On the solution of a graph partitioning problem under capacity constraints

Viet Hung Nguyen

LIP6 - University Pierre et Marie Curie Paris 6

joint work with P. Bonami, M. Klein and M. Minoux

seminar of the PGMO, October 2012

Outline of the talk

1. Introduction
2. 0/1 quadratic models for GPCC
3. Solution techniques
4. Numerical results

What we are going to talk about...

1. Introduction
2. 0/1 quadratic models for GPCC
3. Solution techniques
4. Numerical results

Graph Partitioning under Capacity Constraints (GPCC)

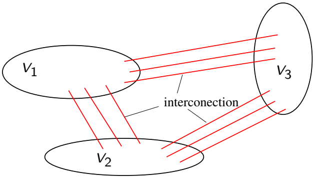
- Given a graph $G = (V, E)$ of n nodes and m edges. The edges are weighted by a vector $t \in \mathbb{R}_+^E$.

Graph Partitioning under Capacity Constraints (GPCC)

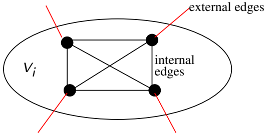
- Given a graph $G = (V, E)$ of n nodes and m edges. The edges are weighted by a vector $t \in \mathbb{R}_+^E$.
- *The Graph Partitioning Problem* to find a partition $\mathcal{P} = \{V_1, V_2, \ldots, V_p\}$ of V minimizing the total weight of the edges in the interconnection. Classical versions impose some constraints on the cardinality of the cluster.

Graph Partitioning under Capacity Constraints (GPCC)

- Given a graph $G = (V, E)$ of n nodes and m edges. The edges are weighted by a vector $t \in \mathbb{R}_+^E$.
- *The Graph Partitioning Problem* to find a partition $\mathcal{P} = \{V_1, V_2, \ldots, V_p\}$ of V minimizing the total weight of the edges in the interconnection. Classical versions impose some constraints on the cardinality of the cluster.
- *Capacity constraints.* The capacity of a cluster V_i in the partition is the sum of its internal edges and external edges. The capacity constraints bound the capacity of every cluster to a constant B.

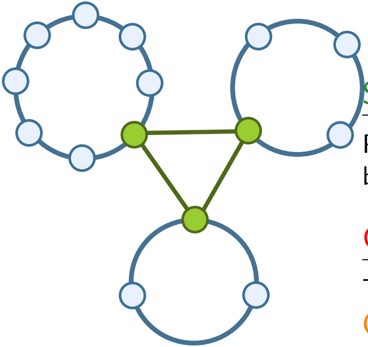


A partition $\mathcal{P} = \{V_1, V_2, V_3\}$.



Capacity of a cluster $V_i \leq B$

Application of GPCC : SONET/SDH network design



Datas

- A set of nodes $V = \{1, \ldots, n\}$ ($n \approx 50$),
- The volume of traffic t_{uv} and distance d_{uv} between every pair u, v of nodes

Solutions

Partition V into rings (*local rings*) connected by a secondary federal ring.

Constraints

The capacity of local ring are bounded.

Objective

A cost l_{uv} aggregated from t_{uv} and d_{uv} is given for each pair u, v of nodes. We aim to minimize the cost of the interconnection.

GPCC as a subproblem

- The problem could be decomposed (in a Benders fashion) into two stages :
  1. Partition nodes into clusters satisfying the capacity constraints.
  2. Find the “cheapest“ TSP tour (ring) over each cluster.

GPCC as a subproblem

- The problem could be decomposed (in a Benders fashion) into two stages :
  1. Partition nodes into clusters satisfying the capacity constraints.
  2. Find the “cheapest“ TSP tour (ring) over each cluster.
- In SONET/SDH norm, the number of nodes in the clusters is around 10. Thus the second stage can be solved easily by current TSP solvers.
- The first stage (master problem) is the most difficult part and it is the GPCC.

GPCC as a subproblem

- The problem could be decomposed (in a Benders fashion) into two stages :
  1. Partition nodes into clusters satisfying the capacity constraints.
  2. Find the “cheapest“ TSP tour (ring) over each cluster.
- In SONET/SDH norm, the number of nodes in the clusters is around 10. Thus the second stage can be solved easily by current TSP solvers.
- The first stage (master problem) is the most difficult part and it is the GPCC.
- Goldschmidt et al. (2003) have shown that GPCC is *NP*-hard even when G is a 3-regular graph.
- We show that even when the number of rings is fixed (k-GPCC), there is no constant approximation algorithm.

Proof of no constant approximability

The idea of the proof is borrowed from [Andrev and Racke 2006] for the balanced graph partitioning problem.

The 3-Partition Problem

- a positive constant M
- $3k$ positive integers $a_1, \ldots, a_{3k}$ where $\frac{M}{4} < a_i < \frac{M}{2}$ $\forall i = 1, \ldots 3k$
- $\sum_{i=1}^{3k} a_i = kM$.

Instance of $\lfloor \frac{n}{3p} \rfloor$-GPCC

- Each integer a_i corresponds to a clique of p vertices.
- Each edge in the cliques is of capacity $\frac{2a_i}{p(p-1)}$ so that the total capacity of the clique is a_i.

Proof of no constant approximability

The idea of the proof is borrowed from [Andrev and Racke 2006] for the balanced graph partitioning problem.

The 3-Partition Problem

- a positive constant M
- $3k$ positive integers $a_1, \ldots, a_{3k}$ where $\frac{M}{4} < a_i < \frac{M}{2}$ $\forall i = 1, \ldots 3k$
- $\sum_{i=1}^{3k} a_i = kM$.

Instance of $\lfloor \frac{n}{3p} \rfloor$-GPCC

- Each integer a_i corresponds to a clique of p vertices.
- Each edge in the cliques is of capacity $\frac{2a_i}{p(p-1)}$ so that the total capacity of the clique is a_i.

An approximation algorithm with finite approximation factor should output a solution of zero objective for k-GPCC and hence solve the 3-Partition problem when the latter is feasible.

What we are going to talk about...

1. Introduction
2. 0/1 quadratic models for GPCC
3. Solution techniques
4. Numerical results

Node-Cluster model [Goldschmidt et al. 2003]

The variables

For every node u and n possible clusters $i = 1, \ldots, n$,

$$x_{ui} = \begin{cases} 1, & \text{if the node } u \text{ is assigned to cluster } i \\ 0, & \text{otherwise} \end{cases}$$

Objective

Maximizing the internal traffics in clusters,

$$\max \sum_{i=1}^{n} \sum_{u=1}^{n-1} \sum_{v=u+1}^{n} t_{uv} x_{ui} x_{vi}$$

Note that the objective is quadratic

Node-Cluster model : The constraints

The capacity constraints

For all clusters $i = 1, \ldots, n$,

$$\sum_{u=1}^{n} x_{ui} W_u - \sum_{u=1}^{n-1} \sum_{v=u+1}^{n} x_{ui} x_{vi} t_{uv} \leq B$$

W_u is the total traffic incident to the node u.

The assignment constraints

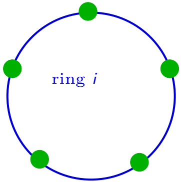
Node-Cluster model : The constraints

The capacity constraints

For all clusters $i = 1, \ldots, n$,

$$\sum_{u=1}^{n} x_{ui} W_u - \sum_{u=1}^{n-1} \sum_{v=u+1}^{n} x_{ui} x_{vi} t_{uv} \leq B$$

W_u is the total traffic incident to the node u.



The assignment constraints

For every node u,

$$\sum_{i=1}^{n} x_{ui} = 1$$

Drawback

High degree of symmetry : the same solution for GPCC may correspond to several distinct solutions for the model.

Node-Cluster model : The constraints

The capacity constraints

For all clusters $i = 1, \ldots, n$,

$$\sum_{u=1}^{n} x_{ui} W_u - \sum_{u=1}^{n-1} \sum_{v=u+1}^{n} x_{ui} x_{vi} t_{uv} \leq B$$

W_u is the total traffic incident to the node u.

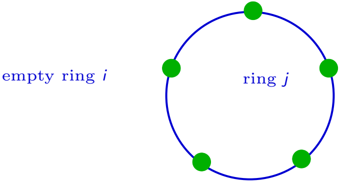
The assignment constraints

For every node u,

$$\sum_{i=1}^{n} x_{ui} = 1$$

Drawback

High degree of symmetry : the same solution for GPCC may correspond to several distinct solutions for the model.



Node-Cluster model : Breaking the symmetry

We propose two constraints which oblige that a solution for GPCC correspond to an unique solution for the model.

The order constraints

For every indices $u, i = 1 \ldots, n$,

$$x_{ui} \leq x_{ii} \;\; u = i + 1, \ldots, n$$

$$x_{ui} = 0 \;\; u = 1, \ldots, i - 1$$

Summary

The Node-Cluster model has n^2 variables, n quadratic constraints, $O(n^2)$ linear constraints and the objective is quadratic.

Node-Node model

This formulation for GPCC is quite similar to the model of Grötschel et Wakabayashi (89) for the clique partitioning problem.

The variables

Node-Node model

This formulation for GPCC is quite similar to the model of Grötschel et Wakabayashi (89) for the clique partitioning problem.

The variables

For every pairs of nodes u and v,

$$x_{uv} = \begin{cases} 0, & \text{if } u, v \text{ are assigned to the same cluster} \\ 1, & \text{otherwise} \end{cases}$$

Objective

Minimizing the traffics in the interconnection,

$$\min \sum_{e \in \mathcal{P}} t_e x_e$$

Note that the objective is linear

Node-Node model : Constraints

The transitivity (or triangle) constraints

For all triplets of nodes u, v and w,

$$x_{(u,v)} + x_{(u,w)} \geq x_{(v,w)}$$

The capacity constraints

For each node u, the capacity of the cluster containing u is bounded by B,

$$\sum_{(v,w)} t_{vw} - \sum_{(v,w)} t_{vw} x_{uv} x_{uw} \leq B$$

Summary

The Node-Node model has n^2 variables, n quadratic constraints, $O(n^3)$ linear constraints and the objective is linear.

Summary on the models

Improved Node-Cluster model

$$\max \sum_{i=1}^{n} \sum_{u=1}^{n-1} \sum_{v=u+1}^{n} t_{uv} x_{ui} x_{vi}$$

$$\sum_{i=1}^{n} x_{ui} = 1 \ \forall \text{ nodes } u, \text{ clusters } i$$

$$x_{ui} \leq x_{ii} \ \ u = i+1, \ldots, n$$

$$x_{ui} = 0 \ \ u = 1, \ldots, i-1$$

$$\sum_{u=1}^{n} x_{ui} W_u - \sum_{u=1}^{n-1} \sum_{v=u+1}^{n} x_{ui} x_{vi} t_{uv} \leq B$$

n^2 variables, n quadratic constraints, $O(n^2)$ linear constraints and quadratically objective.

Node-Node model

$$\min \sum_{e \in E} l_e x_e$$

$$x_{uv} + x_{uw} \geq x_{vw}$$

$$(u, v, w) \in \mathcal{T}$$

$$\sum_{(v,w)} t_{vw} - \sum_{(v,w)} t_{vw} x_{uv} x_{uw} \leq B$$

$$x_e \in \{0, 1\} \ e \in E.$$

n^2 variables, n quadratic constraints, $O(n^3)$ linear constraints and linear objective.

What we are going to talk about...

1. Introduction
2. 0/1 quadratic models for GPCC
3. Solution techniques
4. Numerical results

General unified 0/1 quadratically constrained program

The Node-Cluster and Node-Node models can be cast into the following standard form :

$$\min \; c^T x \text{ (linear objective)}$$
$$Ax \leq b, \text{ (linear constraints)}$$
$$\sum_{i=1}^{p} \sum_{j=i+1}^{p} q_{ij}^k x_i x_j + {d^k}^T x \leq q^k, \text{ (quadratic constraints)}$$
$$k = 1, \ldots, m,$$
$$x \in \{0, 1\}^p$$

For the node-cluster formulations one needs to add an extra variable to put the objective function in the constraints.

Solution techniques for 0/1 quadratically constrained programs (I)

The quadratic capacity constraints are not convex as the traffic matrix is a priori not positive semidefinite. Several known techniques could be used to convexify these constraints, e.g. :

Minimum eigenvalue technique :

By augmenting the diagonal of the matrix representing the quadratic constraints so that

- its smallest eigen values becomes 0,
- and without changing the set of feasible 0/1 solutions.

Implemented in CPLEX, reveals not efficient for solving our models for GPCC.

Solution techniques for 0/1 quadratically constrained programs (II)

Classical linearization (CL) :

Introducing variable y_{ij} to represent each product $x_i x_j$:

$$
\begin{aligned}
\min\ c^T x & \\
Ax \le b, & \\
\sum_{i=1}^{p}\sum_{j=i+1}^{p} q_{ij}^k y_{ij} + {d^k}^T x \le q^k, & \qquad k = 1, \dots, m, \\
\max\{0, x_i + x_j - 1\} \le y_{ij} \le \min\{x_i, x_j\}, & \qquad 1 \le i < j \le n, \\
(x, y) \in \{0,1\}^p \times \mathbb{R}^{\frac{p(p-1)}{2}}, &
\end{aligned}
$$

Projection Technique : An alternative to Classical Linearization

- A drawback of classical linearization is the number of the added variables y's may be at worst the square of the number of the original variables x's . In our case, it raises this number from n^2 to n^3. For $n \approx 50$, the number of variables grows from 2500 to 125000.

Projection Technique : An alternative to Classical Linearization

- A drawback of classical linearization is the number of the added variables y's may be at worst the square of the number of the original variables x's . In our case, it raises this number from n^2 to n^3. For $n \approx 50$, the number of variables grows from 2500 to 125000.
- Saxena et al. (2011) propose a technique which (in a simplified way)
  - dertermine if for a point x^*, there exists a y^* for that (x^*, y^*) feasible for the linear relaxation of the classical linearization.
  - if such y^* does not exist, it generates valid inequalities violated by x^*.

For that, it consists in solving a linear program.

Projection Technique : An alternative to Classical Linearization

- A drawback of classical linearization is the number of the added variables y's may be at worst the square of the number of the original variables x's . In our case, it raises this number from n^2 to n^3. For $n \approx 50$, the number of variables grows from 2500 to 125000.
- Saxena et al. (2011) propose a technique which (in a simplified way)
  - dertermine if for a point x^*, there exists a y^* for that (x^*, y^*) feasible for the linear relaxation of the classical linearization.
  - if such y^* does not exist, it generates valid inequalities violated by x^*.

  For that, it consists in solving a linear program.
- We will see that this technique adapted to presented models will be particularly simple, i.e. not have to solve linear program, cuts generation can be done combinatorially in linear time.

Solution techniques for 0/1 quadratically constrained programs (III)

Projection technique :

It is based on the fact that if $x_i, x_j \in \{0,1\}$,
$y_{ij} = x_i x_j = \min\{x_i, x_j\} = \max\{x_i + x_j - 1, 0\}$. Hence the previous classical linearized program can be rewritten as :

$$\begin{aligned} &\min \ c^T x \\ &Ax \leq b, \\ &\phi^k(x) \leq q^k, k = 1, \dots, m, \\ &x \in \{0,1\}^p. \end{aligned}$$

where $\phi^k(x) := \sum_{\substack{i=1,\dots,p \\ j=i,\dots,p \\ q^k_{ij}<0}} q^k_{ij} \min\{x_i, x_j\} + \sum_{\substack{i=1,\dots,p \\ j=i,\dots,p \\ q^k_{ij}>0}} q^k_{ij} \max\{x_i + x_j - 1, 0\} + {d^k}^T x.$

$\phi^k(x)$ is a convex piecewise linear function with at most $2^{\frac{p(p-1)}{2}}$ pieces.

Classical Linearization (CL)

$$
\begin{aligned}
\min \;\; & c^T x \\
& Ax \leq b, \\
& \sum_{i=1}^{p} \sum_{j=i+1}^{p} q_{ij}^k y_{ij} + {d^k}^T x \leq q^k \\
& k = 1, \ldots, m, \\
\max\{0, x_i + x_j - 1\} \leq y_{ij} \leq \min\{x_i, x_j\}, & \\
& 1 \leq i < j \leq n, \\
& (x, y) \in \{0,1\}^p \times \mathbb{R}^{\frac{p(p-1)}{2}},
\end{aligned}
$$

$\Leftrightarrow$

$\Leftrightarrow$

Classical Linearization (CL)

$$
\begin{aligned}
\min\ & c^T x \\
Ax &\leq b, \\
\sum_{i=1}^{p} \sum_{j=i+1}^{p} q_{ij}^k y_{ij} + {d^k}^T x &\leq q^k \\
k &= 1, \ldots, m, \\
\max\{0, x_i + x_j - 1\} \leq y_{ij} &\leq \min\{x_i, x_j\}, \\
1 \leq i < j &\leq n, \\
(x, y) \in \{0,1\}^p \times \mathbb{R}^{\frac{p(p-1)}{2}}&,
\end{aligned}
$$

$\Leftrightarrow$

$\Leftrightarrow$

Projection Technique (PT) :

$$
\begin{aligned}
&\min\ c^T x \\
&Ax \leq b, \\
&\phi^k(x) \leq q^k, k = 1, \ldots, m, \\
&x \in \{0,1\}^p.
\end{aligned}
$$

Classical Linearization (CL)

$$
\begin{aligned}
\min\ & c^T x \\
Ax &\le b, \\
\sum_{i=1}^{p}\sum_{j=i+1}^{p} q_{ij}^k y_{ij} + {d^k}^T x &\le q^k \\
k &= 1,\dots,m, \\
\max\{0, x_i + x_j - 1\} \le y_{ij} &\le \min\{x_i, x_j\}, \\
1 \le i < j &\le n, \\
(x,y) &\in \{0,1\}^p \times \mathbb{R}^{\frac{p(p-1)}{2}},
\end{aligned}
$$

$\Leftrightarrow$

Projection Technique (PT) :

$$
\begin{aligned}
\min\ & c^T x \\
Ax &\le b, \\
\phi^k(x) &\le q^k, k = 1,\dots,m, \\
x &\in \{0,1\}^p.
\end{aligned}
$$

$$\Updownarrow$$

$\Leftrightarrow$

$$
\begin{aligned}
\min\ & c^T x \\
Ax &\le b, \\
B^k(x) &\le q^k \mathbf{1}, k = 1,\dots,m \\
x &\in \{0,1\}^p.
\end{aligned}
$$

where B^k is the matrix of the $2^{\frac{p(p-1)}{2}}$ linear constraints defining $\phi^k(x)$.

Linear Relaxation of CL (RCL)

$$\min \; c^T x$$
$$Ax \leq b,$$
$$\sum_{i=1}^{p} \sum_{j=i+1}^{p} q_{ij}^k y_{ij} + {d^k}^T x \leq q^k$$
$$k = 1, \ldots, m,$$
$$\max\{0, x_i + x_j - 1\} \leq y_{ij} \leq \min\{x_i, x_j\},$$
$$1 \leq i < j \leq n,$$
$$(x, y) \in [0,1]^p \times \mathbb{R}^{\frac{p(p-1)}{2}},$$

?

Linear Relaxation of PT (RPT) :

$$\min \; c^T x$$
$$Ax \leq b,$$
$$B^k(x) \leq q^k \mathbf{1}, k = 1, \ldots, m$$
$$x \in [0,1]^p.$$

where B^k is the matrix of the $2^{\frac{p(p-1)}{2}}$ linear constraints defining $\phi^k(x)$.

Linear Relaxation of CL (RCL)

$$
\begin{aligned}
\min\ & c^T x \\
Ax &\le b, \\
\sum_{i=1}^{p}\sum_{j=i+1}^{p} q_{ij}^k y_{ij} + {d^k}^T x &\le q^k \\
k &= 1,\dots,m, \\
\max\{0, x_i + x_j - 1\} \le y_{ij} &\le \min\{x_i, x_j\}, \\
1 \le i < j &\le n, \\
(x,y) &\in [0,1]^p \times \mathbb{R}^{\frac{p(p-1)}{2}},
\end{aligned}
$$

$\Leftrightarrow$

Linear Relaxation of PT (RPT) :

$$
\begin{aligned}
\min\ & c^T x \\
Ax &\le b, \\
B^k(x) &\le q^k \mathbf{1}, k = 1,\dots,m \\
x &\in [0,1]^p.
\end{aligned}
$$

where B^k is the matrix of the $2^{\frac{p(p-1)}{2}}$ linear constraints defining $\phi^k(x)$.

Theorem

Suppose that for $k = 1,\dots,m$, q_{ij}^k is either non-negative or non-positive for all $1 \le i < j \le p$, then $RPT = \mathrm{proj}_x(RCL)$.

Summary on the models and their convex reformulations

- Two models : Improved Node-Cluster, Node-Node.
- Three methods of convexification/linearization : minimum eigenvalue technique (default algorithm in CPLEX), CL and PT.

Comparing the strength of the continuous relaxations

In this first experiment, we only seek to compute the continuous relaxations of the classical linearization of the two models. For Node-Node model, we also implement the projection technique.

	Improved N/C		N/N		proj. N/N	
n	CPU	value	CPU	value	CPU	value
16	0.05	1.13E+5	1.24	2.26E+6	0.53	2.26E+6
17	0.05	1.93E+5	1.98	3.33E+6	0.53	3.33E+6
18	0.06	1.93E+5	2.72	1.14E+6	1.10	1.14E+6
19	0.06	2.07E+5	3.90	1.88E+6	1.58	1.88E+6
20	0.06	2.26E+5	5.40	3.90E+6	2.61	3.90E+6
21	0.07	3.24E+5	7.25	5.52E+6	3.01	5.52E+6
22	0.07	4.67E+5	9.11	7.02E+6	3.84	7.02E+6
25	0.07	3.87E+5	29.3	6.13E+6	8.37	6.13E+6
all	0.06	2.64E+5	7.62	3.90E+6	2.69	3.90E+6

Three methods on Node-Node model

We compare the three methods of convexification/linearization for Node-Node models included in a branch-and-bound/cut algorithm for GPCC.

		quad. N/N			N/N			proj. N/N		
n	#in	#sol	CPU	Nodes	#sol	CPU	Nodes	#sol	CPU	Nodes
16	6	2	7979	75233	6	243	3178	6	52	6060
17	6	1	9011	244233	6	412	5827	6	82	6137
18	6	1	8919	45945	6	594	2981	6	117	10993
19	6	2	7617	26129	6	4885	22044	6	900	66267
20	5	0	10800	29919	5	1142	2454	5	336	4542
21	4	0	10800	41544	3	5083	6176	4	1388	5047
22	2	0	10800	42341	1	8683	6584	2	6175	128823
25	2	0	10800	16921	1	9254	1342	2	2869	23982
all	37	6	9585	65283	34	3787	6323	37	1490	31481

Exact Solutions : Improved Node-Cluster model vs Node-Node model

		Improved N/C			proj. N/N		
n	#inst	# sol	CPU	Nodes	# sol	CPU	Nodes
16	6	6	25	6550	6	52	6060
17	6	6	88	21910	6	82	6137
18	6	6	46	6188	6	117	10993
19	6	6	147	20385	6	900	66267
20	5	5	459	44284	5	336	4542
21	4	4	434	29571	4	1388	5047
22	2	2	2467	161504	2	6175	128823
25	2	2	498	13071	2	2869	23982
all	37	37	520	37933	37	1490	31481

TABLE: Statistics on complete solution by branch-and-cut.

Conclusions

- We have studied several models and exact solutions for GPCC.
  - An improvement for Node-Cluster model of Goldschmidt et al.
  - A new Node-Node model for GPCC which have the root in the work of Grötschel and Wakabayashi for the clique partitioning problem.
  - An adaptation particularly simple of the projection technique from Saxena et al. for the presented models.

Conclusions

- We have studied several models and exact solutions for GPCC.
  - An improvement for Node-Cluster model of Goldschmidt et al.
  - A new Node-Node model for GPCC which have the root in the work of Grötschel and Wakabayashi for the clique partitioning problem.
  - An adaptation particularly simple of the projection technique from Saxena et al. for the presented models.
- Numerical results show the advantage of the projection technique over the minimum eigenvalue technique and classical linearization on the Node-Node model. The latter though stronger is less efficient than the Improved Node-Cluster model in exact solutions by branch-and-cut.

Conclusions

- We have studied several models and exact solutions for GPCC.
    - An improvement for Node-Cluster model of Goldschmidt et al.
    - A new Node-Node model for GPCC which have the root in the work of Grötschel and Wakabayashi for the clique partitioning problem.
    - An adaptation particularly simple of the projection technique from Saxena et al. for the presented models.
- Numerical results show the advantage of the projection technique over the minimum eigenvalue technique and classical linearization on the Node-Node model. The latter though stronger is less efficient than the Improved Node-Cluster model in exact solutions by branch-and-cut.
- Future works could be tuning the projection technique for Improved Node-Cluster model and extending this technique to a stronger relaxation than classical linearization.