



JONAS SCHWEIGER¹

Exploiting structure in non-convex quadratic optimization

¹  0000-0002-4685-9748

Zuse Institute Berlin
Takustr. 7
14195 Berlin
Germany

Telephone: +49 30-84185-0
Telefax: +49 30-84185-125

E-mail: bibliothek@zib.de
URL: <http://www.zib.de>

ZIB-Report (Print) ISSN 1438-0064
ZIB-Report (Internet) ISSN 2192-7782

Exploiting structure in non-convex quadratic optimization

Jonas Schweiger*

Zuse Institute Berlin, schweiger@zib.de

Abstract The amazing success of computational mathematical optimization over the last decades has been driven more by insights into mathematical structures than by the advance of computing technology. In this vein, we address applications, where nonconvexity in the model poses principal difficulties.

This paper summarizes the dissertation [Sch17] for the occasion of the GOR dissertation award 2018. We focus on the work on non-convex quadratic programs and show how problem specific structure can be used to obtain tight relaxations and speed up Branch-and-bound methods. Both a classic general QP and the Pooling Problem as an important practical application serve as showcases.

Keywords: Nonconvexity, Quadratic Programming, Relaxations, Cutting Planes, Standard Quadratic Programming, Pooling Problem

Profound knowledge of the mathematical structures is the key to design algorithms that perform practical computations in the most efficient way. A thorough exploitation of such structures is why solver software for so-called linear and mixed-integer linear programs routinely computes optimal solutions for practical problems with hundreds of thousands of variables.

This paper focuses on the contributions to non-convex quadratic programming relaxations from the dissertation [Sch17] for the occasion of the GOR dissertation award 2018. The key results on gas network planning under uncertainty from this dissertation were already presented in this series [Sch16].

Branch-and-bound methods for non-convex quadratic programs (QP) depend on tight relaxations. We contribute in several ways: First, we establish a new way to handle missing linearization variables in the well-known Reformulation-Linearization-Technique (RLT). This is implemented into the commercial software CPLEX. Second, we study the optimization of a quadratic objective over the standard simplex. These basic structure appears as part of many complex models. Exploiting connections to the maximum clique problem and RLT, we derive new valid and strong inequalities. Using exact and heuristic separation methods, we demonstrate the impact of the new inequalities on the relaxation and the global optimization of these problems. Third, we strengthen the state-of-the-art relaxation for the Pooling Problem, a well-known non-convex quadratic

* OrcID: 0000-0002-4685-9748

problem, which is, for example, relevant in the petrochemical industry. We propose a novel relaxation that captures the essential non-convex structure of the problem but is small enough for an in-depth study. We provide a complete inner description in terms of the extreme points as well as an outer description in terms of inequalities defining its convex hull (which is not a polyhedron). We show that the resulting valid convex inequalities significantly strengthen the standard relaxation of the pooling problem.

1 Reformulations and relaxations for quadratic programs

Quadratic programming deals with the following mathematical program:

$$\begin{aligned} \min_{x \in \mathbb{R}^d} \quad & x^T Q_0 x + a_0 x + b_0 \\ \text{s.t.} \quad & x^T Q_i x + a_i x + b_i \leq 0 && \text{for all } i \in \mathcal{M} \\ & \underline{x}_j \leq x_j \leq \bar{x}_j && \text{for all } j \in \mathcal{N} \end{aligned}$$

The matrices $Q_i \in \mathbb{R}^{d \times d}$, $i \in \mathcal{M} \cup \{0\}$ are symmetric, but not necessarily positive semidefinite, such that the problem is in general non-convex.

Termwise linearization of all bilinear and quadratic terms is the most common way to convexify quadratic programs. To this end an auxiliary variable Y_{ij} is introduced for each term $x_i x_j$ in the model together with the constraint $Y_{ij} = x_i x_j$. Quadratic terms $Q_{ii} x_i^2$ are over- and underestimated using gradient and secant hyperplanes that might be separated during the solution process. For bilinear terms, the equation $Y_{ij} = x_i x_j$ is non-convex and can be convexified by the so-called *McCormick* inequalities [McC76]

$$\begin{aligned} \bar{x}_j x_i + \underline{x}_i x_j - \bar{x}_i \bar{x}_j &\leq Y_{ij}, & \bar{x}_j x_i + \underline{x}_i x_j - \underline{x}_i \bar{x}_j &\geq Y_{ij}, \\ \underline{x}_j x_i + \underline{x}_i x_j - \underline{x}_i \underline{x}_j &\leq Y_{ij}, & \underline{x}_j x_i + \bar{x}_i x_j - \bar{x}_i \underline{x}_j &\geq Y_{ij}. \end{aligned}$$

In this fashion all terms appearing in the model are relaxed.

The major drawback of the termwise McCormick relaxation is that each term is relaxed independently, disrespecting the interactions of various constraints. The Reformulation-Linearization Technique (RLT) [SA92] is a way to combine several aspects of the model and derive new valid inequalities.

RLT consists of two steps. In the first step, valid constraints are multiplied by other constraints or by variables yielding an equation or inequality with higher order terms. In the second step, these terms are reformulated using linearization variables to obtain a linear constraint. The result is a valid constraint on the linearization variables that is often a very strong cutting plane [MF13].

Our contribution is to propose a treatment for those quadratic terms that arise from RLT but do not appear in the model so far. To avoid the introduction of new linearization variables and still be able to formulate a valid inequality, we propose to project them out by replacing them with appropriate over- and underestimators, an approach that has not been described or applied before to the best of our knowledge. An implementation of this is implemented into the

commercial solver CPLEX and enabled by default for problems with non-convex quadratic objective. On the CPLEX test set at IBM comprising several hundred problems, projected RLT is instrumental for solving eleven additional instances in comparison to version 12.6.3. When looking at models that are affected by the separation of these inequalities, a run time reduction of 29% is achieved. This rises to 84% when focusing on “hard instances” where either of the two versions takes at least 1000 seconds to solve the problem. This work on RLT and the work of the next section is joint work with Pierre Bonami, Andrea Lodi, and Andrea Tramontani [Bon+16].

2 Motzkin-Straus inequalities for Standard Quadratic Programming and generalizations

Next, we study more specific structures that appear as part of many complex models. The *standard simplex* is the set $\Delta = \left\{x \in \mathbb{R}^d \mid \sum_{i=1}^d x_i = 1, x \geq 0\right\}$. We study the optimization of a quadratic function over the standard simplex; a problem called *Standard Quadratic Program* [Bom98]:

$$\min \{x^T Qx \mid x \in \Delta\}. \quad (\text{StQP})$$

The next theorem connects the clique number of a graph with StQP.

Theorem 1 (Motzkin-Straus [MS65]). *Let A be the adjacency matrix of a simple, undirected graph G with d nodes and $\omega(G)$ its clique number. Then, the following relation holds:*

$$\max \{x^T Ax \mid x \in \Delta\} = 1 - \frac{1}{\omega(G)}.$$

Again, we assume the quadratic function has been reformulated with auxiliary linearization variables Y and we strive to use Theorem 1 to derive strong valid inequalities for the set

$$\Gamma = \{(x, Y) \in \mathbb{R}^d \times (\mathbb{R}^d \times \mathbb{R}^d) \mid Y = xx^T, x \in \Delta\}.$$

Corollary 1. *For any simple, undirected graph G with adjacency matrix A and clique number $\omega(G)$, the following inequality is valid for $(x, Y) \in \Gamma$:*

$$\langle A, Y \rangle \leq 1 - \frac{1}{\omega(G)}$$

In the remainder, we call the inequalities derived from Corollary 1 *Motzkin-Straus Clique inequalities (MSC inequalities)*.

We propose several heuristic methods to separate violated MSC inequalities. For complete bipartite graphs, the separation can be formulated as a binary QP which gives an exact separation algorithm for this class of inequalities. We show that the relaxation is considerably strengthened and that the separation program for complete bipartite graphs can be solved very efficiently by state-of-the-art solvers.

2.1 Generalized MSC inequalities for bipartite graphs

By performing a specific aggregation of RLT inequalities, we can obtain a non-linear inequality that generalizes Motzkin-Straus Clique inequalities for bipartite graphs. Consider a partition (M, \bar{M}) of $\{1, \dots, d\}$. First, we multiply the simplex constraint $\sum_{i=1}^d x_i = 1$ by some x_j and add these equation up for all $j \in M$. Then, we subtract the term $x_i x_j$ for $i, j \in M$ from both side and get

$$\sum_{j \in M} \sum_{i \in \bar{M}} x_i x_j = \sum_{j \in M} x_j - \sum_{j \in M} \sum_{i \in M} x_i x_j.$$

Formulating the left hand side with the linearization variables and noting that the right hand side can be overestimated by tangents on the function $g = z - z^2$ yields the following

Theorem 2. *Let f_α be the tangent of $g = z - z^2$ at $\alpha \in [0, 1]$. Then, for any $M \subset \{1, \dots, d\}$, the following inequality is valid for $(x, Y) \in \Gamma$:*

$$\sum_{j \in M} \sum_{i \in \bar{M}} Y_{ij} \leq f_\alpha \left(\sum_{j \in M} x_j \right)$$

We call these inequalities *generalized MSC bipartite inequalities*. They can also be separated by solving a quadratic programming problem which features continuous and binary variables. An extensive computational study shows that the inequalities considerably strengthen the McCormick relaxation and yield a significant improvement in terms of dual bound and time to optimality. On a large test set, these inequalities close over 86 % of the remaining gap of the root relaxation including the application of RLT. With reference to CPLEX version 12.6.3, combining the strengths of projected RLT and both classes of inequalities, we increased the number of instances solved from 119 to 230. The run time as well as the number of branch-and-bound nodes fall by one or two orders of magnitude depending on the problem size. The proposed inequalities can be adapted to sets described by a knapsack constraint with more general coefficients.

3 Strong Relaxations for the Pooling Problem

The pooling problem is a classic non-convex, nonlinear problem introduced by Haverly in 1978 [Hav78], which is, for example, relevant in the petrochemical industry. The task is to route flow from inputs over pools to outputs. The material has attributes whose concentrations are known at inputs, but are constrained at outputs. The challenge is to route the flow through the network such that the quality constraints at the outputs are met. The blending of the material at pools and outputs differentiates the pooling problem from other multi-commodity flow problems and give rise to non-convex quadratic constraints. This part is joint work with Jim Luedtke, Claudia D’Ambrosio, and Jeff Linderoth [Lue+18b; Lue+18a].

We use the state-of-the-art pq -formulation [TS02] as starting point to derive a novel relaxation for the problem. This is done by reducing the network to its minimal essence comprising only an aggregated version of the inputs, one output, one pool and one material attribute and finally to the study of the feasible set for the following constraints:

$$u - xt = 0 \tag{1a}$$

$$y + u \leq 0 \tag{1b}$$

$$z + x \leq C \tag{1c}$$

$$\underline{\beta}z \leq y \leq \bar{\beta}z \tag{1d}$$

$$z \geq 0, x \in [0, C], t \in [\underline{\gamma}, \bar{\gamma}] \tag{1e}$$

All variables are linear aggregations of variables from the pq -formulation and the set is carefully constructed to be a relaxation. Thus valid inequalities for the set are also valid for the pq -formulation. This set is small enough for a thorough study, but still captures parts of the central non-convex structure of the pooling problem in the non-convex constraint (1a).

After finding the relaxation, we follow the route that has been walked by integer linear programmers for decades, namely to study the extreme points of the relaxation and translate the inner description of the set into an outer description based on the defining inequalities. The set of interest is in general not a polyhedron and has infinitely many extreme points. However, a finite subset of the extreme points leads to a parametrized set of linear inequalities that are used to derive new valid linear and convex nonlinear inequalities for the pooling problem. The inequalities are added directly or by means of the separation of gradient inequalities to the pq -formulation. A computational study shows that the proposed relaxation is tight enough to strengthen the McCormick relaxation of the pq -formulation. On a test set of 360 instances the gap of the state-of-the-art pq -relaxation is reduced from 5.5% to 3.0%. Especially on sparse instances, the additional inequalities provide a significant speed-up of the global solution of the problem and allow to solve several instances that are not solved by the pq -formulation.

4 Conclusion

We focused on applications, where nonconvexity in the model formulation pose principal difficulty and improved the solution methods by exploiting structural particularities. Different techniques were used to achieve these results. For RLT we propose a small enhancement for a well-known general technique. For standard quadratic programming, we use a connection to a combinatorial problem to derive cutting planes that were then generalized using a connection to RLT. In contrast, for the pooling problem, the key contribution was to find and study an appropriate relaxation to derive valid inequalities. We conducted extensive computational experiments in order to fathom the computational impact.

Acknowledgements

The author thanks Pierre Bonami, Claudia D’Ambrosio, Jeff Linderoth, Andrea Lodi, Jim Luedtke, and Andrea Tramontani for the joint work.

References

- [Bom98] I. M. Bomze. “On Standard Quadratic Optimization Problems.” In: *Journal of Global Optimization* 13.4 (1998), pages 369–387.
- [Bon+16] P. Bonami, A. Lodi, J. Schweiger, and A. Tramontani. *Solving Standard Quadratic Programming by Cutting Planes*. Technical report DS4DM-2016-001. Polytechnique Montréal, 2016.
- [Hav78] C. A. Haverly. “Studies of the Behavior of the Recursion for the Pooling Problem.” In: *SIGMAP Bulletin* 25 (1978), pages 19–28.
- [Lue+18a] J. Luedtke, C. D’Ambrosio, J. Linderoth, and J. Schweiger. *Strong Convex Nonlinear Relaxations of the Pooling Problem*. Technical report ZIB-Report 18-12. Zuse Institute Berlin, 2018.
- [Lue+18b] J. Luedtke, C. D’Ambrosio, J. Linderoth, and J. Schweiger. *Strong Convex Nonlinear Relaxations of the Pooling Problem: Extreme Points*. Technical report ZIB-Report 18-13. Zuse Institute Berlin, 2018.
- [McC76] G. P. McCormick. “Computability of Global Solutions to Factorable Nonconvex Programs: Part I—Convex Underestimating Problems.” In: *Mathematical Programming* 10 (1976), pages 147–175.
- [MF13] R. Misener and C. A. Floudas. “GloMIQO: Global mixed-integer quadratic optimizer.” In: *Journal of Global Optimization* 57.1 (2013), pages 3–50.
- [MS65] T. S. Motzkin and E. G. Straus. “Maxima for graphs and a new proof of a theorem of Turán.” In: *Canadian Journal of Mathematics* 17 (1965), pages 533–540.
- [SA92] H. D. Sherali and A. R. Alameddine. “A new reformulation-linearization technique for bilinear programming problems.” In: *Journal of Global Optimization* 2.4 (1992), pages 379–410.
- [Sch16] J. Schweiger. “Gas Network Extension Planning for Multiple Demand Scenarios.” In: *Operations Research Proceedings 2014: Selected Papers of the Annual International Conference of the German Operations Research Society (GOR), RWTH Aachen University, Germany, September 2-5, 2014*. Edited by M. Lübbecke, A. Koster, P. Letmathe, R. Madlener, B. Peis, and G. Walther. Cham: Springer International Publishing, 2016, pages 539–544.
- [Sch17] J. Schweiger. “Exploiting structure in nonconvex quadratic optimization and gas network planning under uncertainty.” PhD thesis. Technische Universität Berlin, 2017.
- [TS02] M. Tawarmalani and N. V. Sahinidis. *Convexification and Global Optimization in Continuous and Mixed-Integer Nonlinear Programming: Theory, Algorithms, Software, and Applications*. Boston MA: Kluwer Academic Publishers, 2002.