Mathematical issues in Bilevel Mixed-Integer Linear Programming applied to systems biology.

Guilherme De Olivera Mendes Ohira*a\**, Galo A. C. Le Roux*a*

aEscola Politécnica daUniversidade de São Paulo (USP), Av. Prof. Luciano Gualberto, trav.3, 380, 05508-020, Brazil

guilhermeohira@gmail.com

Abstract: Gene knockouts can be seen as a tool to improve microorganism performance of an industrial objective function. Despite molecular biology own several state-of-the-art techniques to implement genetic modifications, there is a lack of strategies to guide the decision on what target genes knockouts should be performed. It is possible to tackle the problem by Fluxomic techniques and Constraint-Based Models. To secure the biological objective, while trying to enhance bioprocess performance, it is possible to formulate the problem as a Bilevel Mixed-Integer Linear Programming Problem (BMILP). The first algorithm suggested in literature that makes use of the BMILP strategy, is called OptKnock (2003). It applies the BMILP by enforcing optimality, which translates the problem into a Mathematical Program with Complementarity Constraints (MPCC). OptKnock’s authors suggested to apply dual theory to circumvent complementary and keep linearity (MILP). However, the mathematical description never clear if the optimal achieved by the MPCC is in fact the same of the original BMILP. In the present study an algebraic deduction was suggested. A set of *in silico* experiments were performed in order to compare the approaches. The results show that, in spite of the that KKT conditions hold, the industrial optimal achieved by the resulting MILP may not match the original BMILP. It is possible to conclude that the approach suggested by OptKnock authors can reshape the feasible region when compared to the BMILP. Hence, it is possible to obtain values of industrial that wrongly appear to enhance the objective functions, which can mislead the binary search tree (BST), and then produce suboptimal results instead.

**Keywords**: Systems Biology, Fluxomic, Computational Strain Optimization Models (CSOM), Bi-level Programming Problem (BLPP), MPCC.

* 1. Introduction

Microorganisms can be seen as micro-factories since they can convert substrate into products. There are several products that are desired for industrial scale production, but microorganisms are biased to enhance their survivability, usually maximizing biomass production instead. This behaviour can be seen as a competition between biological objectives (i.e., maximizing survivability) and industrial objectives (i.e., human desire).

 To build algorithms that optimize biochemical networks it is necessary to interpret the metabolism. One way to do this is by Fluxomic techniques, such as Flux Balance Analysis (FBA). This approach looks at the metabolic network as a huge collection of chemical reactions that are represented by their stoichiometry. Hence, it is possible to reconstruct biochemical networks that contain known metabolites, reactions and genes in genomic scale (Orth *et al*., 2010). It is possible to build a space that describes everything that is possible in terms of the fluxes inside a microorganism. It is common to explore that space with a bias, which corresponds to the maximization of biomass growth.

 FBA, when seen mathematically, is a Linear Programming problem (LP), which can have degenerate results. In Systems Biology, and when considering genomic scale, a degenerate result is very likely to take place. This means that there are infinite optimal solutions. This result depicts a key feature in metabolism, the notion of silent phenotypes (Price *et al*., 2004).

 Algorithms such as OptKnock (Burgard *et al*., 2003), hold the biological objective function (i.e., the inner problem – the biological objective) while looking for high-throughput industrial desired phenotypes (i.e., outer problem – industrial objective). These algorithms switch off (i.e., knockout) given reactions, in order to enhance industrial objective function.

 OptKnock-like algorithms formulate a Bilevel Mixed-Integer Linear Programming Problem (BMILP). When first described, authors suggested to enforce KKT conditions to describe the BMILP as a Mathematical Problem with Complementarity Constraints (MPCC), and then use lagrangian duality to finally achieve a Mixed-Integer LP (MILP) (Burgard *et al*., 2003).

OptKnock-like algorithms account for a reasonable number of successful strains and patents. But, even after twenty years, it has not been proved that the resulting MILP has the same result as the first suggested BMILP framework. Dempe and Dutta, 2012, showed that when Slater’s Condition Qualification (CQ) (in the outer problem) is fulfilled both Bilevel and MPCC have the same optimal global solution. Despite both levels are LP, Slater CQ may not be achieved in the outer level.

Our main objective is to review the mathematical framework, initially proposed by Burgard *et al*., 2003, to verify possible issues that can be faced in further studies and highlight them. In the present work we show that KKT conditions hold in the inner problem (FBA), but a gap can be achieved when comparing both BMILP and the resulting MILP.

* 1. Methodology
		1. Experimental (in silico) set-up

A process to convert a bilevel into a single level problem is described in Burgard *et al*., 2003. Briefly, the BMILP, shown in Eq. (1), has an inner level (MILP) that can have its KKT conditions enforced.

|  |  |
| --- | --- |
| $$\begin{matrix}\begin{matrix}max\\y,v\end{matrix}&v\_{obj}=c\_{2}^{t}v&&\\&s.t.&\begin{matrix}max\\v\end{matrix}&v\_{bio}=c\_{1}^{t}v\\&&s.t.&Sv=0\\&&&diag\left(y\right)LB\leq v\leq diag\left(y\right)UB\\&&&LB\leq v\leq UB\\&&&v\in R^{n}\\&&&1^{t}y=n-K\\&&&y\in \{0,1\}^{n}\end{matrix}$$ |  |

This results in a MPCC with mixed-integer variables. This problem can be solved by lagrangian duality, which has linear dependency with KKT conditions. Hence, it is possible to replace complementarity with lagrangian strong duality (which is still non-linear for the mixed-integer case). It is possible to handle the current problem applying an additional (redundant) constraint ($LB\leq v\leq UB$). Due to duality theory and based on how the lagrangian duality works (Bazaraa *et al*., 2003) the redundant constraint unfolds in such way that non-linear terms of strong duality can be replaced by a zero, and hence, circumventing the non-linearity (this approach was first suggested by Xu *et al*., 2013).

 Eq. (1) shows the BMILP, where $v$is the optimization column vector of the inner level that represents flux of every reaction (*n*). $S$is the stoichiometric matrix, which has *m* lines and *n* columns, representing every reaction rule in the metabolism. Both $LB$and $UB$stands for *lower bounds* and *upper bounds*, which are imposed limits estimated by Flux Variability Analysis (FVA). FVA is a pre-processing analysis (details described in Maranas and Zomorrodi, 2016). $y$is a binary optimization vector (of the outer level) that can switch-off (i.e., knockout) reactions. Due to limited computational power, number of knockouts are limited by the constraint $1^{t}y=n-K$. In the present study, only 5 knockouts are allowed ($K=5)$. Additionally, there are reactions that cannot be knocked out since the pre-processing analysis showed that they are fundamental to viability: Biomass formation, ATP maintenance and phosphate transport (diffusion) (algorithm described in Machado *et al*., 2016).

A genomic scale model of *Escherichia coli*, strain K-12, substrate MG1655 was used (Feist *et al*., 2007). This model is available in BiGG database (King *et al*., 2016). After FVA pre-processing the model has $n=1532$ reactions and $m=1032$ metabolites. The resulting MILP is similar to the one presented by Xu *et al*., 2013 (I), but we assumed the stoichiometric matrix constraint as an equality one ($Sv=0)$, and our knockout number is fixed ($1^{t}y=n-K$). The final build had 8692 continuous variables and 1532 binary variables, 2671 equalities constraints and 9192 inequalities constraints. A total of 3 industrial objective functions were chosen: Succinate production, Hydrogen production and Threonine (the difference between each industrial objective is due only to $c\_{2}^{t}$).

* + 1. Slater’s CQ verification and degeneracy (outer level)

A LP always fulfil Slater’s CQ if in standard form, so there is no reason to test the resulting MILP. However, it is possible for a BMILP not to fulfil these conditions. To investigate that, the inner optimal was fixed in both $LB\_{bio}$ and $UB\_{bio}$ and FBA-like optimization was solved for each reaction (primal variables) maximizing and minimizing each of them (Eq. (2)).

|  |  |
| --- | --- |
| $$\begin{matrix}\begin{matrix}min/max\\v\end{matrix}&v\_{i}=c^{t}v\\s.t.&Sv=0\\&LB\leq v\leq UB\\&LB\_{bio}\leq v\_{bio}\leq UB\_{bio}\\&v\in R^{n}\end{matrix}$$ |  |

 Once solved, it is possible to know the possible horizon of each reaction, and hence, to know the domain of variation of each reaction in the degenerate region of FBA. Then, if a given $v\_{i}$is at either $LB\_{i}$ or $UB\_{i}$ bound, while there is a gap between $\overbar{v}\_{i}^{min}$ and $\overbar{v}\_{i}^{max}$, then Slater’s CQ is unfulfilled.

 Both results of $\overbar{v}\_{i}^{min}$ and $\overbar{v}\_{i}^{max}$ show how wide a reaction can vary in the degenerate region of FBA. That kind of analysis is very close to FVA, which shares the same objective, but with a slight difference. In this study this kind of analysis is important to verify if the degenerate region is the same for both conditions.

* + - 1. Medium, growing conditions and implementation

 It was assumed that carbon source, oxygen uptake and ATP maintenance ($LB\_{i}$) take the following values: $-10 mmol/gDW/h$ ($gDW$ stands for grams of Dry Weight biomass), $-18.5 mmol/gDW/h$ and $8.39 mmol/gDW/h$, respectively.

 Optimizations were run in a desktop computer with a Ryzen 5 3600X, implemented in Matlab, solved by Gurobi (v. 10) with a fixed processing time of 90 minutes (deterministic). Due to computational limitations, a heuristic algorithm was chosen for the binary search tree while the node solver chosen was Dual Simplex.

* 1. Results and discussion
		1. Algebraic deductions
			1. Inner optimal

When dealing with Bilevel problems it is possible to enforce KKT conditions of the inner level problem and achieve a MPCC. In LP, complementarity can hold in both contexts: elementwise and as vector multiplication. In the last one it is possible to see that strong duality is linear dependent with KKT conditions, so is possible to circumvent the complementarity non-linearity by replacing it with strong duality (first suggested by Burgard *et al*., 2003). Note that complementarity is implicit in the final framework.

 An important concern is that if a numerical violation eventually crosses a line to be negative (either dual variable associated to an inequality constraint or the inequality constraint itself), then the complementarity seen as a vector multiplication is not equivalent to original KKT condition (hence it is not possible to replace it by a strong duality). In such a way, complementarity can be violated, and the inner result be non-optimal.

* + - 1. Gap between MPCC and Bilevel

Theorem 2.3 and Slater’s CQ showed in Dempe and Dutta (2012) study states that if the outer problem fulfils Slater’s CQ then a bilevel can be replaced by an MPCC and achieve global optimal solution (and hence exhibiting no gap between them).

 Slater condition is a sufficient condition for strong duality, which is mandatory to enforce KKT conditions. Slater condition, in LP, is always fulfilled if it is in its standard form. When deducing strong duality, in LP, it is possible to see that complementarity is only the term which stands between weak and strong duality.

 In bilevel LP context, it is possible not to fulfil Slater CQ in the outer problem. For example, consider an inner level problem that a dual variable (associated to an inequality) is positive, and hence inequality is binding in the optimal solution. That particular inequality does not fulfil Slater CQ (regarding outer level problem) since it becomes an equality and not a strict inequality. This behaviour is usual in the context of Systems Biology.

 Furthermore, also usual in Systems Biology, there is a degeneracy in dual problem, meaning that Lagrange multipliers (i.e., dual variables) are non-unique (i.e., degenerate). Note that, in the process of enforcing KKT conditions, dual variables become optimization variables as well. Then, it is possible to glimpse that some of these variables can reshape constraints related to inner level problem. For example, it is possible that a $LB\_{i}\leq v\_{i}\leq UB\_{i}$ variable gets bonded to a lower bound or upper bound if the degeneracy of dual variables enhances the industrial objective function. Note that the original bilevel, Eq. (1), does not consider dual variables as optimization variables, and hence, it is not possible to observe that kind of behavior.

* + - 1. What cannot be seen only using algebra?

In section 3.1.1. there is a motivation to investigate KKT conditions of the inner level. Obviously numerical violations, below a given threshold ($10^{-9}$), is expected, but there is no guarantee that this value is small enough to hold KKT conditions of the inner level. If so, the matter of how small the gap between an optimal inner level solution (bilevel) and a (possible) non-optimal solution (resulting MILP), due unfulfilled complementarity, can only be evaluated in practice.

 Section 3.1.2. outlines possible issues due to the gap between BMILP and the resulting MILP. The magnitude of this gap must be discussed, as the size of the gap determines the magnitude of this (possible) issue. Additionally, a greater value wrongly observed in the Bilevel Problem, can also be translated into a false degeneracy (regarding to outer problem), which can hinder further analysis on phenotypes.

 In order to verify the magnitude of possible gaps and mathematical issues in the resulting MILP (mathematical simplification of a Bilevel), an experimental (*in silico*) set-up was suggested.

* + 1. Experimental (in silico) results
			1. Inner level optimality

In all studied industrial objective functions, there was no significant numerical violation (>$10^{-7}$). Moreover, once knockouts were identified, original bilevel was solved sequentially (inner level was solved, then $c\_{1}^{t}v=v\_{bio}$ becomes a constraint bond into its respective bounds). Comparing both results (resulting MILP and Bilevel) were basically the same. From these results, it is reasonable to assume that inner level optimal was achieved and the “height” (i.e., the optimal result of the inner level) of a feasible region is the same for both resulting MILP and BMILP.

* + - 1. Slater’s CQ and gap

All conditions did not fulfill Slater’s CQ. Succinate and Hydrogen show 2 reactions (common to both) “ATPM” (maintenance energy requirements) and glucose uptake, that could both be translated into equality constraints. Threonine showed 82 reactions that violated Slater’s CQ. All of them at once cannot be translated into equalities without violating predictive feature of Eq. (1).

 Succinate, Hydrogen and Threonine showed low gap between BMILP and the resulting MILP. Succinate was of the order of $10^{-9}$ and both the others were $10^{-6}$, all had a greater result in the resulting MILP.

 It was possible to verify if the FBA’s degenerate region for BMILP and resulting MILP is the same, as was suggested in Eq. (2). Table 1 shows the results.

Table 1: FBA’s degenerate region differences

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Target | # of LB cases | # of UB cases | Order of greatest gap (LB) | Order of greatest gap (UB) |
| Succinate | 2 | 1 | $$10^{-3}$$ | $$10^{-4}$$ |
| Hydrogen | 14 | 25 | $$10^{-2}$$ | $$10^{-5}$$ |
| Threonine | 277 | 1174 | $$10^{-6}$$ | $$10^{-6}$$ |

 The second and third columns of Table 1 stands for the number of cases where LB and UB were different between BMILP and resulting MILP. The last two columns show the order of greatest values (in absolute) of gap between BMILP and MILP.

 It is notable that the gap between objective functions is reasonably smaller than the gap between these two cases of lower bounds shown in Table 1. Succinate’s lower bound greatest gap ($10^{-3}$) is associated with “GLYLC” (Glycine Cleavage System) reaction. While Hydrogen’s lower bound greatest gap ($10^{-4}$) is associated with “EAR160x” a reaction related to Cell Envelope Biosynthesis. Both need coenzymes in its reactions (NAD and NADH).

* 1. Conclusions

Present work highlights a possible gap between a BMILP and resulting MILP in the context of OptKnock-like algorithms. It is possible to find gaps between BMILP and the resulting MILP as well as in FBA’s degenerate region. That kind of problem can produce unrealistic results in nodes while an algorithm searches the binary tree (BST). Because of it is a MILP, lower node solutions usually are not searched, which can mislead the BST if one node shows a high enough unrealistic optimal. Hence, by the end of the BST, a suboptimal result can be achieved in the resulting MILP.

 Additionally, the cause for these gaps remains unclear. Despite all industrial objective functions show violation of Slater’s CQ, Succinate and Hydrogen can be avoided by removing inequalities from “ATMP”, glucose uptake and oxygen uptake, and introducing equivalent equalities instead. On the other hand, there are gaps in FBA’s degenerate region too.

Probably the non-unicity (degeneracy) of the dual variables can reshape the feasible region of the outer level (i.e., FBA’s degenerate region), turning upper or lower feasible bounds tighter.

Acknowledgements

We gratefully acknowledge grant 2023/09130-6, São Paulo Research Foundation (FAPESP), and support of the RCGI – Research Centre for Gas Innovation, hosted by the University of São Paulo (USP) and sponsored by FAPESP (2014/50279-4 and 2020/15230-5) and Shell Brasil, and the strategic importance of the support given by ANP (Brazil’s National Oil, Natural Gas and Biofuels Agency) through the R&D levy regulation.

References

Bazaraa, M. S., Sherali, H. D., & Shetty, C. M. (2013). *Nonlinear programming: theory and algorithms*. John wiley & sons.

Burgard, A. P., Pharkya, P., & Maranas, C. D. (2003). Optknock: a bilevel programming framework for identifying gene knockout strategies for microbial strain optimization. Biotechnology and bioengineering, 84(6), 647-657.

Dempe, S., & Dutta, J. (2012). Is bilevel programming a special case of a mathematical program with complementarity constraints?. Mathematical programming, 131, 37-48.

Feist, A. M., Henry, C. S., Reed, J. L., Krummenacker, M., Joyce, A. R., Karp, P. D., ... & Palsson, B. Ø. (2007). A genome‐scale metabolic reconstruction for Escherichia coli K‐12 MG1655 that accounts for 1260 ORFs and thermodynamic information. Molecular systems biology, 3(1), 121.

Gurobi Optimization, L.. (2023). Gurobi Optimizer Reference Manual

King ZA, Lu JS, Dräger A, Miller PC, Federowicz S, Lerman JA, Ebrahim A, Palsson BO, and Lewis NE. BiGG Models: A platform for integrating, standardizing, and sharing genome-scale models (2016) Nucleic Acids Research 44(D1):D515-D522. doi:10.1093/nar/gkv1049.

Machado, D., Herrgård, M. J., & Rocha, I. (2016). Stoichiometric representation of gene–protein–reaction associations leverages constraint-based analysis from reaction to gene-level phenotype prediction. PLoS computational biology, 12(10), e1005140.

Maranas, C. D., & Zomorrodi, A. R. (2016). Optimization methods in metabolic networks. John Wiley & Sons.

Orth, J. D., Thiele, I., & Palsson, B. Ø. (2010). What is flux balance analysis?. *Nature biotechnology*, *28*(3), 245-248.

Price, N. D., Reed, J. L., & Palsson, B. Ø. (2004). Genome-scale models of microbial cells: evaluating the consequences of constraints. Nature Reviews Microbiology, 2(11), 886-897.

Xu, Z., Zheng, P., Sun, J., & Ma, Y. (2013). ReacKnock: identifying reaction deletion strategies for microbial strain optimization based on genome-scale metabolic network. PloS one, 8(12), e72150.