation. The work [Wang and Kilinç-Karzan (2021)](#_bookmark172) presents an exactness condition for a

imax QCQPs is addressed (the objective is the maximum of a finite set of quadratic functions). A SOCP reformulation of the Lagrangian dual is presented. Exactness of this SOCP problem is proved when the epi- graphical set (the epigraph of all the quadratic functions involved in the objective and in the constraints) is closed and convex. The Celis– Dennis–Tapia (CDT) problem, where a quadratic function is minimized over the intersection of two ellipsoids, has been investigated in some works. For instance, [Ai and Zhang (2009)](#_bookmark15) gives a necessary and suf- ficient condition for the exactness of the Lagrangian relaxation for this problem. Different papers investigate the possibility of narrowing [(or closing) the duality gap by adding SOC-RLT constraints Burer and Anstreicher (2013), by solving two subproblems with SOC constraints](#_bookmark79) [Yuan et al. (2017)](#_bookmark181), by adding lifted RLT cuts [Yang and Burer (2016)](#_bookmark178), by adding KSOC cuts [Anstreicher (2017)](#_bookmark19) (we refer to [Section 4.2.4](#_bookmark10) for a discussion of all these cuts). It is also worthwhile to mention the re- sult in [Yang et al. (2018)](#_bookmark177) stating that the addition to QCQPs of further reverse convex constraints, imposing that feasible points cannot lie in the interior of non-intersecting ellipsoids, does not lead to more diﬃcult problems. More precisely, the existence of a tight SDP relaxation for the problem without such constraints implies the existence of a tight SDP relaxation also for the problem with these constraints.

For some problems exact convex reformulations are not known but

still the problem can be solved in polynomial time. The polynomial methods usually enumerate all the (polynomially bounded) KKT points for these problems. An example of such enumerative methods is pro- posed in [Bienstock and Michalka (2014b)](#_bookmark51) for problems with a quadratic objective function, some ball constraints, some reverse ball constraints (i.e., constraints which impose that points cannot lie in the interior of a ball), and some linear constraints. The method runs in polynomial time provided that the number of ball and reverse ball constraints is fixed, and that the number of faces of the polyhedron defined by the set of lin- ear inequalities having a nonempty intersection with the set defined by [the ball constraints is polynomially bounded. The papers Consolini and Locatelli (2017) and](#_bookmark39) [Sakaue](#_bookmark176) [et al. (2016) provide a polynomial-time](#_bookmark39) method to solve a generalization of the CDT problem, where one of the two quadratic constraints is allowed to be nonconvex. The proposed

variate polynomial system with polynomials of degree at most 2*𝑛* and method identifies all the KKT points of the problem by solving a bi-

with the two unknowns corresponding to the Lagrange multipliers of the two quadratic constraints. The algorithm has a polynomial com-

plexity but with a large exponent (six) with respect to the number *𝑛* of

variables. Interestingly, a rather different polynomial-time approach for

this problem has been presented in [Bienstock (2016)](#_bookmark47). In this approach a sequence of feasibility problems for systems of quadratic inequalities is solved by a polynomial-time algorithm based on Barvinok’s construction [Barvinok (1993)](#_bookmark30). The algorithm is hard to implement but the approach can also be extended to any fixed number of quadratic constraints, pro- vided that one of them is strictly convex.

We conclude this section by mentioning problems for which a polynomial-time solution algorithm is not available but for which [polynomial-time approximation schemes are available. In Bomze and de Klerk (2002) a Polynomial Time Approximation Scheme (PTAS) is](#_bookmark65) introduced for the StQP problem. The PTAS is based on the evaluation of the objective function over a uniform grid. In [de Klerk et al. (2006)](#_bookmark54),

degree *𝑑* over the unit simplex. In [Depetrini and Locatelli (2011)](#_bookmark61) a Fully the result has been extended to the minimization of polynomials of fixed

Polynomial Time Approximation Scheme (FPTAS) has been proposed for Linear Fractional-Multiplicative Programming problems, where sums or products of a fixed number of ratios of aﬃne functions are minimized over polytopes. The FPTAS is based on the solution of LP problems over a nonuniform grid. The approach has also been extended to a more [general class of problems in](#_bookmark144) [Locatelli (2013)](#_bookmark104) [and Mittel and Schulz (2013).](#_bookmark144)

* 1. *Exact methods for “difficult” GO problems*

Branch-and-Cut (B&C) methods are most widely employed for the exact solution of GO problems. In what follows we briefly sketch how they work. A collection C of subsets of the feasible region is maintained throughout the algorithm. The collection is initialized with a single set corresponding to the whole feasible region. Then, during the execution of the algorithm the problem is subdivided into subproblems by *branch- ing* operations. Each branching operation replaces a subset in C by other subsets which cover it. For each subproblem a *relaxation*, often a con- vex relaxation, is defined whose solution gives a *lower bound* for the subproblem. The relaxation can be strengthened through the introduc- tion of *cutting planes*, which are guaranteed not to remove feasible or, at least, optimal solutions. A global *upper bound* is possibly updated each time a feasible point is detected, e.g., as a result of the solution of the convex relaxation. *Fathoming* rules are applied, and, in particular, all subproblems in the collection C whose lower bound is not lower than

the current upper bound (possibly decreased by a tolerance value *𝜀*) are

removed from the collection C. The algorithm stops as soon as the col-

lection C is empty. In some cases, a further collection g is maintained. Such collection contains all subsets over which the lower bound is not higher than the current upper bound. Then, when the algorithm stops (i.e., when the collection C is empty), the subsets in the collection g

the optimal value by at most the tolerance value *𝜀*. The theoretical is- contain all feasible solutions whose objective function value differ from

sue of the finiteness of B&C methods has been widely investigated in [the past and for this we refer to GO textbooks like Horst and Parda- los (1995),](#_bookmark115) [Horst](#_bookmark116) [et al. (2001) and](#_bookmark115) [Horst](#_bookmark118) [and Tuy (1993). In what follows](#_bookmark115) we discuss the most recent developments related to the main operations of B&C approaches emphasized above. Note that we are not going to discuss operations like upper bounding and fathoming since these are rather standard operations for which significant developments cannot be expected. For what concerns fathoming we only observe that, be- sides the standard rule based on the comparison between lower and upper bounds, in some cases it is possible to introduce fathoming rules based on optimality conditions. In particular, one can remove subsets from C for which it is possible to guarantee that they do not contain points fulfilling necessary optimality conditions.

* + 1. *Branching*

Branching operations can be subdivided into two broad classes, *spa- tial branching* and *KKT branching*. Spatial branching can be applied to generic GO problems. The feasible region of the GO problem is initially enclosed into a region with a simple geometrical form. Then, branch- ing is performed by subdividing this region into smaller regions, often, but not necessarily, with the same geometrical form. The intersections of these smaller regions with the original feasible set give rise to the subsets entering the collection C. The most common geometrical form is a box, since lower and upper bounds for the variables are often already part of the GO problem description or, alternatively, can be easily computed by solving auxiliary problems. But other geometrical forms have been adopted in the literature. For instance, simplices and polyhedral cones [have been often employed, see again the textbooks (Horst and Pardalos, 1995; Horst et al., 2001; Horst and Tuy, 1993). Note that in all these](#_bookmark115) cases (boxes, simplices, polyhedral cones) the branching operation sub- divides a subset into smaller subsets, whose interiors do not overlap, but [which can share some common face. Some papers (Cartis et al., 2015; de Angelis et al., 2004; Fowkes et al., 2013; Hager and Phan, 2009; Le Thi, 2000) employ ellipsoids, exploiting the fact that quadratic func-](#_bookmark35) tions and also some cubic functions can be eﬃciently minimized over these sets. In this case the subsets generated by a branching operation may have overlapping interiors. We also mention [Tóth et al. (2016)](#_bookmark158), where a subdivision into *regular* simplices is proposed, which also leads to subsets with overlapping interiors.

For some specific problems further geometrical forms have been con- sidered. For QCQP problems, [Linderoth (2005)](#_bookmark101) employs cartesian prod-

ucts of rectangles and right-angled triangles. For linear sum-of-ratios problems in [Kuno (2005)](#_bookmark95) cartesian products of trapezoids are used, while for the same problem [Locatelli (2015a)](#_bookmark106) employs cartesian prod- ucts of rectangles and right-angled triangles as in [Linderoth (2005)](#_bookmark101). The use of these geometrical forms is strictly related to the development of tight under- and over-estimators for bilinear and bivariate fractional functions over the two-dimensional regions appearing in the cartesian products, i.e., rectangles, right-angled triangles and trapezoids.

The subdivision can be performed either in a problem-independent or in a problem-dependent way. In the former a subregion is subdivided into subregions of equal size. The most common problem-independent subdivision is bisection, where the subregion, say a box, is split into two subregions through a subdivision of the longest edge performed at its midpoint (for ellipsoids the longest edge is replaced by the longest axis). Instead, problem-dependent subdivisions take the solution of the convex relaxation into account. If the objective function is underestimated by a convex function, the subdivision is performed in such a way that the error at the optimal solution of the convex relaxation, i.e., the difference between the objective function and the convex underestimating function evaluated at such optimal solution, is reduced as much as possible in the newly generated subregions. Note that it is important to reduce the error at the optimal solution of the convex relaxation, since the value of the convex underestimating function at this point is equal to the lower bound computed over the subregion. Then, reducing the error usually means improving (increasing) the lower bound of the newly generated subregions with respect to the lower bound over the original subregion. KKT branching has a more limited applicability with respect to spa- tial branching. It is employed for QP problems with linear constraints. In particular, for Box QP problems (problems with a quadratic objec- [tive function and box constraints) in Vandenbussche and Nemhauser (2005a,b) it is observed that these problems can be reformulated by re-](#_bookmark163) placing the original feasible region with the one defined by the KKT conditions. This requires the addition of the variables corresponding to the Lagrange multipliers of the box constraints, while the original quadratic objective function can be replaced by a linear one, involv- ing also the additional variables. The only nonconvex constraints of the reformulation are those corresponding to the complementarity condi- tions. These are initially omitted, while KKT branching first selects a constraint according to some rule (usually the one with the largest vi- olation of the complementarity condition at the solution of the convex relaxation), and then splits the current subregion into two new subre- gions by imposing that in one subregion the constraint is active, while in the other the corresponding Lagrange multiplier is set equal to 0. The work [Burer and Vandenbussche (2008)](#_bookmark86) extends the approach to feasi- ble regions which are general polytopes. While the idea is the same, the simple approach of omitting the complementarity conditions in order to get a convex (linear) relaxation does not work in the general case [since it leads to a trivial bound. Therefore, in Burer and Vandenbuss- che (2008) an SDP relaxation is introduced. A different SDP relaxation](#_bookmark86) has been proposed in [Burer and Vandenbussche (2009)](#_bookmark87) for Box QP prob- lems with further developments in [Chen and Burer (2012)](#_bookmark36), which reports very good computational results for problems with a dense Hessian ma- trix. KKT branching is also employed in [Audet et al. (1999)](#_bookmark24) for a special QP case, the case of disjoint bilinear programming. It is also worthwhile to point out here that the reformulation of QP problems based on KKT conditions has also been used to convert the solution of these problems into the solution of Mixed Integer Linear Programming (MILP) prob- lems, where additional binary variables are used to establish whether a constraint is active or, alternatively, its Lagrange multiplier is equal to zero (see, e.g., [Xia et al. (2020)](#_bookmark176) where good results, especially for StQP

problems, are reported).

* + 1. *Convex underestimating functions*

Convex underestimating functions or, analogously, concave overesti- mating functions, are essential for the definition of convex relaxations. The ability of detecting them and evaluating their tightness, i.e., how

close they are to the original function, strictly depends on the proper- ties of the original function and on the region over which the underesti- mation takes place. For some functions and regions of relatively simple

*𝑖,𝑗*=1*, 𝑖*≠*𝑗*

by inequalities where the left-hand side is imposed to be not lower than the convex envelope (over a suitable rectangle or interval) of the right- hand side, and not larger than its concave envelope. A simple example

form it is possible to compute the tightest convex underestimator, also

of factorable function is the sum of bilinear terms ∑*𝑛*

*𝐴𝑖𝑗*

*𝑥𝑖 𝑥𝑗* . A

known as convex envelope. More formally, given a function *𝑓* and a re-

gion *𝑋*, the convex envelope of *𝑓* over *𝑋* is defined as follows for each

*𝑥* ∈ *𝑋*:

*𝑐𝑜𝑛𝑣𝑓,𝑋* (*𝑥*) = sup{*𝑐*(*𝑥*) ∶ *𝑓* (*𝑦*) ≥ *𝑐*(*𝑦*)*,* ∀*𝑦* ∈ *𝑋, 𝑐* is convex}*,*

i.e., *𝑐𝑜𝑛𝑣𝑓,𝑋* is the largest convex underestimator of *𝑓* over *𝑋* (note that in the above definition ’*𝑐* is convex’ can be replaced with ’*𝑐* is

aﬃne’). Analogously, we can define concave envelopes, i.e., tightest

vex envelopes, since the concave envelope of *𝑓* over *𝑋* can easily be concave overestimators. In what follows we will always refer to con- seen to be the opposite of the convex envelope of −*𝑓* over *𝑋*. Besides

some results for specific functions, the best known of which is probably McCormick convex envelope for bilinear terms over rectangles, some

mostly related to so called polyhedral convex envelopes. A function *𝑓* general results appeared in the literature. The first general results were admits a polyhedral convex envelope over a region *𝑋* if its convex en-

ular, if *𝑋* is a polytope, a function *𝑓* is said to admit a *vertex* polyhedral velope is the maximum of a finite number of aﬃne functions. In partic-

lope of the same function over the vertex set of *𝑋*. In fact, according to convex envelope if its convex envelope is equal to the convex enve-

convex envelopes for continuously differentiable functions *𝑓* . The condi- [Rikun (1997)](#_bookmark173) polyhedral convex envelopes are always vertex polyhedral

of *𝑋*) guarantees the existence of a vertex polyhedral convex envelope tion of edge-concavity (concavity along each segment parallel to an edge

([Meyer and Floudas, 2005a; Tardella, 2003; 2008](#_bookmark141)). For instance, multi- linear functions always admit vertex polyhedral convex envelopes over

are strictly related to triangulations of the polytope *𝑋* with a number boxes. In [Meyer and Floudas (2005a)](#_bookmark141) it is shown that these envelopes of simplices which can be very large (e.g., of size *𝑛*! for *𝑛*-dimensional

except for some special cases. In particular, when *𝑋* is the unit box and boxes). Identification of the triangulation may be a rather diﬃcult task,

*𝑓* is submodular over the vertex set of *𝑋*, the convex envelope is the Lovász extension of *𝑓* (see [Tawarmalani et al., 2013](#_bookmark149)). Recently, differ-

Most of them require that *𝑋* is a box, while *𝑓* is required to satisfy differ- ent results about non-polyhedral convex envelopes have been presented. ent assumptions. In [Tawarmalani and Sahinidis (2001)](#_bookmark151) *𝑓* is assumed to

ing *𝑛* − 1 variables are fixed. In [Jach et al. (2008)](#_bookmark124) the Hessian of *𝑓* is be convex if we fix the value of one variable and concave if the remain- assumed to have at least one negative eigenvalue over the box and *𝑓* is

[(2012, 2013) *𝑓* is a product function *ℎ*(**𝐳**)*𝑔*(**𝐲**), where *ℎ* and *𝑔* must fulfill convex if the value of one variable is fixed. In Khajavirad and Sahinidis](#_bookmark90) some conditions (e.g., *ℎ* must be nonnegative and convex and with some specific form, *𝑔* must be nonnegative and component-wise concave). In [Ballerstein and Michaels (2014)](#_bookmark27) and [Locatelli (2016b)](#_bookmark111) *𝑓* is required to be (strictly) convex if we fix the values of *𝑛* − 1 variables, while if we

is attained at a vertex of the corresponding (*𝑛* − 1)-dimensional box. In fix the value of the remaining variable, the minimum of the function all the above works, the required properties for *𝑓* imply that the con- vex envelope of *𝑓* over the box is equivalent to the convex envelope

of the same function over the border of the box. In [Locatelli (2020)](#_bookmark114) a case where this does not hold has been discussed, namely the convex envelope of bivariate cubic functions over rectangles.

The convex (and concave) envelopes of some functions of relatively simple form (usually, univariate or bivariate functions) can also be em- ployed to underestimate and/or overestimate functions with a more complicated form. This is the case for factorable functions, i.e., functions which can be progressively decomposed into the sum and product of simple univariate and bivariate functions, for which convex and concave [envelopes are available (see, e.g., Khajavirad et al., 2014; Scott et al., 2011; Tawarmalani and Sahinidis, 2004). For instance, given the fac-](#_bookmark142)

torable function *𝑒𝑥*1 *𝑥*2 *𝑥*3 , we first introduce the new variables *𝑥*4 = *𝑥*1 *𝑥*2 ,

*𝑥*5 = *𝑒𝑥*4 , and *𝑥*6 = *𝑥*5 *𝑥*3 , and then these equality constraints are replaced

possible convex underestimator (concave overestimator) for this func- tion over the unit hypercube can be obtained by summing McCormick’s convex envelope (McCormick’s concave envelope) of each single bi- linear term over the unit square. Unfortunately, the convex underes- timator is *not* the convex envelope of the overall sum (in general, the convex envelope of a sum is not the sum of the convex envelopes of the single terms of the sum). Similarly, for the concave overestimator. However, there are some results which bound the difference between the concave overestimator and the convex underestimator. In particu- lar, in [Luedtke et al. (2012)](#_bookmark121) it is shown that such difference cannot be larger than the difference between the concave and the convex enve-

[lope over the unit hypercube times a *𝑂*(*𝑛*) constant, while in Boland](#_bookmark56)

[et √al. (2017) the order of magnitude of the constant has been refined to](#_bookmark56)

*𝑂*( *𝑛*).

For a poorly structured function *𝑓* the best known approach for un- derestimating *𝑓* is the one proposed within the framework of the *𝛼*-BB approach (see [Adjiman et al. (1998a, 1996, 1998b)](#_bookmark16)). In this case, *𝑓* is

function *𝑞* is introduced such that the Hessian of *𝑓* − *𝑞* is semidefinite only assumed to be twice-continuously differentiable and a nonnegative positive over a box *𝑋* or, stated in another way, *𝑓* − *𝑞* is a convex un- derestimator of *𝑓* over *𝑋*. In the original *𝛼*-BB approach *𝑞* was defined

as

∑*𝑛*

*𝑞*(*𝑥*) = *𝛼𝑖* (*𝑥𝑖* − 𝓁*𝑖* )(*𝑢𝑖* − *𝑥𝑖* )*,*

*𝑖*=1

where 𝓁*𝑖 , 𝑢𝑖* are, respectively, the lower and upper bound for variable

*𝑥𝑖* within the box *𝑋*, while the values *𝛼𝑖* are suitably chosen in or- der to guarantee convexity of *𝑓* − *𝑞* over *𝑋*. Later on, different func- tions *𝑞* have been proposed in the literature, like, e.g., spline functions [Meyer and Floudas (2005b)](#_bookmark143), obtained by first subdividing the box *𝑋* into smaller sub-boxes and then computing different (and sharper) *𝛼* val-

[ues over these sub-boxes, or exponential functions (Akrotirianakis and Floudas, 2004). In the recent work](#_bookmark17) [Kazazakis](#_bookmark138) [and Adjiman (2018) a fur-](#_bookmark17) ther variant has been proposed, where the original function is replaced

by a so called *𝜇*-subenergy function, with the property that its eigen-

values have a smaller magnitude with respect to those of the original

to choose tighter *𝛼* values in those regions. function in the regions far away from global minimizers, thus allowing

For poorly structured GO problems we should also mention meth- ods based on interval arithmetic. These are rigorous methods, i.e., they guarantee the detection of a solution within a prescribed tolerance even in the presence of rounding errors. The overall number of contributions in this field in the last two decades is relatively limited, but there are [some interesting works. We first mention the two surveys (Araya and Reyes, 2016; Neumaier, 2004), to which we refer for a more detailed](#_bookmark22) discussion about the topic. In some works first and second order infor- mation are employed both to define linear and quadratic underestimat- ing functions and to fathom (or shrink) boxes which are guaranteed not to include points satisfying first and second order necessary optimality [conditions (see, e.g., Hansen et al., 2007; Markot and Schichl, 2014;](#_bookmark105)

[Martinez et al., 2004). Note that the original *𝛼*-BB approach exploits](#_bookmark105)

second order information to derive a convex quadratic underestimating

function. In [Borradaile and Van Hentenryck (2005)](#_bookmark73), [Kearfott (2011)](#_bookmark140) and [Neumaier and Shcherbina (2004)](#_bookmark150) safe linear underestimators are dis- cussed. These are linear relaxations which take into account numeri- cal errors in the computation of the coeﬃcients of the linear functions. The paper [Berenguel et al. (2013)](#_bookmark44) presents ways to take into account separability of the objective function. In [Carrizosa et al. (2004)](#_bookmark34) it is shown that simple translations of the variables allow to derive bet- ter inclusion functions through interval arithmetic. A discussion about different branching rules for the subdivision of a box into sub-boxes can be found in [Csallner et al. (2000)](#_bookmark41) and [Markót et al. (2006)](#_bookmark128). In

[Jansson et al. (2007)](#_bookmark125) interval arithmetic is applied to compute rigor- ous error bounds for the optimal value of semidefinite programs. In [Domes and Neumaier (2016)](#_bookmark70) it is shown that the information extracted from local optimization, namely the approximated local optimizer as well as the corresponding Lagrange multipliers, can be exploited to form an aggregated redundant constraint which turns out to be power- ful in order to tight the bounds on the variables (see also the following [Section 4.2.5](#_bookmark12)). The authors also remark that such aggregate constraints are able to reduce the so called *cluster eﬀect*, i.e., the presence of a large cluster of unfathomed small boxes in the regions around the global min- imizer, which is a serious obstacle to the eﬃciency of B&C approaches. We also mention the successful application of an interval method to the circle packing problem (see [Markót and Csendes, 2005](#_bookmark127)).

We conclude this section by observing that up to now we have talked

about *convex* underestimators. However, as discussed in [Section 4.1](#_bookmark6), there are some nonconvex problems which can be solved eﬃciently. So, in some cases nonconvex underestimators can be employed. For in- stance, in [Cartis et al. (2015)](#_bookmark35) and [Fowkes et al. (2013)](#_bookmark83) the objective function is underestimated by quadratic and cubic functions, whose min- imizers can be eﬃciently computed over spheres, which are the regions generated by the branching operation adopted in those works. If the branching operation generates polyhedral sets with a limited number of vertices, then concave underestimators can be employed. Indeed, the minimum value of a concave function over a polytope is attained at a vertex of the polytope. This is done, e.g., in the context of Lipschitz op- [timization (see, for instance, Hendrix and Tóth, 2010; Paulavic˘ius and Zi˘linskas, 2014), where a concave lower bounding function based on](#_bookmark107) the Lipschitz condition is minimized over a simple polyhedral region (in particular, a simplex).

* + 1. *Problem reformulations and convex relaxations*

For some nonconvex problems a convex relaxation can be derived by first reformulating the problem, usually with the addition of new vari- ables, and then removing or relaxing the nonconvex constraints of the

QCQPs, where the additional matrix variable *𝑋* and the equality con- reformulation. A typical example is the well known Shor relaxation for

straint *𝑋* = *𝑥𝑥⊤* are introduced. This way, each quadratic form *𝑥⊤𝑄𝑥*

successors. The predecessors are copositive reformulations for some sub- classes of QP problems. In particular, the first of such results is a copos- itive reformulation of StQP problems presented in [Bomze et al. (2000)](#_bookmark66). In the successors of [Burer (2009)](#_bookmark77), completely positive and copositive reformulations have been extended to other classes of optimization problems. In fact, in [Burer (2009)](#_bookmark77) itself the result is extended also to [cases where some quadratic equality constraints appear. In Burer and Dong (2012) it is shown that QCQPs can be reformulated as *general-*](#_bookmark80) *ized* copositive programs, i.e., problems defined over the generalized

completely positive cone, where the requirement *𝑥𝑘* ∈ ℝ*𝑛 ,* ∀*𝑘* ∈ *𝐾*, is

+

replaced by *𝑥𝑘* ∈ G*,* ∀*𝑘* ∈ *𝐾*, where G is a convex cone. The paper

[Amaral et al. (2014)](#_bookmark18) reformulates the problems of minimizing the ratio

of two quadratic functions over a polyhedral region as problems over the cones of completely positive matrices and of copositive matrices. In

problems whose variables are constrained to belong to G ∩ n, where [Bai et al. (2016)](#_bookmark28) it is shown that, under suitable assumptions, QCQP

G is a convex cone and n is a region defined by linear equality con- straints, can be reformulated as problems over the cone of completely positive matrices. In [Bomze et al. (2017)](#_bookmark60) it is shown that, besides the one presented in [Burer (2009)](#_bookmark77), other equivalent completely positive re- formulations for mixed-binary QPs are possible. The interest of these equivalent reformulations lies in the fact that, once the completely pos- itive cone is relaxed into a tractable one, the new reformulations may lead to tighter bounds. In [Bomze et al. (2018)](#_bookmark62) it is shown that for QP problems with quadratic objective function, two quadratic constraints and some further linear constraints, under suitable assumptions a copos- itive reformulation exists. For some problems, a completely positive re- formulation is not available but still they admit completely positive re- laxations. In [Bomze (2015)](#_bookmark58) it is proved that for QCQPs bounds returned by completely positive relaxations dominate Lagrangian dual bounds. This result has been extended to polynomial programming problems in [Kuang and Zuluaga (2018)](#_bookmark92) after the introduction of completely positive tensors.

Unfortunately, though convex, the cones of copositive and com-

pletely positive matrices are not tractable, i.e., problems over these cones cannot be solved in polynomial time by interior point methods because the computation of their self-concordant barrier functions can-

can be replaced by the linear term *𝑄* ∙ *𝑋* = ∑*𝑛*

*𝑖,𝑗*=1

*𝑄𝑖𝑗*

*𝑋𝑖𝑗*

. Nonconvex-

not be performed in polynomial time. All the same, reformulations over the completely positive or copositive cone allow to define polynomi-

ity only lies in the rank-one constraint *𝑋* = *𝑥𝑥⊤*. A convex relaxation is

straint *𝑋* ⪰ *𝑥𝑥𝑇* , i.e., it is required that the matrix *𝑋* − *𝑥𝑥⊤* is positive obtained by replacing the equality constraint with the semidefinite con-

semidefinite. The resulting SDP bound has a good quality but is costly. Conversely, SOCP relaxations, obtained by replacing quadratic functions with convex quadratic underestimating functions, have a poorer qual- ity but are faster to compute. For this reason in [Burer et al. (2014)](#_bookmark83) it is proposed to construct mixed SOCP-SDP relaxations which allow for a balance between the quality of the bound and the time needed to com- pute it.

The work [Burer (2009)](#_bookmark77) introduces an exact reformulation for QP problems with quadratic objective function and linear constraints, pos-

vex cone of *𝑛* × *𝑛* completely positive matrices sibly with some binary variables. The reformulation is based on the con-

C∗ = {*𝑋* ∈ ℝ*𝑛*×*𝑛* ∶ *𝑋* = ∑ *𝑥 𝑥⊤,* |*𝐾*| has finite cardinality*, 𝑥* ∈ ℝ*𝑛 ,*

ally solvable convex relaxations by replacing these cones with tractable

matrices are tractable and they outer approximate C∗ and, being self- cones. The cone of nonnegative matrices and the cone of semidefinite dual, they inner approximate C*𝑛* . The intersection of these two cones

*𝑛*

is the cone of doubly nonnegative matrices, which inner approximates

C*𝑛* (in fact, equality holds for *𝑛 <* 5), while its dual cone, i.e., the cone

ative and a semidefinite matrix, outer approximates C∗. Different hi- made up by matrices which can be obtained by summing a nonneg- erarchies of cones {G*𝑟* } have been proposed in the literature such that

*𝑛*

*𝑛*

∀*𝑟* ∶ G*𝑟 ⊃* G*𝑟*+1 *⊃* C∗, i.e., the cones are finer and finer outer approxima-

*𝑛 𝑛 𝑛*

tions of the completely positive cone and, conversely, their dual cones [are finer and finer inner approximations of C*𝑛* (see de Klerk and Pasech- nik, 2002; Parrillo, 2000; Peña et al., 2007). Although the bounds based](#_bookmark57)

on these hierarchies tend to become sharper as *𝑟* increases, the main

drawback is that the size of these problems tends to increase rapidly

*𝑛*

∀*𝑘* ∈ *𝐾*}*,*

*𝑘*∈*𝐾*

*𝑘 𝑘*

*𝑘* +

with *𝑟*, since they involve *𝑛𝑟*+1 × *𝑛𝑟*+1 matrices or a comparable number of *𝑛* × *𝑛* matrices. Thus, from the computational point of view only the use of small values for *𝑟* is feasible. In [Bundfuss and Duer (2009)](#_bookmark75), after

and its dual cone, the dual cone of copositive matrices:

observing that a matrix *𝐴* is copositive if and only if *𝑥⊤𝐴𝑥* ≥ 0 for all

C*𝑛*

= {*𝑋* ∈ ℝ*𝑛*×*𝑛* ∶ *𝑥⊤𝑋𝑥* ≥ 0 ∀*𝑥* ∈ ℝ*𝑛* }*.*

[*𝑥* ∈ Δ*𝑛*, where Δ*𝑛* is the *𝑛*-dimensional unit simplex (see also Tóth et al.,](#_bookmark160)

[2021), polyhedral inner and outer approximations of the copositive](#_bookmark160)

In [Burer (2009)](#_bookmark77) it is proved that, under mild assumptions, any QP prob- lem with linear constraints and, possibly, some binary variables, can be reformulated as a problem over the convex cone of completely positive matrices (we refer, e.g., to the survey paper [Bomze et al. (2012)](#_bookmark64) for a discussion about problems over this cone and over the dual cone of copositive matrices). This result has some predecessors and different

+

cone are proposed, based on simplicial subdivisions of the unit sim- plex. It is shown that, under suitable assumptions, bounds computed by replacing the copositive cone with these approximations tend to con- verge to the optimal value of the copositive problem, provided that the length of the largest edge of the simplices in the partition converges to 0.

In order to strengthen the bound obtained by replacing C∗ with a tractable cone, we may also proceed as follows. Let *𝑋⋆* be the opti-

*𝑛*

mal solution obtained by solving the relaxation over the tractable cone

G *⊃* C∗. Then, either *𝑋⋆* ∈ C∗ or *𝑋⋆* ∈ G ⧵ C∗. In the former case *𝑋⋆* is

*𝑥𝑖* of the problem. Again for Box QP, in [Burer and Letchford (2009)](#_bookmark84) first the objective function is linearized with the addition of the variables *𝑋𝑖𝑗* and the related constraints *𝑋𝑖𝑗* = *𝑥𝑖 𝑥𝑗* , and then the convex hull of the

set

*𝑛 𝑛 𝑛*

also an optimal solution of the completely positive problem. Otherwise, {

*𝑛*+*𝑛*(*𝑛*+1)∕2 }

by definition of dual cone, there exists some *𝐶* ∈ C*𝑛* such that *𝐶* ∙ *𝑋⋆ <* 0.

(*𝑥, 𝑋*) ∈ [0*,* 1]

∶ *𝑋𝑖𝑗* = *𝑥𝑖 𝑥𝑗 ,* 1 ≤ *𝑖* ≤ *𝑗* ≤ *𝑛 ,*

Thus, adding the inequality *𝐶* ∙ *𝑋* ≥ 0 to the previous relaxation, we are

able to strengthen the previously computed bound. In other words, a

separation problem is solved. Such an approach has been explored in

is analyzed. Many facet-inducing inequalities for this set are de- rived from valid inequalities for the Boolean quadric polytope (see [Padberg (1989)](#_bookmark161)):

[Burer and Dong (2013)](#_bookmark81) and also in [Bomze et al. (2010)](#_bookmark59) for what con- cerns the completely positive reformulation of the max-clique problem.

{

(*𝑥, 𝑋*) ∈ {0*,* 1}

*𝑛*+*𝑛*(*𝑛*−1)∕2

}

*𝑖𝑗 𝑖 𝑗* ≤ *𝑖 < 𝑗* ≤ *𝑛 .*

∶ *𝑋* = *𝑥 𝑥 ,* 1

Note that this topic would also fit into the following [Section 4.2.4](#_bookmark10) about cutting planes.

Reformulations and relaxations have also been proposed for poly-

nomial programming problems. The unconstrained minimization of a

For QP problems or, more generally, for polynomial programming prob- lems, a convex relaxation can be strengthened by the Reformulation- [Linearization Technique (RLT), introduced in Sherali and Tunc-](#_bookmark181)

[bilek (1992). If the linear constraints *𝑎⊤𝑥* ≥ *𝑏*1 and *𝑎⊤𝑥* ≥ *𝑏*2 are present,](#_bookmark181)

[1 2](#_bookmark181)

polynomial *𝑓* with degree *𝑑* can easily be reformulated as:

then first the additional quadratic constraint (*𝑎⊤𝑥* − *𝑏*1 )(*𝑎⊤𝑥* − *𝑏*2 ) ≥ 0,

1 2

sup *𝛼*

*𝑓* (*𝑥*) − *𝛼* ∈ g*𝑛,𝑑 ,*

is linearized by replacing the terms *𝑥𝑖 𝑥𝑗* with the additional variables implied by the two linear constraints, is added, and then this constraint

*𝑋𝑖𝑗* . A typical example are the RLT constraints obtained by multiplying

where g

*𝑛,𝑑*

is the set of nonnegative polynomials of degree at most

the box constraints 0 ≤ *𝑥𝑖 , 𝑥𝑗* ≤ 1:

*𝑑*. However, the set of nonnegative polynomials is not easily repre-

sentable. Therefore, a tractable relaxation can be obtained by replac-

most *𝑑*, i.e., polynomials which can be written as a sum of a finite num- ing it with the set of Sum-Of-Squares (SOS) polynomials of degree at ber of square of polynomials of degree at most *𝑑* . The nice feature of

2

SOS polynomials is that they are representable by a positive semidefi-

.

*𝑋𝑖𝑗* ≥ 0*, 𝑋𝑖𝑗* ≥ *𝑥𝑖* + *𝑥𝑗* − 1*, 𝑋𝑖𝑗* ≤ *𝑥𝑖 , 𝑋𝑖𝑗* ≤ *𝑥𝑗 ,*

which also correspond to the classical McCormick envelopes. In

including both the semidefinite condition on the variable matrix *𝑋* [Anstreicher (2009)](#_bookmark21) it is shown that for Box QPs and some QCQPs,

and RLT constraints, leads to significatively better bounds than us-

nite condition imposed over matrices with dimension

*𝑂*(*𝑛𝑑*∕2) In fact,

ing the semidefinite condition or the RLT constraints alone. The paper

[Anstreicher and Burer (2010)](#_bookmark23) derives convex hulls for quadratic forms

[Papp and Yildiz (2019)](#_bookmark162) discusses an alternative and cheaper way to rep- resent SOS polynomials. The bound based on the SOS relaxation (and its dual counterpart, the moment relaxation, see [Lasserre, 2001](#_bookmark97)) has been strengthened with the definition of hierarchies of bounds both for the unconstrained and the (polynomially) constrained case, which have been discussed in different papers such as, e.g., [de Klerk et al. (2017a,b)](#_bookmark53), [Lasserre (2001, 2005, 2006)](#_bookmark97), [Laurent (2007)](#_bookmark100), [Nie (2013)](#_bookmark156), [Nie (2014)](#_bookmark157),

[Nie et al. (2006)](#_bookmark159) and [Vui and So’n (2008)](#_bookmark168).

* + 1. *Outer approximation and cutting planes*

When the feasible region *𝑆* is not convex, we need to outer approx-

The tightest convex outer approximation of *𝑆* is called *convex hull* of imate it with a convex region in order to derive a convex relaxation.

*𝑆* and is denoted by *𝑐ℎ𝑢𝑙𝑙*(*𝑆*). One possibility to derive a convex outer approximation is to replace each constraint *𝑔*(*𝑥*) ≤ 0 defining *𝑆*, *𝑔* non- convex, with a constraint *𝑐*(*𝑥*) ≤ 0, where *𝑐* is a convex underestimator of *𝑔* over *𝑆*, so that all the material discussed in [Section 4.2.2](#_bookmark9) could

be applied here. The tighter the convex underestimator, the tighter is the convex outer approximation. However, replacing each nonconvex

function *𝑔* with its convex envelope over a suitable region containing

*𝑆* leads to a convex outer approximation, but this is not necessarily

the convex hull. A simple example is the following. Given the region

*𝑆* = {(*𝑥, 𝑦*) ∈ [1*,* 2] ∶ *𝑥𝑦* ≥ 3}, according to the previous discussion the

over small-dimensional regions (triangles and boxes) based on semidefi- nite and nonnegative conditions over matrices and, possibly, additional

defined for the characterization of *𝑐ℎ𝑢𝑙𝑙*(*𝑆*) when RLT constraints. In [Bienstock and Michalka (2014a)](#_bookmark50) cutting planes are

*𝑆* = {(*𝑥, 𝑞*) ∈ ℝ*𝑑* × *𝑅* ∶ *𝑞* ≥ *𝑄*(*𝑥*)*, 𝑥* ∈ ℝ*𝑛* ⧵ *𝑖𝑛𝑡*(*𝑃* )}*,*

where *𝑄* is convex and differentiable and *𝑖𝑛𝑡*(*𝑃* ) denotes the interior of set *𝑃* . A polynomial separation algorithm is proposed for the case when *𝑄* is quadratic and strictly convex, while *𝑃* is a polyhedron or

an ellipsoid. The work [Burer and Kilinç-Karzan (2017)](#_bookmark82) derives convex relaxations and, under suitable assumptions, even convex hulls for the intersection of the following sets:

* a SOC representable cone G, i.e., given a matrix *𝐵* ∈ ℝ*𝑛*×(*𝑛*−1) and

*𝑏* ∈ ℝ*𝑛*, G = {*𝑥* ∶ ‖*𝐵⊤𝑥*‖ ≤ *𝑏⊤𝑥*};

* a cone g defined by a homogeneous quadratic function, i.e., g =

{*𝑥* ∶ *𝑥⊤𝐴𝑥* ≤ 0}, for some matrix *𝐴* ∈ ℝ*𝑛*×*𝑛*;

* an aﬃne hyperplane.

Based on the observation that *𝐴* ⪰ *𝑂* and *𝐵* ⪰ *𝑂* implies *𝐴 ⊗ 𝐵* ⪰ *𝑂*, where *⊗* denotes the Kronecker product, in [Anstreicher (2017)](#_bookmark19) so called Kronecker product constraints are introduced. Let *𝐻* (*𝑥*) ⪰ *𝑂* and *𝐺*(*𝑥*) ⪰

*𝑂* be semidefinite constraints where single components of both matrices

are aﬃne functions of *𝑥*. Then, a Kronecker product constraint is ob-

nonconvex function *𝑥𝑦* should be replaced by its concave envelope over

[1*,* 2]2 , i.e., by min{*𝑥* + 2*𝑦* − 2*,* 2*𝑥* + *𝑦* − 2}. But the resulting convex re-

tained by replacing in *𝐻* (*𝑥*) *⊗ 𝐺*(*𝑥*) ⪰ *𝑂* each term *𝑥𝑖 𝑥𝑗*

with *𝑋*

*𝑖𝑗*

. These

gion {(*𝑥, 𝑦*) ∈ [1*,* 2] ∶ *𝑥* + 2*𝑦* − 2 ≥ 3*,* 2*𝑥* + *𝑦* − 2 ≥ 3} is not the convex hull of *𝑆*, since *𝑆* is already a convex set (just note that *𝑥𝑦* ≥ 3 can be rewritten as *𝑥* ≥ 3∕*𝑦*). More generally, in [Anstreicher (2012)](#_bookmark20) for QCQPs, with additional linear constraints defining a polytope *𝑃* , it is shown that replacing the quadratic terms with their convex envelopes over *𝑃* leads

the convex hull of the set {(1 *𝑥*)(1 *𝑥*)*⊤, 𝑥* ∈ *𝑃* } (note, however, that both to a convex relaxation which is dominated by the one obtained through

the convex envelope and the convex hull may be hard to compute).

Many results about convex hulls have been presented in the recent literature. The work [Vandenbussche and Nemhauser (2005b)](#_bookmark165), within the framework of a branch-and-bound approach for Box QP problems

constraints are generalization both of the classical RLT constraints ob- tained from two linear inequality constraints, and of the SOC-RLT con- straints obtained from one linear inequality constraint and a SOC con- straint. In [Wang and Kilinç-Karzan (2021)](#_bookmark172) for some QCQPs, conditions are given under which the convex hull of the epigraph of the QCQP is [the projection of the epigraph of its Shor relaxation. In Del Pia and Kha- javirad (2017, 2018) in the context of binary polynomial problems, i.e.,](#_bookmark58) problems with a polynomial objective function, some binary variables and some other variables constrained to belong to the interval [0,1], valid inequalities are derived to characterize the convex hull of multi- linear sets:

based on KKT branching (see [Section 4.2.1](#_bookmark8)), considers convex hulls of the regions defined by the KKT conditions related to each single variable

*𝑆* =

{

(*𝑥, 𝑦*) ∶ *𝑦𝐼*

}

∏

= *𝑥𝑖 , 𝐼* ∈ S*, 𝑥* ∈ {0*,* 1}*𝑛 ,*

*𝑖*∈*𝐼*

where S is a collection of subsets of {1*,* … *, 𝑛*} with cardinality not lower

than two.

convex hulls for the intersection of cubes with regions {(*𝑥, 𝑦, 𝑧*) ∶ The paper ([Nguyen et al., 2018)](#_bookmark153) presents the (non polyhedral)

*𝑥𝑏*1 *𝑦𝑏*2 ≥ *𝑧*} and with regions {(*𝑥, 𝑦, 𝑧*) ∶ *𝑥𝑦𝑏*2 ≤ *𝑧*} for *𝑏*1 *, 𝑏*2 ≥ 1. In

[Davarnia et al. (2017)](#_bookmark49) a description is given for the convex hull of sets

*𝑆* = {(*𝑥, 𝑦, 𝑧*) ∶ *𝑥* ∈ [0*,* 1]*𝑛, 𝐸𝑥* ≥ *𝑓, 𝑦* ∈ ℝ*𝑚, 𝑦* ≥ 0*, 𝑒⊤𝑦* ≤ 1*,*

*𝑚*

*𝑧𝑘* = *𝑦⊤𝐴𝑘 𝑥, 𝑘* = 1*,* … *, 𝐾*}*,*

where *𝐸* ∈ ℝ*𝑡*×*𝑛*, *𝑓* ∈ ℝ*𝑡*, *𝑒𝑚* is the *𝑚*-dimensional vector with all compo- nents equal to one, *𝐴𝑘* ∈ ℝ*𝑚*×*𝑛*. Then, *𝑥* belongs to a polyhedral subset of the unit box, *𝑦* belongs to the *𝑚*-dimensional unit simplex, while *𝑧* is obtained by bilinear terms involving *𝑥* and *𝑦*. An application to net-

work interdiction problems is presented. In [Bonami et al. (2019)](#_bookmark71) some cutting planes to strengthen linear relaxations of QPs with linear con- straints are proposed. The cutting planes are based on a well known result by Motzkin and Straus (see [Motzkin and Strauss (1965)](#_bookmark147)) about max clique problems. For this reason, they are called Motzkin-Straus clique inequalities. The work [Santana and Dey (2020)](#_bookmark177) establishes that the convex hull of the intersection of the region defined by a quadratic equality constraint and a polytope is SOCP representable. In

[Bienstock et al. (2020)](#_bookmark48) valid inequalities are introduced for sets *𝑆* ∩ *𝑃* ,

where *𝑆* is a closed set and *𝑃* is a polyhedron. A cutting plane algorithm

is proposed. Given an oracle returning the distance of some point from

*𝑆*, the algorithm generates cutting planes which are able to approximate

*𝑐ℎ𝑢𝑙𝑙*(*𝑆* ∩ *𝑃* ) in an arbitrarily precise way. An application to polynomial

vex hulls are given for subsets of ℝ5, arising in pooling problems and programming is presented. In [Luedtke et al. (2020)](#_bookmark119) non polyhedral con-

defined by some linear constraints and a nonconvex bilinear constraint.

* + 1. *Bound tightening*

Rather special cutting planes are those involving lower and upper bounds on single variables. Strengthening these bounds is also known

as *bound tightening*. Given a lower and upper bound 𝓁 and *𝑢* for vari-

ables *𝑥*, a bound tightening procedure can be viewed as a function re- ceiving the box [𝓁*, 𝑢*] in input and returning a box [𝓁′*, 𝑢*′] *⊆* [𝓁*, 𝑢*], such that [𝓁*, 𝑢*] ⧵ [𝓁′*, 𝑢*′] does not contain any feasible point (feasibility-based

bound tightening) or, more powerfully, does not contain any optimal so- lution (optimality-based bound tightening). The importance of bound- tightening is due to the fact that it does not only reduce the feasible re- gion of a convex relaxation, but it also improves the quality of (convex) underestimating functions for the objective and constraint functions. In- deed, in many cases the underestimating functions depend on the lower and upper limits for the variables and they tend to become tighter as these limits become tighter. The importance of bound-tightening for an eﬃcient solution of nonconvex problems is proved by the fact that many

GO solvers (such as BARON [Sahinidis (2017)](#_bookmark175), BMIBNB [Lofberg (2004)](#_bookmark117),

COUENNE [Belotti et al. (2006)](#_bookmark35), SCIP [Gamrath et al. (2020)](#_bookmark88)) include

procedures to perform it.

A simple way to perform bound tightening is by minimizing and maximizing each variable over a convex relaxation of the feasible re- gion (feasibility-based bound tightening), or over the same relaxation with an additional constraint imposing that a convex underestimating function for the objective function is not larger than the current known upper bound (optimality-based bound tightening, since the additional constraint may remove feasible points but no optimal solution). Since, as previously commented, convex underestimating functions usually de- pend on variable bounds, once all bounds have been tightened, the new bounds improve the quality of the underestimating functions so that a further round of bound-tightening may allow to further reduce the bounds. This can be iteratively repeated until convergence. Such itera- tive procedure has been theoretically and computationally investigated in [Caprara and Locatelli (2010)](#_bookmark32) and [Caprara et al. (2016)](#_bookmark33). Note, how- ever, that such iterative procedure, while providing tight bounds, is quite expensive, requiring the solution of many convex subproblems.

In [Gleixner et al. (2017)](#_bookmark98) three techniques are introduced to keep the computational cost of these bound tightening techniques under control. The work [Tawarmalani and Sahinidis (2004)](#_bookmark154) presents a general theo- retical framework for bound-tightening techniques and discusses ways to exploit dual solutions of convex subproblems in order to perform the tightening. Constraint propagation techniques are also widely employed to tighten bounds. For factorable functions, first a directed acyclic graph (DAG) is built, where the nodes correspond to the variables (both the original variables and the additional variables introduced to model fac- torable functions), while the arcs represent the dependencies between the variables. Then, bound tightening is performed by forward and back- ward propagation along this graph. This technique is discussed in dif- [ferent papers such as](#_bookmark178) [Belotti et al. (2009)](#_bookmark37)[,](#_bookmark178) [Messine (2004)](#_bookmark139)[, Schichl and Neumaier (2005),](#_bookmark178) [Vu](#_bookmark166) [et al. (2009) and](#_bookmark178) [Wechsung](#_bookmark174) [et al. (2015). In](#_bookmark178) [Puranik and Sahinidis (2017)](#_bookmark171) it is observed that, in order to perform optimality-based bound tightening, besides the previously mentioned additional constraint involving a convex underestimating function of the objective, one could also add constraints imposing necessary opti- mality conditions or, more precisely, a convex relaxation of such con- ditions. This is also done in [Zhang et al. (2020)](#_bookmark184) where some topics are addressed such as how to bound the dual variables appearing in the optimality conditions, for which explicit bounds are not given.

# Computational aspects, test problems and solvers

Most of the papers cited in this work report computational experi- ments on different sets of test problems and with different solvers. Dis- cussing in detail computational experiences, test functions and solvers is beyond the scope of the current work. However, we make a few ob- servations and, for the interested reader, we provide pointers to papers and web sites where these aspects are presented in more detail.

* 1. *Computational aspects*

Here we briefly discuss a couple of observations which are well known but should always be taken into account when evaluating and comparing different solution approaches.

The first observation is that we should never search for the ’best’ approach to solve GO problems. This is obviously true for the whole class of GO problems, since such wide class encompasses problems with fairly different properties and characteristics, which lead to fairly differ- ent approaches for their solution. But it is also true for much narrower subclasses of GO problems. In particular, we mention here a subclass of problems which attracted a lot of attention in the recent literature, namely the class of nonconvex QP problems with linear constraints. Well

known commercial software products like CPLEX and GUROBI have re-

cently introduced solvers for the solution of problems within this class. These QP problems can be tackled in many different ways. As already mentioned at the end of [Section 4.2.1](#_bookmark8), in [Xia et al. (2020)](#_bookmark176) the problem is reformulated as a Mixed Integer Linear Program (MILP) after refor- mulating it with the inclusion of the KKT conditions. Binary variables are included to model the nonconvex complementarity conditions. In [Chen and Burer (2012)](#_bookmark36) the problem is still reformulated with the inclu- sion of KKT conditions and KKT branching is applied, but semidefinite relaxations are considered. In [Bonami et al. (2018)](#_bookmark69) a spatial branch-and- cut approach is proposed with the addition of valid cuts for the Boolean Quadric Polytope. In [Liuzzi et al. (2021)](#_bookmark102) another spatial branch-and- bound approach with intensive bound-tightening has been applied to a class of QP problems arising from an application in game theory.

The outcome of the computational experiments reported in all these works is that none of the proposed methods strictly dominates the oth- ers. The method proposed in [Xia et al. (2020)](#_bookmark176), as well as problem spe- cific ones proposed in [Gondzio and Yildrim (2021)](#_bookmark119), [Liuzzi et al. (2019)](#_bookmark103), are the best performing over the subclass of StQP problems. The ap- proach proposed in [Bonami et al. (2018)](#_bookmark69) performs quite well over the subclass of Box QP problems and techniques proposed in that paper have

been successfully incorporated in CPLEX. However, for Box QP prob- lems with dense Hessian matrices the best approach appears to be the one proposed in [Chen and Burer (2012)](#_bookmark36). Finally, the approach proposed in [Liuzzi et al. (2021)](#_bookmark102) is the best performing over the subclass of QP

problems presented in that work (but ongoing experiments prove its ef- fectiveness over more general QP problems with linear constraints).

The second observation we should always keep in mind is that when- ever we deal with a class of special structured GO problems, incorporat- ing as much as possible the special structure into a solution approach allows for significant improvements. That holds true both for exact and for heuristic methods. We just mention a couple of examples taken from our own personal experiences, but many other examples could be given. Circle packing problems are special structured QCQPs with many nonconvex quadratic constraints, namely the nonoverlapping con- [straints. The best exact methods for this problem (see, e.g., Markót and Csendes (2005) for circle packing into a unit square) are those which](#_bookmark127) incorporate tools based on special properties of the circle packing prob- lem, such as symmetry-breaking tools. Special purpose methods strongly outperform general purpose methods for QCQPs when applied to circle

packing problems.

Molecular conformation problems lead to challenging GO problems, with a number of local minimizers which grows exponentially with the number of atoms. Exact methods can be applied only at very small dimensions (i.e., with a small number of atoms). But many excellent heuristic approaches exist. Also in this case the best performing ap- proaches strongly rely on special properties (in particular, geometrical [properties) of the molecular conformation problems (see, e.g., Ferreiro- Ferreiro et al., 2019; Wales, 2018; Zhao et al., 2017).](#_bookmark78)

* 1. *Test problems*

Test problems for GO have been proposed in different papers, like, e.g., [Furini et al. (2019)](#_bookmark85) for QP problems. We also recall the book [Floudas et al. (1999)](#_bookmark81). Currently, there are different web sites which pro- vide large sets of GO test problems. In [Neumaier (2021)](#_bookmark148) many academic as well as real-life GO test problems are reported. Large collections of test functions have been reported in [Gavana (2021)](#_bookmark93); this quite recent web site reports results of some GO algorithms over the presented test functions and, based on these results, a classification of the diﬃculty of the test functions is proposed. Through different editions of the GECCO Workshop on Real-Parameter Black-Box Optimization Benchmarking (BBOB), a collection of test problems has been collected. Details can be found in [Auger et al. (2019)](#_bookmark29). Another conference, the IEEE Congress on Evolutionary Computation (CEC), organised different competitions on Large Scale Global Optimization, providing many test functions (see, e.g., [Škvorc et al. (2019)](#_bookmark185) and the web site [IEEE TfLsgo (2021)](#_bookmark122)).

* 1. *Solvers and their comparison*

In previous sections we mentioned some GO solvers such as BARON, BMIBNB, COUENNE, SCIP and for nonconvex QPs with linear con- straints also CPLEX and GUROBI. But others are available (see, e.g.,

[Neumaier](#_bookmark55) [(2021)](#_bookmark148) [for quite an extensive list). The work Biscani and Izzo (2020) describes pagmo and pygmo, C++/Python libraries for](#_bookmark55) massively parallel global, possibly multi-objective, optimization. The web site [Johnson (2021)](#_bookmark132) makes available NLopt, another quite large

set of software tools for nonlinear and global optimization. Even SciPy

[Jones et al. (2001–](#_bookmark134)) includes a set of implemented general purpose GO algorithms. Most of the sites containing test problems also present detailed numerical comparisons among different solvers. Many papers compare a newly proposed approach with a limited set of existing GO approaches. But only few papers make a systematic comparison between a large set of different solvers. Here we mention [Neumaier et al. (2005)](#_bookmark152), [where different GO exact methods are compared, and Rios and Sahini- dis (2013), which presents many different derivative-free algorithms for](#_bookmark174) bound constrained problems. We also mention [Beiranvand et al. (2017)](#_bookmark34),

which discusses some guidelines to perform a fair comparison be- tween different solvers. Finally, CEC competitions offer the opportunity of comparing with many different algorithms, while within the con- text of GECCO-BBOB workshops, the Comparing Continuous Optimizers (COCO) platform allows for automated benchmarking. Users can bench- mark their own solvers over a wide set of test functions and compare their results with those of other solvers.

# Conclusions

In this paper we presented our view on what we consider as relevant in the recent GO literature. It can be immediately seen, by simply brows- ing the rich list of references below, how the field is attracting more and more active research and novel computational paradigms. Large scale GO problems are no more out of the scope of solution algorithms, to the

point that even professional solvers like GUROBI and CPLEX now in-

clude some GO solver, at least for nonconvex quadratic optimization. Al- though we tried our best to cover many recent advances, we do not claim to have given account of all relevant papers in the field. This survey re- flects our personal points of view on the subject and we are perfectly aware that omissions are inevitable. Besides alternative approaches to the GO problems presented in this paper, we deliberately did not even mention other relevant fields, like multi-objective GO, stochastic GO, bi-level optimization, parallel, distributed, GPU based or quantum com- puting. These topics might become the subject of a different survey.

# Acknowledgements

We would like to thank the co-editors of this issue, Miguel Anjos, Tibor Illés, and Tamás Terlaky, for their kind invitation to contribute to this important collection of papers. We also thank the anonymous reviewers for their careful reading and constructive suggestions.

# References

[Adjiman,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0001) [C.S., Androulakis, I.P., Floudas, C.A., 1998. A global optimization method, *𝛼*BB, for general twice-differentiable constrained NLPs - II. Implementation and computa- tional results. Comput. Chem. Eng. 22, 1159–1179.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0001)

[method, *𝛼*BB, for process design. Comput. Chem. Eng. Suppl. 20, S419–S424.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0002) [Adjiman,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0002) [C.S., Androulakis, I.P., Maranas, C.D., Floudas, C.A., 1996. A global optimization](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0002)

[method, *𝛼*BB, for general twice differentiable NLPs- I. Theoretical advances. Comput.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0003) [Adjiman,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0003) [C.S., Dallwig, S., Floudas, C.A., Neumaier, A., 1998. A global optimization](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0003)

[Chem. Eng. 22, 1137–1158.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0003)

[Ahmed,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0004) [M.O., Vaswani, S., Schmidt, M., 2020. Combining Bayesian optimization and Lip- schitz optimization. Mach. Learn. 109 (1), 79–102.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0004)

[Ai,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0005) [W., Zhang, S., 2009. Strong duality for the CDT subproblem: a necessary and suﬃcient condition. SIAM J. Optim. 19, 1735–1756.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0005)

[Akrotirianakis, I.G., Floudas, C.A., 2004. A new class of improved convex underestimators for twice continuously differentiable constrained NLPs. J. Global Optim. 30, 367–390.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0006)

[Alexandropoulos, S.-A.N., Pardalos, P.M., Vrahatis, M.N., 2020. Dynamic search trajectory methods for global optimization. Ann. Math. Artif. Intell. 88 (1), 3–37.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0007)

[Amaral,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0008) [P., Bomze, I., Júdice, J., 2014. Copositivity and constrained fractional quadratic problems. Math. Program. 146, 325–350.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0008)

[Anstreicher,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0009) [K., 2012. On convex relaxations for quadratically constrained quadratic pro- gramming. Math. Program. 136, 233–251.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0009)

[Anstreicher,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0010) [K., 2017. Kronecker product constraints with an application to the two-trust-region subproblem. SIAM J. Optim. 27, 368–378.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0010)

[Anstreicher,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0011) [K.M., 2009. Semidefinite programming versus the reformulation-lineariza- tion technique for nonconvex quadratically constrained quadratic programming. J. Global Optim. 43, 471–484.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0011)

[Anstreicher,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0012) [K.M., Burer, S., 2010. Computable representations for convex hulls of low-di- mensional quadratic forms. Math. Program. 124, 33–43.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0012)

[Araya,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0013) [I., Reyes, V., 2016. Interval Branch-and-Bound algorithms for optimization and constraint satisfaction: a survey and prospects. J. Global Optim. 65, 837–866.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0013)

[Audet,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0014) [C., Hansen, P., Jaumard, B., Savard, G., 1999. A symmetrical linear maxmin ap- proach to disjoint bilinear programming. Math. Program. 85, 573–592.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0014)

Auger, A., Brockhoff, D., Hansen, N., Tusa˘ r, T., GECCO workshop on real-parameter black-box optimization benchmarking (BBOB2019). <http://numbbo.github.io/workshops/BBOB-2019/>Accessed: 2021-04-02.

[Bagattini,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0016) [F., Schoen, F., Tigli, L., 2018. Clustering methods for the optimization of atomic cluster structure. J. Chem. Phys. 148 (14).](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0016)

[Bagattini,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0017) [F., Schoen, F., Tigli, L., 2019. Clustering methods for large scale geometrical global optimization. Optim. Methods Softw. 34 (5), 1099–1122.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0017)

[Bai,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0018) [L., Mitchell, J.E., Pang, J.-S., 2016. On conic QPCCs, conic QCQPs and completely positive programs. Math. Program. 159, 109–136.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0018)

[Ballerstein, M., Michaels, D., 2014. Extended formulations for convex envelopes. J. Global Optim. 60, 217–238.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0019)

[Barvinok, A., 1993. Feasibility testing for systems of real quadratic equations. Discr. Com- put. Geom. 10, 1–13.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0020)

[Beale, E.M.L., Forrest, J.J.H., 1976. Global optimization using special ordered sets. Math.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0021)

[Program. 10 (1), 52–69.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0021)

[Beck, A., Teboulle, M., 2009. A convex optimization approach for minimizing the ratio of indefinite quadratic functions over an ellipsoid. Math. Program. 118, 13–35.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0022)

[Becker, R.W., Lago, G.V., 1970. A global optimization algorithm. In: Proceedings of the 8th Allerton Conference on Circuits and Systems Theory, pp. 3–12.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0023)

[Beiranvand, V., Hare, W., Lucet, Y., 2017. Best practices for comparing optimization al- gorithms. Optim. Eng. 18, 815–848.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0024)

Belotti, P., Bonami, P., Vigerske, S., Wächter, A., 2006. Couenne, an exact solver for non- convex MINLPs.

[Belotti, P., Lee, J., Liberti, L., Margot, F., Wächther, A., 2009. Branching and bounds tightening techniques for non-convex MINLP. Optim. Methods Softw. 24, 597–634.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0026)

[Bemporad, A., 2020. Global optimization via inverse distance weighting and radial basis functions. Comput. Optim. Appl..](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0027)

[Ben-Tal, A., den Hertog, D., 2014. Hidden conic quadratic representation of some non- convex quadratic optimization problems. Math. Program. 143, 1–29.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0028)

[Ben-Tal, A., Nemirovski, A., 2001. Lectures on Modern Convex Optimization. Society for Industrial and Applied Mathematics.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0029)

[Berenguel, J., Casado, L., Garcia, I., Hendrix, E., Messine, F., 2013. On interval branch-and-bound for additively separable functions with common variables. J. Global Optim. 56, 1101–1121.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0030)

[Bienstock, D., 2016. A note on polynomial solvability of the CDT problem. SIAM J. Optim.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0031)

[26, 488–498.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0031)

[Bienstock, D., Chen, C., Muñoz, G., 2020. Outer-product-free sets for polynomial optimiza- tion and oracle-based cuts. Math. Program. 183, 105–148.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0032)

[Bienstock, D., Michalka, A., 2014. Cutting-planes for optimization of convex functions over nonconvex sets. SIAM J. Optim. 24, 643–677.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0033)

[Bienstock, D., Michalka, A., 2014. Polynomial solvability of variants of the trust-region subproblem. In: SODA 14 Proceedings of the Twenty-Fifth Annual ACM-SIAM Sym- posium on Discrete Algorithms, pp. 380–390.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0034)

[Binois, M., Ginsbourger, D., Roustant, O., 2020. On the choice of the low-dimensional domain for global optimization via random embeddings. J. Global Optim. 76 (1), 69–90.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0035)

[Biscani, F., Izzo, D., 2020. A parallel global multiobjective framework for optimization: PAGMO. J. Open Source Softw. 5 (53), 2338.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0036)

[Boland, N., Dey, S., Kalinowski, T., Molinaro, M., Rigterink, F., 2017. Bounding the gap between the McCormick relaxation and the convex hull for bilinear functions. Math. Program. 162, 523–535.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0037)

[Bomze, I., 2015. Copositive relaxation beats lagrangian dual bounds in quadratically and linearly constrained quadratic optimization problems. SIAM J. Optim. 25, 1249–1275.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0038) [Bomze, I., Cheng, J., Dickinson, P., Lisser, A., 2017. A fresh CP look at mixed-binary QPs:](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0039)

[new formulations and relaxations. Math. Program. 166, 159–184.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0039)

[Bomze, I., Frommlet, F., Locatelli, M., 2010. Copositive bounds for improving SDP bounds on the clique number. Math. Program. 124, 13–32.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0040)

[Bomze, I., Jeyakumar, V., Li, G., 2018. Extended trust-region problems with one or two balls: exact copositive and Lagrangian relaxations. J. Global Optim. 71, 551–569.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0041)

[Bomze, I., Schachinger, W., Uchida, G., 2012. Think co(mpletely)positive! matrix prop- erties, examples and a clustered bibliography on copositive optimization. J. Global Optim. 52, 423–445.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0042)

[Bomze, I.M., de Klerk, E., 2002. Solving standard quadratic optimization problems via linear, semidefinite and copositive programming. J. Global Optim. 24, 163–185.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0043)

[Bomze, I.M., Duer, M., de Klerk, E., Roos, C., Quist, A., Terlaky, T., 2000. On copositive programming and standard quadratic optimization problems. J. Global Optim. 18, 301–320.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0044)

[Bonami, P., Günlük, O., Linderoth, J., 2018. Globally solving nonconvex quadratic pro- gramming problems with box constraints via integer programming methods. Math. Program. Comput. 10, 333–382.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0045)

[Bonami, P., Lodi, A., Schweiger, J., Tramontani, A., 2019. Solving quadratic programming by cutting planes. SIAM J. Optim. 29, 1076–1105.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0046)

[Borradaile, G., Van Hentenryck, P., 2005. Safe and tight linear estimators for global opti- mization. Math. Program. 102, 495–517.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0047)

[Bundfuss, S., Duer, M., 2009. An adaptive linear approximation algorithm for copositive programs. SIAM J. Optim. 20, 30–53.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0048)

[Burer, S., 2009. On the copositive representation of binary and continuous nonconvex quadratic programs. Math. Program. 120, 479–495.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0049)

[Burer, S., Anstreicher, K.M., 2013. Second-order-cone constraints for extended trust-region subproblems. SIAM J. Optim. 23 (1), 432–451.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0050)

[Burer, S., Dong, H., 2012. Representing quadratically constrained quadratic programs as generalized copositive programs. Oper. Res. Lett. 40, 203–206.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0051)

[Burer, S., Dong, H., 2013. Separation and relaxation for cones of quadratic forms. Math.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0052)

[Program. 137, 343–370.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0052)

[Burer, S., Kilinç-Karzan, F., 2017. How to convexify the intersection of a second order cone and a nonconvex quadratic. Math. Program. 162, 393–429.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0053)

[Burer, S., Kim, S., Kojima, M., 2014. Faster, but weaker, relaxations for quadratically constrained quadratic programs. Comput. Optim. Appl. 59, 27–45.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0054)

[Burer, S., Letchford, A.N., 2009. On nonconvex quadratic programming with box con- straints. SIAM J. Optim. 20, 1073–1089.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0055)

[Burer, S., Vandenbussche, D., 2008. A finite branch-and-bound algorithm for nonconvex quadratic programming via semidefinite relaxations. Math. Program. 113, 259–282.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0056)

[Burer, S., Vandenbussche, D., 2009. Globally solving box-constrained nonconvex quadratic programs with semidefinite-based finite branch-and-bound. Comput. Op- tim. Appl. 43, 181–195.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0057)

[Burer,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0058) [S., Yang, B., 2015. The trust region subproblem with non-intersecting linear con- straints. Math. Program. 149, 253–264.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0058)

[Burer,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0059) [S., Ye, Y., 2020. Exact semidefinite formulations for a class of (random and non-ran- dom) nonconvex quadratic programs. Math. Program. 181, 1–17.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0059)

[Cabassi,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0060) [F., Locatelli, M., 2016. Computational investigation of simple memetic ap- proaches for continuous global optimization. Comput. Oper. Res. 72, 50–70.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0060)

[Caprara,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0061) [A., Locatelli, M., 2010. Global optimization problems and domain reduction strategies. Math. Program. 125, 125–137.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0061)

[Caprara,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0062) [A., Locatelli, M., Monaci, M., 2016. Theoretical and computational results about optimality-based domain reductions. Comput. Optim. Appl. 64, 513–533.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0062)

[Carrizosa,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0063) [E., Hansen, P., Messine, F., 2004. Improving interval analysis bounds by trans- lations. J. Global Optim. 29, 157–172.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0063)

[Cartis,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0064) [C., Fowkes, J., Gould, N., 2015. Branching and bounding improvements for global optimization algorithms with Lipschitz continuity properties. J. Global Optim. 61, 429–457.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0064)

[Chen,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0065) [J., Burer, S., 2012. Globally solving nonconvex quadratic programming problems via completely positive programming. Math. Program. Comput. 4, 33–52.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0065)

[Consolini,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0066) [L., Locatelli, M., 2017. On the complexity of quadratic programming with two quadratic constraints. Math. Program. 164, 91–128.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0066)

[Csallner,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0067) [A.E., Csendes, T., Markót, M.C., 2000. Multisection in interval branch-and-bound methods for global optimization I. Theoretical results. J. Global Optim. 16, 371–392.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0067) [Cui,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0068) [L., Li, G., Lin, Q., Chen, J., Lu, N., 2016. Adaptive differential evolution algorithm with novel mutation strategies in multiple sub-populations. Comput. Oper. Res. 67,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0068)

[155–173.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0068)

[Dantzig,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0069) [G.B., 1960. On the significance of solving linear programming problems with some integer variables. Econometrica 28 (1), 30–44.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0069)

[Das,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0070) [S., Mullick, S., Suganthan, P., 2016. Recent advances in differential evolution-an updated survey. Swarm Evol. Comput. 27, 1–30.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0070)

[Davarnia,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0071) [D., Richard, J.-P., Tawarmalani, M., 2017. Simultaneous convexification of bi- linear functions over polytopes with application to network interdiction. SIAM J. Op- tim. 27, 1801–1833.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0071)

[de](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0072) [Angelis, P.L., Bomze, I.M., Toraldo, G., 2004. Ellipsoidal approach to box-constrained quadratic problems. J. Global Optim. 28, 1–15.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0072)

[de](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0073) [Klerk, E., Hess, R., Laurent, M., 2017. Improved convergence rates for Lasserre-type hierarchies of upper bounds for box-constrained polynomial optimization. SIAM J. Optim. 27, 347–367.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0073)

[de](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0074) [Klerk, E., Lasserre, J., Laurent, M., Sun, Z., 2017. Bound-constrained polynomial opti- mization using only elementary calculations. Math. Oper. Res. 42, 834–853.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0074)

[de](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0075) [Klerk, E., Laurent, M., Parrillo, P.A., 2006. A PTAS for the minimization of polynomials of fixed degree over the simplex. Theor. Comput. Sci. 361, 210–225.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0075)

[de](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0076) [Klerk, E., Pasechnik, D.V., 2002. Approximation of the stability number of a graph via copositive programming. SIAM J. Optim. 12, 875–892.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0076)

[Del](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0077) [Pia, A., Khajavirad, A., 2017. A polyhedral study of binary polynomial programs.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0077)

[Math. Oper. Res. 42, 389–410.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0077)

[Del](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0078) [Pia, A., Khajavirad, A., 2018. On decomposability of multilinear sets. Math. Program.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0078)

[170, 387–415.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0078)

[Del Ser, J.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0079), [Osaba, E.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0079), [Molina, D.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0079), [Yang, X.-S.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0079), [Salcedo-Sanz, S.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0079), [Camacho, D.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0079), [Das, S.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0079), [Suganthan, P., Coello Coello, C., Herrera, F., 2019. Bio-inspired computation: where we stand and what’s next. Swarm Evol. Comput 48, 220–250.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0079)

[Depetrini,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0080) [D., Locatelli, M., 2011. Approximation of linear fractional/multiplicative prob- lems. Math. Program. 128, 437–443.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0080)

[Di](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0081) [Carlo, M., Vasile, M., Minisci, E., 2020. Adaptive multi-population inflationary differ- ential evolution. Soft comput 24 (5), 3861–3891.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0081)

[Diouane,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0082) [Y., Gratton, S., Vicente, L.N., 2015. Globally convergent evolution strategies.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0082)

[Math. Program. 152 (1–2), 467–490.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0082)

[Diouane,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0083) [Y., Gratton, S., Vicente, L.N., 2015. Globally convergent evolution strategies for constrained optimization. Comput. Optim. Appl. 62 (2), 323–346.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0083)

[Dixon,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0084) [L., 1975. Towards global optimisation. In: Dixon, L., Szegö, G. (Eds.), Towards Global Optimization. North-Holland, pp. 29–54.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0084)

[Dixon, L., Szegö, G. (Eds.), Towards global optimisation. proceedings of a workshop at the University of Cagliari, Italy, October 1974, 1975. North-Holland.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0085)

[Dixon, L., Szegö, G. (Eds.), 1978. Towards Global Optimisation 2. North-Holland](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0086). [Domes,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0087) [F., Neumaier, A., 2016. Constraint aggregation for rigorous global optimization.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0087)

[Math. Program. 155, 375–401.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0087)

[Evtushenko,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0088) [Y.P., 1971. Numerical method for finding the global extremum of a function.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0088)

[Vyl. Mat. I Mat. Fiz. 11, 1390–1403.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0088) [(in Russian)](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0088)

[Falk,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0089) [J.E., Soland, R.M., 1969. An algorithm for separable nonconvex programming prob- lems. Manage. Sci. 15 (9), 550–569.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0089)

[Feliot,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0090) [P., Bect, J., Vazquez, E., 2017. A Bayesian approach to constrained single- and multi-objective optimization. J. Global Optim. 67 (1), 97–133.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0090)

[Ferreiro-Ferreiro, A.M., García-Rodríguez, J.A., Souto, L., Vázquez, C., 2019. Basin Hop- ping with synched multi L-BFGS local searches. Parallel implementation in multi-CPU and GPUs. Appl. Math. Comput. 356, 282–298.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0091)

[Floudas, C.A.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0092), [Pardalos, P.M.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0092), [Adjiman, C.S.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0092), [Esposito, W.R.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0092), [Gumus, Z.H.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0092), [Harding, S.T.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0092), [Klepeis, J.L., Meyer, C.A., Schweiger, C., 1999. Handbook of Test Problems in Local and Global Optimization. Nonconvex Optimization and its Applications, 33. Kluwer Academic Publishers, Dordrecht.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0092)

[Fowkes,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0093) [J., Gould, N., Farmer, C., 2013. A branch and bound algorithm for the global opti- mization of Hessian Lipschitz continuous functions. J. Global Optim. 56, 1791–1815.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0093) [Frazier,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0094) [P.I., 2018. Bayesian optimization. In: Gel, E., Ntaimo, L., Shier, D., Greenberg, H.J. (Eds.), Recent Advances in Optimization and Modeling of Contemporary Problems.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0094)

[INFORMS, pp. 255–278.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0094)

[Furini, F.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0095), [Traversi, E.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0095), [Belotti, P.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0095), [Frangioni, A.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0095), [Gleixner, A.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0095), [Gould, N.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0095), [Liberti, L.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0095), [Lodi, A.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0095), [Misener, R., Mittelmann, H., Sahinidis, N., Vigerske, S., Wiegele, A., 2019. QPLIB: A library of quadratic programming instances. Math. Program. Comput. 11, 237– 265.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0095)

[Gamrath, G.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0096), [Anderson, D.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0096), [Bestuzheva, K.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0096), [Chen, W.-K.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0096), [Eifler, L.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0096), [Gasse, M.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0096), [Gemander, P.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0096), [Gleixner, A., Gottwald, L., Halbig, K., Hendel, G., Hojny, C., Koch, T., Le Bodic, P., Ma- her, S.,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0096) [Matter, F.,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0096) [Miltenberger, M.,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0096) [Mühmer, E.,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0096) [Müller, B.,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0096) [Pfetsch, M.,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0096) [Schlösser, F.,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0096) [Serrano, F.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0096), [Shinano, Y.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0096), [Tawfik, C.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0096), [Vigerske, S.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0096), [Wegscheider, F.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0096), [Weninger, D.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0096), [Witzig, J., 2020. The SCIP Optimization Suite 7.0. Technical Report. Optimization Online.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0096)

[Garrido-Merchán, E.C., Hernández-Lobato, D., 2020. Dealing with categorical and integer– valued variables in Bayesian optimization with Gaussian processes. Neurocomputing 380, 20–35.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0097)

Gavana, A., Global Optimization Benchmarks. <http://infinity77.net/go_2021/>. Accessed: 2021-04-02.

[Gaviano, M., 1975. Some general results on the convergence of random search algorithms in minimisation problems. In: Dixon, L., Szegö, G. (Eds.), Towards Global Optimiza- tion. North-Holland, pp. 149–157.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0099)

[Ghosh, S., Das, S., Vasilakos, A.V., Suresh, K., 2012. On convergence of differential evo- lution over a class of continuous functions with unique global optimum. IEEE Trans. Syst. Man Cybern.Part B (Cybernetics) 42 (1), 107–124.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0100)

[Gleixner, A., Berthold, T., Müller, B., Weltge, S., 2017. Three enhancements for optimiza- tion-based bound tightening. J. Global Optim. 67, 731–757.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0101)

[Grosso, A., Locatelli, M., Schoen, F., 2007. A population-based approach for hard global optimization problems based on dissimilarity measures. Math. Program. 110 (2), 373–404.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0103)

[Hadi, A.A., Mohamed, A.W., Jambi, K.M., 2019. LSHADE-SPA Memetic framework for solving large-scale optimization problems. Complex Intell. Syst. 5 (1), 25–40.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0104)

[Hager, W., Phan, D., 2009. An ellipsoidal branch and bound algorithm for global opti- mization. SIAM J. Optim. 20, 740–758.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0105)

[Hansen, N., 2006. The CMA evolution strategy: a comparing review. In: Lozano, J.A., Lar- rañaga, P., Inza, I., Bengoetxea, E. (Eds.), Towards a New Evolutionary Computation: Advances in the Estimation of Distribution Algorithms. Springer, Berlin, Heidelberg,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0106)

[pp. 75–102.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0106)

[Hansen, N., Ostermeier, A., 2001. Completely derandomized self-adaptation in evolution strategies. Evol. Comput. 9 (2), 159–195.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0107)

[Hansen, P., Lagouanelle, J.-L., Messine, F., 2007. Comparison between Baumann and ad- missible simplex forms in interval analysis. J. Global Optim. 37, 215–228.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0108)

[Hendrix, E.M.T., Tóth, B., 2010. Introduction to Nonlinear and Global Optimization.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0109)

[Springer, New York, NY.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0109)

[Hernández-Lobato, J.M., Gelbart, M.A., Adams, R.P., Hoffman, M.W., Ghahramani, Z., 2016. A general framework for constrained bayesian optimization using informa- tion-based search. J. Mach. Learn. Res. 17 (160), 1–53.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0110)

[Ho-Nguyen, N., Kilinç-Karzan, F., 2017. A Second-Order Cone based approach for solving the trust-region subproblem and its variants. SIAM J. Optim. 27, 1485–1512.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0111)

[Holland, J.H., 1975. Adaptation in Natural and Artificial Systems: An Introductory Anal- ysis with Applications to Biology, Control, and Artificial Intelligence. University of Michigan Press.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0112)

[Horst, R., 1976. An algorithm for nonconvex programming problems. Math. Program. 10 (1), 312–321.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0113)

[Horst, R., Pardalos, P.M. (Eds.), 1995. Handbook of Global Optimization. Kluwer Aca- demic Publishers, Dordrecht.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0114)

[Horst, R., Pardalos, P.M., Thoai, N.V., 2001. Introduction to Global Optimization. Non- convex Optimization and Its Applications, 48. Springer Verlag.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0115)

[Horst, R., Tuy, H., 1993. Global Optimization: Deterministic Approaches, 2nd edition Springer Verlag.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0116)

[Gondzio, J., Yildrim, E., 2021. Global solutions of nonconvex Standard Quadratic Pro- grams via Mixed Integer Linear Programming reformulations. J. Glob. Optim 81, 293–321.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0113q)

IEEE TfLsgo. IEEE Task Force on Large-Scale Global Optimization. [https://www.tflsgo.org](https://www.tflsgo.org/)Accessed: 2021-04-02.

[Jach, M., Michaels, D., Weismantel, R., 2008. The convex envelope of (*𝑛*–1)-convex func-](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0118)

[tions. SIAM J. Optim. 19, 1451–1466.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0118)

[Jansson, C., Chaykin, D., Keil, C., 2007. Rigorous error bounds for the optimal value in semidefinite programming. SIAM J. Numer. Anal. 46, 180–200.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0119)

[Jeyakumar, V., Li, G., 2014. Trust-region problems with linear inequality constraints: ex- act SDP relaxation, global optimality and robust optimization. Math. Program. 147, 171–206.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0120)

[Jeyakumar, V., Li, G., 2018. Exact second-order cone programming relaxations for some nonconvex minimax quadratic optimization problems. SIAM J. Optim. 28, 760– 787.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0121)

[Jiang, R., Li, D., Wu, B., 2018. SOCP reformulation for the generalized trust region subproblem via a canonical form of two symmetric matrices. Math. Program. 169, 531–563.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0122)

[Johnson, S. G., The NLopt nonlinear-optimization package. http://github.com/stevengj/ nlopt. Accessed: 2021-04-02.](http://github.com/stevengj/nlopt)

[Jones, D.R., Schonlau, M., Welch, W.J., 1998. Eﬃcient global optimization of expensive black-box functions. J. Global Optim. 13 (4), 455–492.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0124)

Jones, E., Oliphant, T., Peterson, P., et al., 2001–. SciPy: Open source scientific tools for Python. Accessed: 2021-04-02.

[Kandasamy, K.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0126), [Vysyaraju, K.R.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0126), [Neiswanger, W.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0126), [Paria, B.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0126), [Collins, C.R.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0126), [Schneider, J.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0126), [Poczos, B., Xing, E.P., 2020. Tuning hyperparameters without grad students: scalable and Robust Bayesian optimisation with dragonfly. J. Mach. Learn. Res. 21, 1–27.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0126)

[Kawaguchi, K., 2016. Deep Learning without Poor Local Minima. Technical Report arXiv:1605.07110v3. Massachusetts Institute of Technology.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0127)

[Kazazakis, N., Adjiman, C., 2018. Arbitrarily tight *𝛼*-BB underestimators of general non–](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0128)

[linear functions over sub-optimal domains. J. Global Optim. 71, 815–844.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0128)

[Kearfott, R.B., 2011. Interval computations, rigor and non-rigor in deterministic continu- ous global optimization. Optim. Methods Softw. 26, 259–279.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0129)

[Khajavirad,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0130) [A., Michalek, J., Sahinidis, N., 2014. Relaxations of factorable functions with convex-transformable intermediates. Math. Program. 144, 107–140.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0130)

[Khajavirad,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0131) [A., Sahinidis, N., 2012. Convex envelopes of products of convex and compo- nent-wise concave functions. J. Global Optim. 52, 391–409.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0131)

[Khajavirad,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0132) [A., Sahinidis, N.V., 2013. Convex envelopes generated from finitely many compact convex sets. Math. Program. 137, 371–408.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0132)

[Kim,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0133) [M.J., 2020. Variance regularization in sequential Bayesian optimization. Math. Oper.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0133)

[Res. 45 (3), 966–992.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0133)

[Kim,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0134) [S., Kojima, M., 2003. Exact solutions of some nonconvex quadratic optimization problems via SDP and SOCP relaxations. Comput. Optim. Appl. 26, 143–154.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0134)

[Kuang,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0135) [X., Zuluaga, L., 2018. Completely positive and completely positive semidefinite tensor relaxations for polynomial optimization. J. Global Optim. 70, 551–577.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0135)

[Kuno,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0136) [T., 2005. A revision of the trapezoidal branch-and-bound algorithm for linear sum-of-ratios problems. J. Global Optim. 33, 215–234.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0136)

[Kushner,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0137) [H.J., 1964. A new method of locating the maximum point of an arbitrary multi- peak curve in the presence of noise. J. Basic Engi.g 97–106.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0137)

[Lasserre,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0138) [J.B., 2001. Global optimization with polynomials and the problem of moments.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0138)

[SIAM J. Optim. 11, 796–817.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0138)

[Lasserre,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0139) [J.B., 2005. Polynomial programming: LP-relaxations also converge. SIAM J. Op- tim. 15, 383–393.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0139)

[Lasserre,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0140) [J.B., 2006. A sum of squares approximation of nonnegative polynomials. SIAM](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0140)

[J. Optim. 16, 751–765.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0140)

[Laurent,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0141) [M., 2007. Semidefinite representations for finite varieties. Math. Program. 109, 1–26.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0141)

[Le](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0142) [Thi, H.A., 2000. An eﬃcient algorithm for globally minimizing a quadratic function under convex quadratic constraints. Math. Program. 87, 401–426.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0142)

[Linderoth,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0143) [J., 2005. A simplicial branch-and-bound algorithm for solving quadratically constrained quadratic programs. Math. Program. 103, 251–282.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0143)

[Liuzzi,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0144) [G., Locatelli, M., Piccialli, V., 2019. A new branch-and-bound algorithm for stan- dard quadratic programming problems. Optim. Methods Softw. 34, 79–97.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0144)

[Liuzzi,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0145) [G., Locatelli, M., Piccialli, V., Rass, S., 2021. Computing mixed strategies equilibria in presence of switching costs by the solution of nonconvex QP problems. Comput. Optim..](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0145) [To appear](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0145)

[Locatelli,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0146) [M., 2013. Approximation algorithm for a class of global optimization problems.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0146)

[J. Global Optim. 55, 13–25.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0146)

[Locatelli,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0147) [M., 2015. Alternative branching rules for some nonconvex problems. Optim.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0147)

[Methods Softw. 30, 365–378.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0147)

[Locatelli,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0148) [M., 2015. Some results for quadratic problems with one or two quadratic con- straints. Oper. Res. Lett. 43, 126–131.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0148)

[Locatelli,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0149) [M., 2016. Exactness conditions for an SDP relaxation of the extended trust region problem. Optim. Lett. 10, 1141–1151.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0149)

[Locatelli,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0150) [M., 2016. Non polyhedral convex envelopes for 1-convex functions. J. Global Optim. 65, 637–655.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0150)

[Locatelli,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0151) [M., 2020. Convex envelope of bivariate cubic functions over rectangular regions.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0151)

[J. Global Optim. 76, 1–24.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0151)

[Locatelli,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0152) [M., Schoen, F., 2013. Global Optimization: Theory, Algorithms, and Applica- tions. SIAM-MOS. SIAM.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0152)

[Locatelli,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0153) [M., Vasile, M., 2015. (Non) convergence results for the differential evolution method. Optim. Lett. 9 (3), 413–425.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0153)

[Lofberg,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0154) [J., 2004. YALMIP : a toolbox for modeling and optimization in MATLAB. In: 2004 IEEE International Conference on Robotics and Automation (IEEE Cat. No.04CH37508), pp. 284–289.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0154)

[Luedtke,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0155) [J., D’Ambrosio, C., Linderoth, J., Schweiger, J., 2020. Strong convex nonlinear relaxations of the pooling problem. SIAM J. Optim. 30, 1582–1609.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0155)

[Luedtke,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0156) [J., Namazifar, M., Linderoth, J., 2012. Some results on the strength of relaxations of multilinear functions. Math. Program. 136, 325–351.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0156)

[Ma,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0157) [X., Li, X., Zhang, Q., Tang, K., Liang, Z., Xie, W., Zhu, Z., 2019. A survey on cooperative co-evolutionary algorithms. IEEE Trans. Evol. Comput. 23 (3), 421–441.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0157)

[Mansueto,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0158) [P., Schoen, F., 2021. Memetic differential evolution methods for clustering problems. Pattern Recognit 114, 107849.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0158)

[Markot,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0159) [M., Schichl, H., 2014. Bound constrained interval global optimization in the CO- CONUT environment. J. Global Optim. 60, 751–776.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0159)

[Markót,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0160) [M.C., Csendes, T., 2005. A new verified optimization technique for the “packing circles in a unit square” problem. SIAM J. Optim. 16, 193–219.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0160)

[Markót,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0161) [M.C., Fernandez, J., Casado, L.G., Csendes, T., 2006. New interval methods for constrained global optimization. Math. Program. 106, 287–318.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0161)

[Martinez,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0162) [J., Casado, L., Garcia, I., Sergeyev, Y., Toth, B., 2004. On an eﬃcient use of gradient information for accelerating interval global optimization algorithms. Numer. Alg. 37, 61–69.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0162)

[Martinez-Cantin, R., 2014. Bayesopt: a bayesian optimization library for nonlinear opti- mization, experimental design and bandits. J. Mach. Learn. Res. 15 (115), 3915–3919.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0163) [Mathesen,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0164) [L., Pedrielli, G., Ng, S.H., Zabinsky, Z.B., 2020. Stochastic optimization with adaptive restart: a framework for integrated local and global learning. J. Global Op-](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0164)

[tim..](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0164)

[McCormick,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0165) [G., 1972. Attempts to calculate global solutions of problems that may have local minima. In: Lootsma, F. (Ed.), Numerical Methods for Non-Linear Optimization. Academic Press, pp. 209–221.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0165)

[McCormick,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0166) [G.P., 1976. Computability of global solutions to factorable nonconvex pro- grams: Part I Convex underestimating problems. Math. Program. 10 (1), 147–175.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0166)

[Meldgaard,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0167) [S.A., Kolsbjerg, E.L., Hammer, B., 2018. Machine learning enhanced global optimization by clustering local environments to enable bundled atomic energies. J. Chem. Phys. 149 (13), 134104.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0167)

[Messine,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0168) [F., 2004. Deterministic global optimization using interval constraint propagation techniques. RAIRO-Oper. Res. 38 (4), 277–293.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0168)

[Meyer,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0169) [C.A., Floudas, C.A., 2005. Convex envelopes for edge-concave functions. Math.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0169)

[Programm. B 103, 207–224.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0169)

[tiable functions by piecewise quadratic perturbation: spline *𝛼*BB underestimators. J. Meyer, C.A., Floudas, C.A., 2005. Convex underestimation of twice continuously differen-](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0170)

[Global Optim. 32, 221–258.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0170)

[Mittel, S., Schulz, A., 2013. An FPTAS for optimizing a class of low-rank functions over a polytope. Math. Program. 141, 103–120.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0171)

[Mockus, J., 1975. On Bayesian methods of optimisation. In: Dixon, L., Szegö, G. (Eds.), Towards Global Optimization. North-Holland, pp. 166–181.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0172)

[Moriconi, R., Kumar, K.S.S., Deisenroth, M.P., 2020. High-dimensional Bayesian optimiza- tion with projections using quantile Gaussian processes. Optim. Lett. 14 (1), 51–64.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0173)

[Motzkin, T.S., Strauss, E.G., 1965. Maxima for graphs and a new proof of a theorem of Turán. Can. J. Math. 17, 533–540.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0174)

Neumaier, A., Global optimiztion. <https://www.mat.univie.ac.at/~neum/glopt.html> Ac- cessed: 2021-04-02.

[Neumaier, A., 2004. Complete search in continuous global optimization and constraint satisfaction. In: Acta Numerica 2004. Cambridge University Press, pp. 271–369.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0176)

[Neumaier, A., Shcherbina, O., 2004. Safe bounds in linear and mixed-integer program- ming. Math. Program. 99, 283–296.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0177)

[Neumaier, A., Shcherbina, O., Huyer, W., Vinkó, T., 2005. A comparison of complete global optimization solvers. Math. Program. 103 (2), 335–356.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0178)

[Nguyen, T., Richard, J.-P., Tawarmalani, M., 2018. Deriving convex hulls through lifting and projection. Math. Program. 169, 377–415.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0179)

[Nie, J., 2013. Certifying convergence of Lasserres hierarchy via flat truncation. Math.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0180)

[Program. 142, 485–510.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0180)

[Nie, J., 2014. Optimality conditions and finite convergence of Lasserres hierarchy. Math.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0181)

[Program. 146, 97–121.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0181)

[Nie, J., Demmel, J., Sturmfels, B., 2006. Minimizing polynomials via sum of squares over the gradient ideal. Math. Program. 106, 587–606.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0182)

[Padberg, M., 1989. The boolean quadric polytope: some characteristics, facets and rela- tives. Math. Program. 45, 139–172.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0183)

[Papp, D., Yildiz, S., 2019. Sum-of-squares optimization without semidefinite program- ming. SIAM J. Optim. 29, 822–851.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0184)

[Parrillo, P.A., 2000. Structured Semidefinite Programs and Semi-algebraic Geome- try Methods in Robustness and Optimization. California Institute of Technology, Pasadena.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0185)

[Paulavic˘ ius, R., Zi˘linskas, J., 2014. Simplicial Global Optimization. Springer, New York,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0186)

[NY.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0186)

[Peña, J., Vera, J., Zuluaga, L.F., 2007. Computing the stability number of a graph via linear and semidefinite programming. SIAM J. Optim. 18, 87–105.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0187)

[Piyavskii, S.A., 1972. An algorithm for finding the absolute extremum of a function. USSR Comput. Math. Math.Phys. 12 (4), 57–67.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0188)

[Pong, T., Wolkowicz, H., 2014. The generalized trust region subproblem. Comput. Optim.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0189)

[Appl. 58, 273–322.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0189)

[Puranik, Y., Sahinidis, N., 2017. Bounds tightening based on optimality conditions for nonconvex box-constrained optimization. J. Global Optim. 67, 59–77.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0190)

[Rendl, F., Wolkowicz, H., 1997. A semidefinite framework for trust region subproblems with applications to large scale minimization. Math. Program. 77, 273–299.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0191)

[Rikun, A.D., 1997. A convex envelope formula for multilinear functions. J. Global Optim.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0192)

[10, 425–437.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0192)

[Rios, L., Sahinidis, N., 2013. Derivative-free optimization: a review of algorithms and comparison of software implementations. J. Global Optim. 56, 1247–1293.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0193)

Sahinidis, N., 2017. BARON 17.8.9: Global Optimization of Mixed-Integer Nonlinear Pro- grams, *User’s Manual*.

[Sakaue, S., Nakatsukasa, Y., Takeda, A., Iwata, S., 2016. Solving generalized CDT problems via two-parameter eigenvalues. SIAM J. Optim. 26, 1669–1694.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0195)

[Santana, A., Dey, S., 2020. The convex hull of a quadratic constraint over a polytope.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0196)

[SIAM J. Optim. 30, 2983–2997.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0196)

[Schichl, H., Neumaier, A., 2005. Interval analysis on directed acyclic graphs for global optimization. J. Global Optim. 33, 541–562.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0197)

[Schoen, F., Tigli, L., 2021. Eﬃcient large scale global optimization through cluster- ing-based population methods. Comput. Oper. Res. 127 (105165).](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0198)

[Scott, J.K., Stuber, M.D., Barton, P.I., 2011. Generalized McCormick relaxations. J. Global Optim. 51, 569–606.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0199)

[Sherali, H.D., Tuncbilek, C.H., 1992. A global optimization algorithm for polynomial pro- gramming problems using a reformulation-linearization technique. J. Global Optim. 2, 101–112.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0200)

[Shubert, B.O., 1972. A sequential method seeking the global maximum of a function. SIAM](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0201)

[J. Numer. Anal. 9.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0201)

[Škvorc, U., Eftimov, T., Korošec, P., 2019. CEC real-parameter optimization competitions: Progress from 2013 to 2018. In: 2019 IEEE Congress on Evolutionary Computation (CEC), pp. 3126–3133.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0202)

[Snoek,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0203) [J., Ripped, O., Swersky, K., Kiros, R., Satish, N., Sundaram, N., Patwary, M., Prab-](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0203) [hat, Adams, R., 2015. Scalable Bayesian optimization using deep neural networks, 3,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0203)

[pp. 2161–2170.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0203)

[Sojoudi,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0204) [S., Lavaei, J., 2014. Exactness of semidefinite relaxations for nonlinear optimiza- tion problems with underlying graph structure. SIAM J. Optim. 24, 1746–1778.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0204)

[Soland,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0205) [R.M., 1971. An algorithm for separable nonconvex programming problems II: nonconvex constraints. Manage. Sci. 17 (11), 759–773.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0205)

[Sturm,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0206) [J.F., Zhang, S., 2003. On cones of nonnegative quadratic functions. Math. Oper.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0206)

[Res. 28 (2), 246–267.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0206)

[Sun,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0207) [Y., Li, X., Ernst, A., Omidvar, M.N., 2019. Decomposition for large-scale optimiza- tion problems with overlapping components. In: 2019 IEEE Congress on Evolutionary Computation (CEC), pp. 326–333.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0207)

[Tardella,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0208) [F., 2003. On the existence of polyhedral convex envelopes. In: Frontiers in Global Optimization. Kluwer Academic Publishers, pp. 563–574.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0208)

[Tardella,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0209) [F., 2008. Existence and sum decomposition of vertex polyhedral convex en- velopes. Optim. Lett. 2, 363–375.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0209)

[Tawarmalani,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0210) [M., Richard, J.-P.P., Xiong, C., 2013. Explicit convex and concave envelopes through polyhedral subdivisions. Math. Program. 138, 531–577.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0210)

[Tawarmalani,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0211) [M., Sahinidis, N.V., 2001. Semidefinite relaxations of fractional programs via novel convexification techniques. J. Global Optim. 20, 137–158.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0211)

[Tawarmalani,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0212) [M., Sahinidis, N.V., 2004. Global optimization of mixed-integer nonlinear programs: a theoretical and computational study. Math. Program. 99 (3), 563–591.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0212)

[Törn,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0213) [A.A., 1978. A search–clustering approach to global optimization. In: Towards Global Optimization 2. North–Holland, pp. 49–62.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0213)

[Tóth,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0214) [B., Hendrix, E., Casado, L., Garcia, I., 2016. On refinement of the unit simplex using regular simplices. J. Global Optim. 64, 305–323.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0214)

[Tóth,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0215) [B., Hendrix, E.M.T., Casado, L.G., 2021. On monotonicity and search strategies in face-based copositivity detection algorithms. Central Eur. J. Oper. Res. available on- line.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0215)

[Tran,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0216) [N., Schneider, J.-G., Weber, I., Qin, A.K., 2020. Hyper-parameter optimization in classification: to-do or not-to-do. Pattern Recognit. 103, 107245.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0216)

[Vandenbussche, D., Nemhauser, G.L., 2005. A branch-and-cut algorithm for nonconvex quadratic programs with box constraints. Math. Program. 102, 559–575.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0217)

[Vandenbussche, D., Nemhauser, G.L., 2005. A polyhedral study of nonconvex quadratic programs with box constraints. Math. Program. 102, 531–557.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0218)

[Vinkó,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0219) [T., Gelle, K., 2017. Basin Hopping Networks of continuous global optimization problems. Central Eur. J. Oper. Res. 25 (4), 985–1006.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0219)

[Vu,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0220) [X.-H., Schichl, H., Sam-Haroud, D., 2009. Interval propagation and search on directed acyclic graphs for numerical constraint solving. J. Global Optim. 45, 499–531.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0220)

[Vui,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0221) [H.H., So’n, P.T., 2008. Global optimization of polynomials using the truncated tan- gency variety and sums of squares. SIAM J. Optim. 19, 941–951.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0221)

[Wales, D.J., 2018. Exploring energy landscapes. Annu. Rev. Phys. Chem. 69 (1), 401–425](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0222). [Wang,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0223) [A., Kilinç-Karzan, F., 2020. The generalized trust region subproblem: solution com-](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0223)

[plexity and convex hull results. Math. Program. available online.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0223)

[Wang,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0224) [A., Kilinç-Karzan, F., 2021. On the tightness of SDP relaxations of QCQPs. Math.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0224)

[Program. to appear.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0224)

[Wang,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0225) [Z., Hutter, F., Zoghi, M., Matheson, D., Feitas, N.d., 2016. Bayesian optimization in a billion dimensions via random embeddings. J. Artif. Intell. Res. 55, 361–387.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0225)

[Wechsung,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0226) [A., Scott, J., Watson, H., Barton, P., 2015. Reverse propagation of McCormick relaxations. J. Global Optim. 63, 1–36.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0226)

[Wu,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0227) [G., Mallipeddi, R., Suganthan, P.N., Wang, R., Chen, H., 2016. Differential evolution with multi-population based ensemble of mutation strategies. Inf. Sci. 329, 329–345.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0227)

[Xia,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0228) [W., Vera, J., Zuluaga, L., 2020. Globally solving nonconvex quadratic programs via linear integer programming techniques. INFORMS J. Comput. 32, 40–56.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0228)

[Yang,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0229) [B., Anstreicher, K., Burer, S., 2018. Quadratic programs with hollows. Math. Pro- gram. 170, 541–553.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0229)

[Yang,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0230) [B., Burer, S., 2016. A two-variable approach to the two-trust region subproblem.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0230)

[SIAM J. Optim. 26, 661–680.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0230)

[Yang,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0231) [L., Shami, A., 2020. On hyperparameter optimization of machine learning algo- rithms: theory and practice. Neurocomputing 415, 295–316.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0231)

[Ye,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0232) [Y., Zhang, S., 2003. New results on quadratic minimization. SIAM J. Optim. 14, 245–267.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0232)

[Yuan,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0233) [J., Wang, M., Ai, W., Shuai, T., 2017. New results on narrowing the duality gap of the extended Celis-Dennis-Tapia problem. SIAM J. Optim. 27, 890–909.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0233)

[Zhan,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0234) [D., Xing, H., 2020. Expected improvement for expensive optimization: a review. J. Global Optim. 78 (3), 507–544.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0234)

[Zhang,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0235) [Y., Sahinidis, N., Nohra, C., Rong, G., 2020. Optimality-based domain reduction for inequality-constrained NLP and MINLP problems. J. Global Optim. 77, 425–454.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0235)

[Zhao,](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0236) [Y., Chen, X., Li, J., 2017. TGMin: a global-minimum structure search program based on a constrained basin-hopping algorithm. Nano Res. 10 (10), 3407–3420.](http://refhub.elsevier.com/S2192-4406(21)00139-8/sbref0236)