II.4 Other Dual-Ascent Approaches

The Bundle approach of the two previous Chapters is only one possible Dual-Ascent method based on the theory developed in Part I: barring the (LP)-based versions, that will be (hopefully) tested in a near future, several other possibilities can be considered. In this Chapter, preliminary theoretical and computational results about three alternative approaches are described: although none of them have yet reached the status of a mature technology, they offer interesting directions for future research. In particular, §II.4.1 and §II.4.2 are devoted to the development of ε-Steepest Ascent approaches to the Cost Decomposition and Potential Decomposition of (MMCF) respectively, and in the first one preliminary computational results are presented: then, in §II.4.3 some ideas are presented about a more "combinatorial" approach to Potential Decomposition, that might be extended to other forms of network-structured problems such as the Min-Cost Flow problem on Hypergraphs.

II.4.1 An ε-Steepest Ascent Cost Decomposition approach

One of the basic hypotheses in the development of Bundle methods is that the $\varphi()$ is available only as a "black box", i.e. that the whole subdifferential $\partial\varphi()$ is not known: this is not actually true in our case, since Theorem I.1.2.1 implies that

$$\partial\varphi() = \left\{ \Sigma_{h \in \mathcal{K}} \boldsymbol{x}^h - \boldsymbol{u} : \boldsymbol{x}^h \in X^h(\boldsymbol{y}) \ \forall\ h \in \mathcal{K} \right\}$$

where $X^h(\boldsymbol{y})$ is the set of all optimal solutions of the h-th (MCF) subproblem

$$(\mathrm{RMF}_{\boldsymbol{y}}^h) \qquad \min \left\{ (\boldsymbol{c}^h + \boldsymbol{y})\boldsymbol{x}^h : \boldsymbol{E}\boldsymbol{x}^h = \boldsymbol{b}^h,\ \mathbf{0} \le \boldsymbol{x}^h \le \boldsymbol{u}^h \right\}$$

and it has a nice description in terms of the original problem. In fact, let $\pi^h = \pi^h(\boldsymbol{y})$ be *any* vector of *optimal potentials* for $(\mathrm{RMF}_{\boldsymbol{y}}^h)$ and

$$\bar{c}_{ij}^h = \bar{c}_{ij}^h(\pi^h) = c_{ij}^h - \pi_i^h + \pi_j^h$$

be the corresponding *reduced costs*: some general duality results (Cf. Sec. II.1.4.2) show that $\bar{\boldsymbol{x}}^h \in X^h(\boldsymbol{y})$ if and only if it is feasible, i.e. $\boldsymbol{E}\bar{\boldsymbol{x}}^h = \boldsymbol{b}^h$ and $\mathbf{0} \le \bar{\boldsymbol{x}}^h \le \boldsymbol{u}^h$, and it satisfies the *Complementary Slackness Conditions* (*CSC*)

$$\bar{c}_{ij}^h < 0 \quad ((i,j) \text{ is } \textit{active}) \quad \Rightarrow \quad x_{ij}^h = u_{ij}^h$$

$$\bar{c}_{ij}^h > 0 \quad ((i,j) \text{ is } \textit{inactive}) \quad \Rightarrow \quad x_{ij}^h = 0.$$

Hence, $X^h(\boldsymbol{y})$ is the set of all the flows in which the active and inactive arcs are *blocked* at their

upper and lower bounds respectively, while the flow on *balanced* arcs (s.t. $\bar{c}_{ij}^h = 0$) is *free* to have any value in $[0, u_{ij}^h]$, provided that the flow conservation equations are respected: then, the *steepest ascent direction* finding problem

$$(\mathrm{SAD}_{\boldsymbol{y}}) \qquad \operatorname{argmin}_{\boldsymbol{g}} \left\{ \left\| \boldsymbol{g} \right\|_2^2 : \boldsymbol{g} \in \partial\varphi(\boldsymbol{y}) \right\} =$$
$$\min_{\boldsymbol{x}} \left\{ \left\| \sum_{h \in \mathcal{K}} \boldsymbol{x}^h - \boldsymbol{u} \right\|_2^2 : \boldsymbol{x}^h \in X^h(\boldsymbol{y}) \; \forall \, h \in \mathcal{K} \right\}$$

is a Quadratic single-commodity Min-Cost Flow with separable feasible region but nonseparable objective function, that can be used in a *steepest ascent* approach for maximizing $\varphi()$. Actually, $(\mathrm{SAD}_{\boldsymbol{y}})$ is not enough, since the *constrained maximization* of $\varphi()$ is required: however, with the usual device of introducing the (concave) *characteristic function* of the nonnegative orthant

$$\varphi^+(\boldsymbol{y}) = \begin{cases} 0 & \text{if } \boldsymbol{y} \geq \boldsymbol{0} \\ -\infty & \text{otherwise} \end{cases}$$

the Lagrangean Dual of (MMCF) can be restated as the unconstrained maximization problem

$$(\mathrm{DMMCF}) \qquad \max\{ \varphi(\boldsymbol{y}) + \varphi^+(\boldsymbol{y}) : \boldsymbol{y} \in \mathcal{R}^m \}$$

and the issue becomes that of the characterisation of $\partial[\varphi + \varphi^+](\boldsymbol{y})$; this is the approach that leads to the general definition of *conditional subdifferential* [LPS96], but a direct derivation is preferable here. For any fixed $\boldsymbol{y} \geq \boldsymbol{0}$, $\partial\varphi^+(\boldsymbol{y})$ is the Cartesian product $V(y_1) \times V(y_2) \times \ldots \times V(y_m)$, where

$$V(y) = \begin{cases} \{ 0 \} & \text{if } y > 0 \\ \mathcal{R}_+ & \text{otherwise} \end{cases}$$

so that, since $\partial[\varphi + \varphi^+](\boldsymbol{y}) = \partial\varphi(\boldsymbol{y}) + \partial\varphi^+(\boldsymbol{y})$, the *constrained* steepest ascent direction finding problem is

$$(\mathrm{SAD}_{\boldsymbol{y}}') \qquad \min_{\boldsymbol{x},\boldsymbol{v}} \left\{ \left\| \sum_{h \in \mathcal{K}} \boldsymbol{x}^h - \boldsymbol{u} + \boldsymbol{v} \right\|_2^2 : \boldsymbol{v} \in \partial\varphi^+(\boldsymbol{y}) \, , \boldsymbol{x}^h \in X^h(\boldsymbol{y}) \; \forall \, h \in \mathcal{K} \right\}.$$

This can be put in a more natural form by defining the projection operator "modulo y"

$$[\, w \,]_+^y = \begin{cases} 0 & \text{if } w < 0 \text{ and } y = 0 \\ w & \text{otherwise} \end{cases}$$

and its obvious componentwise extension to vectors of $\mathcal{R}^m$: then,

$$(\mathrm{SAD}_{\boldsymbol{y}}') \qquad \min_{\boldsymbol{x}} \left\{ \left\| [\, \sum_{h \in \mathcal{K}} \boldsymbol{x}^h - \boldsymbol{u} \,]_+^{\boldsymbol{y}} \right\|_2^2 : \boldsymbol{x}^h \in X^h(\boldsymbol{y}) \; \forall \, h \in \mathcal{K} \right\}$$

since any negative component of $\sum_{h \in \mathcal{K}} \boldsymbol{x}^h - \boldsymbol{u}$ corresponding to a zero component of $\boldsymbol{y}$ can be turned to zero by suitably choosing the vector $\boldsymbol{v} \in \partial\varphi^+(\boldsymbol{y})$, while positive components are not modified; hence, $(\mathrm{SAD}_{\boldsymbol{y}}')$ requires finding the least-norm *projected* subgradient w.r.t. the *active* constraints. In this form, the strong similarities with the Augmented Lagrangean problem $(\mathrm{AL}_{\boldsymbol{y},t})$ (Cf. Sec. II.1.4.2) are apparent, as well as the differences:

- the costs of (MMCF) have disappeared, followed by the penalty parameter t, because the sets $X^h(\boldsymbol{y})$ contain all equivalent solutions for $(\mathrm{RMF}_{\boldsymbol{y}}^h)$;
- hence, the costs now induce "hard" constraints on the feasible region;
- the projection operator now takes into account the current point $\boldsymbol{y}$.

Intuitively, (SAD_y') should be "easier" than ($AL_{y,t}$), since it has a very nicely conditioned quadratic objective function and a smaller feasible region: in fact, if the optimal solution of (RMF_y^h) is *unique*, the $\boldsymbol{x}^h$ variables are fixed and disappear from the problem, and this is true for *all* $h \in \mathcal{K}$ (i.e. $\varphi()$ is differentiable) with probability 1 in a random point $\boldsymbol{y} \in \mathcal{R}^m$ (unfortunately, this is very unlikely to be true in the optimal point $\boldsymbol{y}^*$).

A "pure" steepest-ascent approach is prone to the typical problems of first-order methods, i.e. zig-zagging and slow convergence: a possible way to avoid them is to use *steepest* ε-*ascent* directions instead, i.e. work with the ε-*subdifferential* (Cf. §I.1.1) of $[\varphi + \varphi^+](\boldsymbol{y})$. It is well-known (see e.g. [Be91]) that every $\bar{\boldsymbol{x}}^h \in X_\varepsilon^h(\boldsymbol{y})$, i.e feasible ($\boldsymbol{E}\bar{\boldsymbol{x}}^h = \boldsymbol{b}^h$ and $\boldsymbol{0} \le \bar{\boldsymbol{x}}^h \le \boldsymbol{u}^h$) and satisfying the ε-*Complementary Slackness Conditions* (ε-*CSC*)

$$\bar{c}_{ij}^h < \varepsilon \quad ((i,j) \text{ is } \varepsilon\text{-}active) \;\Rightarrow\; x_{ij}^h = u_{ij}^h$$

$$\bar{c}_{ij} > \varepsilon \quad ((i,j) \text{ is } \varepsilon\text{-}inactive) \;\Rightarrow\; x_{ij}^h = 0$$

is ($n\varepsilon$)-*optimal* for (RMF_y^h), so that

$$\partial_{kn\varepsilon}\varphi() = \left\{ \boldsymbol{w} = \Sigma_{h \in \mathcal{K}} \boldsymbol{x}^h - \boldsymbol{u} : \boldsymbol{x}^h \in X_\varepsilon^h(\boldsymbol{y}) \;\forall\, h \in \mathcal{K} \right\};$$

the ($m\delta$)-*subdifferential* of $\varphi^+()$ is the Cartesian product $V_\delta(y_1) \times V_\delta(y_2) \times \ldots \times V_\delta(y_m)$ (see Figure 13) for

$$V_\delta(y) = \left[\, 0 \;,\; \frac{\delta}{y} \,\right] \qquad \text{with the customary assumption that } \delta/0 = +\infty \text{ for any } \delta \ge 0.$$

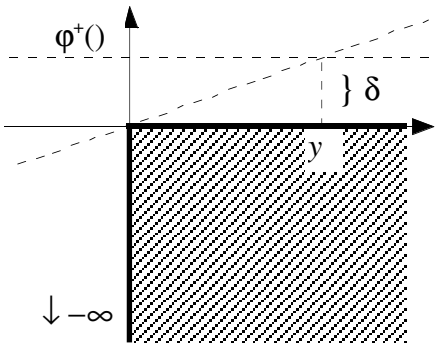

Figure 13: two "extreme" δ-subgradients of $\varphi^+()$ in y

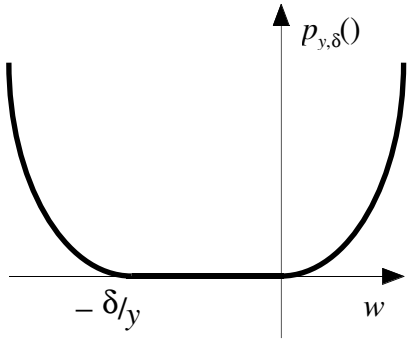

Figure 14: an interpretation of ($\varepsilon\delta$-SAD_y')

Hence, the ($kn\varepsilon + m\delta$)-*steepest* ascent direction finding problem is

($\varepsilon\delta$-SAD_y') $\quad \min_{\boldsymbol{x}} \left\{ \left\| \left[\Sigma_{h \in \mathcal{K}} \boldsymbol{x}^h - \boldsymbol{u} \right]_+^{y,\delta} \right\|_2^2 : \boldsymbol{x}^h \in X_\varepsilon^h(\boldsymbol{y}) \;\forall\, h \in \mathcal{K} \right\}$

having a larger feasible region (the *free* arcs are now the ε-*balanced* ones having $\left| \bar{c}_{ij}^h \right| < \varepsilon$) and a different "projection" operator

$$[w]_+^{y,\delta} = \begin{cases} \min\{ w + \delta/y \,,\, 0 \} & \text{if } w < 0 \\ w & \text{otherwise} \end{cases};$$

the function $p_{y,\delta}(w) = ([w]_+^{y,\delta})^2$ depicted in Figure 14 can be thought as a different form of nonquadratic smooth penalty function, and different ε^h and δ_{ij} could be used for each different commodity and mutual capacity constraint. This leads to the definition of a scaling approach, where the problem is solved for a pair (ε, δ) until the termination criterion $z(\varepsilon\delta\text{-}SAD_y') = 0$ signals that at

least one of the two has to be decreased: a nice feature of such a scheme is that, upon termination of a "scaling phase", an ($kn\varepsilon + m\delta$)-*optimal* solution of (MMCF) is available, since the optimal solution $\bar{\boldsymbol{x}}$ of ($\varepsilon\delta$-SAD$'_{\boldsymbol{y}}$) must clearly satisfy

$$-\delta/y_{ij} \leq \Sigma_{h \in \mathcal{K}} \bar{x}^h_{ij} \leq u_{ij} \qquad \forall\ (i, j) \in \mathcal{A}$$

and hence it is *feasible*.

($\varepsilon\delta$-SAD$'_{\beta,\boldsymbol{y}}$) can be solved with a Simplicial Decomposition [VHo77] approach, where vertices of the sets $X^h_\varepsilon(\ \boldsymbol{y}\ )$ are efficiently generated by exploiting the (MCF) structure as in the Bundle and Linear-Quadratic Penalty function methods [PZ92] [PZ94]: the Master problem of the Simplicial Decomposition is

($\varepsilon\delta$-SAD$'_{\beta,\boldsymbol{y}}$) $\quad \min_{\theta,\mu} \left\{ \tfrac{1}{2} \left\| \boldsymbol{G}_\beta \theta + \mu \right\|^2_2 : \theta \in \Theta\ ,\ \delta \boldsymbol{Y}^{-1} \boldsymbol{e} \geq \mu \geq \boldsymbol{0} \right\}$

where $\boldsymbol{Y}$ is the diagonal matrix having $\boldsymbol{y}$ as the diagonal elements, and $\boldsymbol{G}_\beta$ is the matrix of the subgradients $\boldsymbol{g}_i = \Sigma_{h \in \mathcal{K}} \boldsymbol{x}^h_i - \boldsymbol{u}$ corresponding to the vertices $\boldsymbol{x}^h_i \in X^h_\varepsilon(\ \boldsymbol{y}\ )$ found so far, and kept in the "local Bundle" β. For the special case $\delta = 0$, the upper bound on μ_i is zero if $y_i > 0$ and $+\infty$ otherwise, so that the problem only has the $\mu_\chi \geq \boldsymbol{0}$ variables, where $\chi = \chi(\ \boldsymbol{y}\ ) = \{\ i : y_i = 0\ \}$: hence, it can be (efficiently) solved with the BTT algorithm developed in Chapter I.3. In fact, since (Cf. §I.3.5) its Quadratic Dual is

($\varepsilon\delta$-DSAD$'_{\beta,\boldsymbol{y}}$) $\quad \min_{\boldsymbol{d},v} \left\{ v + \tfrac{1}{2} \left\| \boldsymbol{d} \right\|^2_2 : v\boldsymbol{e} \geq \boldsymbol{d}\boldsymbol{G}_\beta\ ,\ \boldsymbol{0} \leq \boldsymbol{d}_\chi \right\}$

a specialized version of BTT has been developed for dealing with just nonnegativity (instead of general "box") constraints, where some of the operations are performed more efficiently since the RHS is known to be zero. At each step, the optimal solution ($\bar{\theta}$, $\bar{\mu}_\chi$) of ($\varepsilon\delta$-SAD$'_{\beta,\boldsymbol{y}}$) determines a tentative optimal point $\bar{\boldsymbol{x}}^h = \Sigma_{i \in \beta} \boldsymbol{x}^h_i$, and the dual optimal solution $\bar{\boldsymbol{d}} = \boldsymbol{G}_\beta \bar{\theta} + \bar{\mu}_\chi$ is the *gradient* $\Sigma_{h \in \mathcal{K}} \bar{\boldsymbol{x}}^h - \boldsymbol{u}$ of the objective function at $\bar{\boldsymbol{x}}$ projected onto the $\boldsymbol{0} \leq \boldsymbol{d}_\chi$ constraints: hence, the *linearized subproblem* that needs to be solved is

(ε-LP$'_{\bar{\boldsymbol{d}},\boldsymbol{y}}$) $\quad \min_{\boldsymbol{x}} \left\{ \bar{\boldsymbol{d}} (\Sigma_{h \in \mathcal{K}} \boldsymbol{x}^h - \boldsymbol{u}\ ) : \boldsymbol{x}^h \in X^h_\varepsilon(\ \boldsymbol{y}\ )\ \forall\ h \in \mathcal{K} \right\}$

which can be restated as

(ε-LP$'_{\bar{\boldsymbol{d}},\boldsymbol{y}}$) $\quad \min_{\boldsymbol{w}} \left\{ \bar{\boldsymbol{d}}\boldsymbol{w} : \boldsymbol{w} \in \partial_{kn\varepsilon} \varphi(\ \boldsymbol{y}\ ) \right\}$

i.e. the problem of determining the *kn*ε-*directional derivative* of $\varphi()$ in $\boldsymbol{y}$ along $\bar{\boldsymbol{d}}$ (Cf. Chapter I.1); in other words, z(ε-LP$'_{\bar{\boldsymbol{d}},\boldsymbol{y}}$) is the *exact slope* of $\varphi()$ along the (feasible) direction $\bar{\boldsymbol{d}}$, clearly giving an excellent stopping criterion for the early termination of the Simplicial Decomposition.

From the practical viewpoint, the typical implementation of a Relaxation Min-Cost Flow solver can be easily modified in order to "block" the flow on the arcs that are active/inactive in the optimal solution: for instance, in the `RelaxIV` class three vectors `X[]`, `Cap[]` and `b[]` contain the current flows x_{ij}, residual capacities $u_{ij} - x_{ij}$ and node imbalances $\Sigma_{j:\,(i,j) \in \mathcal{A}} x_{ij} - \Sigma_{j:\,(j,i) \in \mathcal{A}} x_{ji} - b_i$ respectively, so that an arc (i, j) will have `X[` *ij* `]` $= u_{ij}$ and `Cap[` *ij* `]` $= 0$ if active, or `X[` *ij* `]` $= 0$ and `Cap[` *ij* `]` $= u_{ij}$ if inactive. Hence, to block an active/inactive arc it is sufficient to set `X[` *ij* `]` $= 0$ or `Cap[` *ij* `]` $= 0$ respectively *without* changing the node balance vector `b[]`: since the code reoptimizes from the current solution, it will keep the arcs to the current value until they are "unblocked" by resetting the

correct values of the `X[]` and `Cap[]` vectors.

Despite these promising features, the computational results obtained so far are rather disappointing: in Table 29, the "pure" Steepest Ascent approach (*SA*), its ε-variant (ε-*SA*) and MMCFB are compared for the solution of a subset of the Canad problems (Cf. Sec. II.2.3.1 and II.2.4.1), aggregated in groups of similar sizes.

				SA	ε- SA					MMCFB	
k	*n*	*m*	*size*	Time	Time	φ# SD	Time SD	φ# LS	Time LS	Time	φ#
10	50	469	4685	13.28	13.58	110	8.85	70	1.58	1.08	22
10	50	625	6250	3600*	1580.00	10078	1160.40	976	47.30	36.00	37
40	20	252	10064	20.64	15.20	84	7.66	46	1.64	0.92	17
10	100	1600	16000	3600*	239.00	700	169.00	293	20.50	18.00	69
10	100	2500	25000	24.00	25.50	26	15.50	25	3.90	2.55	15
200	20	267	53360	18.68	18.20	31	6.38	12	2.14	1.32	7
100	35	545	54450	39.15	40.00	46	11.68	22	4.21	2.10	10
100	50	625	62500	412.00	277.00	316	190.30	65	2.00	18.00	33
100	100	1600	160000	136.50	88.00	10	74.50	9	3.75	5.25	6

Table 29: comparative results of the [ε-]Steepest Ascent and Bundle approach

The "pure" Steepest Ascent version has serious convergence problems on the more difficult instances (on the entries marked with "*" it was stopped after having run for one hour without producing the result), that are solved by the introduction of the ε-scaling mechanism: nevertheless, the ε-SA algorithm is always one order of magnitude slower than the Bundle approach, being up to a factor of 50 worse on the difficult instances. The results show that the solution of (ε-SAD$'_y$) is very costly, in terms of the number of (ε-LP$'_{\bar{d},y}$) subproblems solved, if compared with both the Line Search phase and the *overall* MMCFB: moreover, the surprising negative result is that the solution times of (ε-LP$'_{\bar{d},y}$) are from two to three times *larger* than those of the Lagrangean subproblems (RMF^h_y), despite that on average the 95% of the arcs is blocked. Hence, although substantially negative, these results suggest a number of possible implementative and algorithmic improvements:

- a different implementation should be used for solving (ε-LP$'_{\bar{d},y}$), that be capable of exploiting the large number of blocked arcs: for instance, the optimal solution of (RMF^h_y) is primal but not dual feasible for the *h*-th component of (ε-LP$'_{\bar{d},y}$), so that a primal (MCF) solver might be more efficient than the (dual) Relaxation method, at least for the first call;
- on the difficult instances, the direction is not of a good quality, resulting in a too large number of iterations: this may be in part due to the "myopic" view of the constraints, i.e. the $\delta = 0$ choice, so that one sensible (although not theoretically "clean") alternative could be to solve the slightly different (Dual) Master problem

  $$\bar{\boldsymbol{d}} = \text{argmin}_{d,v} \left\{ v + 1/2 \left\| \boldsymbol{d} \right\|_2^2 : v\boldsymbol{e} \geq \boldsymbol{d}\boldsymbol{G}_\beta \, , -\boldsymbol{y} \leq \boldsymbol{d} \right\}$$

  that would ensure $\boldsymbol{y} + \bar{\boldsymbol{d}}$ to be feasible, in the spirit of the Bundle approach;
- analogously, it may be possible to replace the Line Search with a "one step" approach where the steplength be derived by the available information on $\boldsymbol{d}$: this is benefical in the Bundle case;
- the ε-scaling mechanism can be improved, for instance by gradually decreasing ε without waiting for an ε-optimal solution to be obtained: this is analogous to the gradual increase of the penalty

parameter t in the Augmented Lagrangean approach, that is known to be crucial for the performances;

- finally, it is possible to replace the Simplicial Decomposition method with an entirely different, more efficient algorithm: for instance, the Block-Jacobi method used in [ZaA95] within an Augmented Lagrangean approach could be used in this setting as well.

To conclude, ε-Steepest Ascent approaches to the maximization of the Lagrangean function could be an interesting alternative, due to their capability of greatly reducing the size of the Quadratic Problem to be solved (e.g. with respect to Augmented Lagrangean approaches) and to produce an approximate optimal (primal) solution at the end of each ε-scaling phase: however, further research is needed for obtaining good computational results in practice.

II.4.2 An ε-Steepest Ascent Potential Decomposition approach

Although the most natural, the Multicommodity structure is not the only source of decomposition in (MMCF): a seldom noticed alternative is the *Potential Decomposition*, i.e. the Lagrangean relaxation of the flow conservation constraints with the vector of potentials $\Pi = [\, \pi^1 \,..\, \pi^k \,]$

$$(\mathrm{PDMF}_\Pi) \quad \left\{ \begin{array}{l} \min \sum_{h \in \mathcal{K}} (\boldsymbol{c}^h - \pi^h \boldsymbol{E}) \boldsymbol{x}^h \\ \boldsymbol{0} \le \boldsymbol{x}^h \le \boldsymbol{u}^h \quad \forall\ h \in \mathcal{K} \\ \sum_{h \in \mathcal{K}} \boldsymbol{x}^h \le \boldsymbol{u} \end{array} \right\} + \sum_{h \in \mathcal{K}} \pi^h \boldsymbol{b}^h,$$

that decomposes into m independent *continuous Knapsack* problems

$$(\mathrm{K}_\Pi^{ij}) \qquad \psi_{ij}(\Pi) = \min\Big\{ \sum_{h \in \mathcal{K}} (c_{ij}^h - \pi_i^h + \pi_j^h) x_{ij}^h : 0 \le x_{ij}^h \le u_{ij}^h \ \forall\ h \in \mathcal{K}, \sum_{h \in \mathcal{K}} x_{ij}^h \le u_{ij} \Big\},$$

and the corresponding Lagrangean Dual

$$(\mathrm{PDMCF}) \quad \max_\Pi \Big\{ \psi(\Pi) = \sum_{(i,j) \in \mathcal{A}} \psi_{ij}(\Pi) + \sum_{h \in \mathcal{K}} \pi^h \boldsymbol{b}^h : \Pi \in \mathcal{R}^{n \times k} \Big\}.$$

This approach differs from the Relaxation and Row Action methods (Cf. Sec. II.1.4.5) in that (PDMF_Π) decomposes in only m independent "complex" subproblems, while in the Relaxation method the mutual capacity constraints are relaxed as well, so that there are km "simple" independent subproblems, and in the Row Action method the bounds $[0, u_{ij}^h]$ are also relaxed at the expense of introducing $2km$ more Lagrangean variables.

Requiring the (unconstrained) maximization of a NonDifferentiable function, (PDMCF) can be attacked with the same approaches devised for (DMMCF), such as a Bundle method: this has not been done in the past since $\psi(\Pi)$ can be expected to be "more difficult" to maximize than $\varphi(\boldsymbol{y})$. In fact, if k is large then Π is much longer than $\boldsymbol{y}$ (kn instead of m variables); furthermore, being Lagrangean variables of *equality* constraints all the Πs can be expected to be "active" in the optimum (almost, since $\boldsymbol{E}$ is well-known to be $c(G)$-rank deficitary, where $c(G)$ is the number of distinct maximal connected components of G). However, it is also likely that $\psi()$ can be computed (much) faster than $\varphi()$; furthermore, the "difficulty" of (PDMCF) can be expected to be less dependent from the "tightness" of the mutual capacity constraints, so that it could be competitive on instances that are

"hard" for (DMMCF); finally, there are cases where kn is not "too larger" than m. Actually, the PDS problems (Cf. Sec. II.2.3.3) have all the above characteristics: they have large m and small k, so that $\varphi(\boldsymbol{y})$ requires the solution of (few) large (MCF)s while $\psi(\Pi)$ requires the solution of (many) tiny Knapsacks; they are known (Cf. Sec. II.2.4.3) to be difficult; they are "sparse" ($m / n \approx 3$) and have "active" mutual capacity constraints on about the 30% of the arcs, so that the $kn \approx 3m$ and the "active" Πs can be expected to be "only" about 10 times the "active" $\boldsymbol{y}$s.

Hence, a Bundle approach to (PDMCF) might be sensible, but in this Paragraph the development of a ε-Steepest Ascent (Cf. §II.4) will be sketched instead: with the natural definition of *reduced cost* $\bar{c}^h_{ij} = c^h_{ij} - \pi^h_i + \pi^h_j$, and dropping the superscript (i, j) for ease of notation, the subproblems are

$$(K_\Pi) \qquad \psi(\Pi) = \min\left\{ \sum_{h \in \mathcal{K}} \bar{c}^h x^h : 0 \le x^h \le u^h \ \forall\, h \in \mathcal{K}, \sum_{h \in \mathcal{K}} x^h \le u \right\}$$

(where no confusion should arise with the "global" $\psi()$) and their feasible region will be denoted by Φ_{ij}; it will also be sometimes convenient to consider the multiflow $\boldsymbol{x} = [\, \boldsymbol{x}^1 \,..\, \boldsymbol{x}^k \,]$ as partitioned into m vectors $\boldsymbol{x}_{ij} = [\, \boldsymbol{x}^1_{ij} \,..\, \boldsymbol{x}^k_{ij} \,]$. (K_Π) is a "simple" continuous knapsack problem (all the items have "size"1), and its optimal solutions set can be easily characterised: if $\sigma = \sigma_{ij} = \sigma_{ij}(\Pi) : \mathcal{K} \to \mathcal{K}$ is (one of) the permutation(s) that orders the reduced costs in nonincreasing sense, i.e.

$$\bar{c}^{\sigma(1)} \le \bar{c}^{\sigma(2)} \le \ldots \le \bar{c}^{\sigma(k)},$$

then an optimal solution of (K_Π) can be - roughly speaking - obtained by "picking entirely" the objects $\sigma(1), \sigma(2)$ ... until either an object with $\bar{c} \ge 0$ or the total capacity u is reached, eventually choosing only a fraction of the last object (the one with the higher cost). More formally, a flow vector $\boldsymbol{x}$ is a *pseudoflow* if $\boldsymbol{x}_{ij} \in \Phi_{ij}\ \forall\ (i, j) \in \mathcal{A}$, and for any pseudoflow $\boldsymbol{x}$ let the *aggregated (pseudo)flow* be $\bar{\boldsymbol{x}} = \sum_{h \in \mathcal{K}} \boldsymbol{x}^h$ $(\le \boldsymbol{u})$: given a pair $(\Pi, \boldsymbol{x})$, the vector of *critical reduced costs* η is

$$\eta_{ij} = \begin{cases} 0 & \text{if } \bar{x}_{ij} < u_{ij} \\ \max_{h \in \mathcal{K}}\{\, \bar{c}^h_{ij} : x^h_{ij} > 0 \,\} & \text{if } \bar{x}_{ij} = u_{ij} \end{cases}$$

and the arc (i, j) is *saturated* if $\bar{x}_{ij} = u_{ij}$, and *nonsaturated* otherwise; note that any nonsaturated arc has $\eta_{ij} = 0$, but the converse is not true, i.e. η_{ij} can be zero even if (i, j) is saturated. Then, for a given pair $(\Pi, \boldsymbol{x})$, the *commodity h* over the arc (i, j) is said

active if $\bar{c}^h_{ij} < \eta_{ij}$

balanced if $\bar{c}^h_{ij} = \eta_{ij}$

inactive if $\bar{c}^h_{ij} > \eta_{ij}$

and let A^{ij}_Π (respectively B^{ij}_Π, I^{ij}_Π) be the set of active (resp. balanced, inactive) commodities on (i, j): $(\Pi, \boldsymbol{x})$ satisfies the Complementary Slackness Conditions if and only if for each (i, j)

$$(CSC_{ij}) \qquad \begin{cases} h \in A^{ij}_\Pi & \Rightarrow\ x^{ij}_\Pi = u^{ij}_\Pi \\ h \in I^{ij}_\Pi & \Rightarrow\ x^{ij}_\Pi = 0 \end{cases}$$

i.e. $\boldsymbol{x}_{ij}$ is optimal for the Lagrangean subproblem (K^{ij}_Π).

Lemma II.4.2.1: *for any fixed* Π, $\boldsymbol{x}_{ij}$ *satisfies* (CSC_{ij}) *if and only if* $\boldsymbol{x}_{ij}$ *is optimal for* (K_Π^{ij}).

Proof: A flow (possibly not a pseudoflow) $\boldsymbol{x}$ is optimal for the Lagrangean relaxation of (K_Π)

($RK_{\Pi,y}$) $\quad \min_{\boldsymbol{x}} \{ \sum_{h \in \mathcal{K}} (\bar{c}^h + y) x^h : 0 \le x^h \le u^h \; \forall \; h \in \mathcal{K} \} - yu$

(for notational simplicity, the index ij is disregarded) if and only if

$$x^h \in \begin{cases} \{ \; 0 \; \} & \text{if } \bar{c}^h > -y \\ \{ \; u^h \; \} & \text{if } \bar{c}^h < -y \\ [\; 0, \; u^h \;] & \text{if } \bar{c}^h = -y \end{cases} ;$$

since $y \ge 0$, all the x^h with positive reduced costs will be zero, while possibly not all the x^h with negative reduced costs will be nonzero. y^* is an optimal solution of the corresponding Lagrangean Dual if and only if there exists such an $\boldsymbol{x}$ (w.r.t. y^*) that is also a pseudoflow; for any permutation σ as defined above, it is clear that $\bar{c}^{\sigma(h)} < -y \Rightarrow \bar{c}^{\sigma(p)} < -y \; \forall \; p \le h$, and therefore the *critical* index $\bar{h}$

$$\bar{h} = \max \{ \; h \in \mathcal{K} : \sum_{p \le h} u^p < u \; , \; \bar{c}^{\sigma(h)} < \bar{c}^{\sigma(h+1)} \; , \; \bar{c}^{\sigma(h)} < 0 \; \}$$

(or, more precisely, $\sigma(\bar{h})$) indicates the last "object" that *must* be entirely into the knapsack in *any* optimal solution of (K_Π). In fact, the object $p = \sigma(\bar{h}) + 1$

- either cannot be put entirely into the knapsack without exceeding the capacity u;
- or has $\bar{c}^p \ge 0$, hence it must not (necessarily, if $\bar{c}^p = 0$) be put into the knapsack;
- or, finally, can be put entirely into the knapsack without exceeding the capacity u, but it has exactly the same *strictly negative* reduced cost of other objects, not all of which can be put entirely into the knapsack without exceeding u: in this case, p has exactly the same "right" of the others to be put into the knapsack, so it *can* be chosen but it *must not be necessarily* chosen.

Then, it is easy to see that $y^* = \max \{ -\bar{c}^p , 0 \} = -\eta_{ij}$, and the proof is over. ♦

Hence, (Π, $\boldsymbol{x}$) satisfies (CSC) and $\boldsymbol{x}$ satisfies the flow conservation constraints if and only if Π and $\boldsymbol{x}$ are optimal solutions of (PDMCF) and (MMCF) respectively: also, the optimal solutions set $X(\Pi)$ of (K_Π) is the set of all pseudoflows $\boldsymbol{x}$ such that

$$\sum_{h \in B_\Pi} x^h \le u - \sum_{h \in A_\Pi}$$
$$x^h = u^h \qquad \forall \; h \in A_\Pi$$
$$x^h = 0 \qquad \forall \; h \in I_\Pi$$

so that the Steepest Ascent direction finding problem

(SAD_Π) $\quad \min_{\boldsymbol{x}} \{ \sum_{h \in \mathcal{K}} \sum_{i \in \mathcal{N}} (\sum_{j:\,(i,j) \in \mathcal{A}} x_{ij}^h - \sum_{j:\,(j,i) \in \mathcal{A}} x_{ji}^h - b_i^h)^2 : \boldsymbol{x}_{ij} \in X_{ij}(\boldsymbol{y}) \; \forall \; (i,j) \in \mathcal{A} \}$

is again a (QP) with separable feasible region and nonseparable objective function. (SAD_Π) has a more "complicated" Hessian than ($SAD_{\boldsymbol{y}}'$) (Cf. §II.4.1), but the feasible region is much "simpler": hence, the subproblems of a block-coordinate descent approach to the former would be simpler than the Quadratic (MCF)s needed for the latter, yelding for instance a finer-grain parallel block-Jacobi method than in [ZaA95]. As in §II.4.1, the extension to an ε-steepest ascent approach is quite straightforward, essentially requiring only the definition of ε-*Complementary Slackness Conditions*

characterising the set of ε-optimal solutions $X_{ij}^{\varepsilon}(\Pi)$: rather than further elaborating on the details, however, we prefer to move to a related but different idea, that will be presented in the next Paragraph.

II.4.3 Towards a combinatorial approach

The above Potential Decomposition approach suggests to extend to the Linear (MMCF) the Primal-Dual methods for the single-commodity (MCF) (Cf. Sec. II.1.4.6): some steps in this direction - but, unfortunately, not yet a complete algorithm - will be presented in this Paragraph. A more detailed description of the algorithm for the single-commodity case is in order: Primal-Dual methods attack the "Potential Decomposition" (into *m* univariate subproblems) of (MCF)

(PD) $\quad \max_{\pi} \{ \psi(\pi) : \pi \in \mathcal{R}^n \}$

where

(PR$_\pi$) $\quad \psi(\pi) = \Sigma_{(i,j) \in \mathcal{A}} \min_x \{ (c_{ij} - \pi_i + \pi_j) x_{ij} : 0 \le x_{ij} \le u_{ij} \} - \Sigma_{i \in \mathcal{N}} \pi_i b_i$

but, in the linear case, the coordinate ascent method is not sufficient due to the nondifferentiability of $\psi()$ or, equivalently, due to the fact that $X(\pi)$ - the set of optimal solutions of (PR$_\pi$) - is not necessarily a singleton. As in §II.4.1, however, $X(\pi)$ (and therefore $\partial\psi(\pi)$) can be described in terms of Complementary Slackness Conditions, i.e. of the partition $\mathcal{A} = A_\pi \cup I_\pi \cup B_\pi$ of the arcs among active, inactive and balanced: it is easy to see that $X(\pi)$ is the set of all pseudoflows where active and inactive arcs are *fixed* to their (upper and lower) bounds, while balanced arcs are *free*. Therefore, Primal-Dual methods maintain a pair $(\pi, \boldsymbol{x})$, where $\boldsymbol{x} \in X(\pi)$ is changed in the *primal phase* to reduce the total infeasibility $\| \boldsymbol{Ex} - \boldsymbol{b} \|_1$, i.e. the ($L_1$-)norm of the corresponding subgradient: at the end of the primal phase, either an $\boldsymbol{x}' \in X(\pi)$ is found such that $\boldsymbol{Ex}' = \boldsymbol{b}$, or an ascent direction is identified that is used in the *dual phase* to obtain better potentials π'. From the "dual" viewpoint, the distinguishing feature of Primal-Dual methods is the use of *elementary* ascent directions only, corresponding to the characteristic vector $\boldsymbol{e}_{\mathcal{S}}$ of some cutset $\mathcal{S} \subset \mathcal{N}$: the following nice characterisation of the *directional derivative* $\psi'(\pi, \boldsymbol{e}_{\mathcal{S}})$ with respect to an elementary direction is a key point in their analysis. The *imbalance* of a cutset $\mathcal{S}$ w.r.t. the pseudoflow $\boldsymbol{x}$ is

$$\Delta_{\mathcal{S}}(\boldsymbol{x}) = \Sigma_{i \in \mathcal{S}} \Delta_i(\boldsymbol{x})$$

where $\quad \Delta_i(\boldsymbol{x}) = \sum_{j:(j,i) \in \mathcal{A}} x_{ji} - \sum_{i:(i,j) \in \mathcal{A}} x_{ij} - b_i$

and the sets of *forward* and *backward* arcs in the induced cut are

$$\mathcal{S}^+ = \{ (i,j) : i \in \mathcal{S}, j \notin \mathcal{S} \} \qquad \text{and} \qquad \mathcal{S}^- = \{ (i,j) : i \notin \mathcal{S}, j \in \mathcal{S} \}$$

(with a little abuse of notation, the cut $\mathcal{S}^+ \cup \mathcal{S}^-$ will also be denoted by $\mathcal{S}$): then, with the "capacity" of $\mathcal{S}$ w.r.t. $\boldsymbol{x}$ defined as

$$\beta_{\mathcal{S}}(\boldsymbol{x}) = \left(\sum_{(i,j) \in \mathcal{S}^+ \cap B} u_{ij} - x_{ij} \right) + \left(\sum_{(i,j) \in \mathcal{S}^- \cap B} x_{ij} \right);$$

it is possible to show that

$$\psi'(\pi, \boldsymbol{e}_{\mathcal{S}}) = \Delta_{\mathcal{S}}(\boldsymbol{x}) - \beta_{\mathcal{S}}(\boldsymbol{x}).$$

Clearly, $\Delta_S(\boldsymbol{x})$ units of flow *must* be "pushed away" from $\mathcal{S}$ to make all the nodes balanced, and $\beta_S(\boldsymbol{x})$ measures how much flow *can* be "pushed away" from $\mathcal{S}$ by only using balanced arcs (Figure 15): if the imbalance of a cut exceeds its capacity, no feasible $\boldsymbol{x}$ (in which all the nodes, and hence all those in $\mathcal{S}$, are balanced) can exist in $X(\pi)$, and the potentials π_S must be increased to create some new balanced arcs that allow more flow to be "extracted" from $\mathcal{S}$. Relaxation methods try to use "small" cutsets $\mathcal{S}$: some node i with $\Delta_i(\boldsymbol{x}) > 0$ is used as a tentative $\mathcal{S}$, and a coordinate ascent step is performed if the above relation holds. Otherwise, the *reduced graph* $G_\pi(\mathcal{N}, B_\pi)$ (containing the balanced arcs only) is visited starting from i in search of a node j with $\Delta_j(\boldsymbol{x}) < 0$, where (i, j) is used "forward" if it is not full ($x_{ij} < u_{ij}$) and "backward" if it is not empty ($x_{ij} > 0$): any balanced node h reached during the process is added to $\mathcal{S}$ (Figure 16), each time checking for strict positivity of $\psi'()$ and performing a dual step in case. If a suitable j is found, the imbalance of both i and j is strictly decreased by sending flow on the (balanced) path joining them: otherwise, the visit terminates because no "forward nonfull" or "backward nonempty" arc exists for the current $\mathcal{S}$, but then $\beta_S(\boldsymbol{x}) = 0$ while $\Delta_S(\boldsymbol{x}) > 0$ and $\mathcal{S}$ provides an ascent direction. Hence, the search for an (elementary) ascent direction is performed via efficient combinatorial operations on G_π, that are in fact a *labelling* (or *augmenting path*) approach to the Max-Flow problem aimed at finding the pseudoflow $\boldsymbol{x} \in X(\pi)$ that minimizes the total imbalance $\Delta_{\mathcal{N}}(\boldsymbol{x})$: the Steepest Ascent method just solves the Max-Flow exactly, obtaining a Min-Cut with $\beta_S(\boldsymbol{x}) = 0$ and therefore a "normalized" L_1-Steepest Ascent direction (despite the "negative" result in §I.4.1).

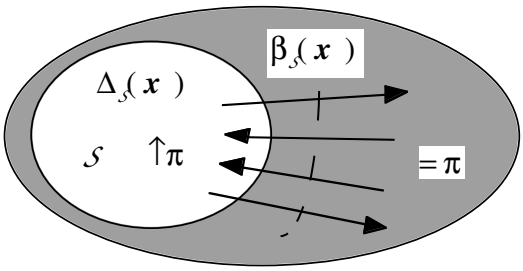

Figure 15: the "dual" phase of Relaxation methods

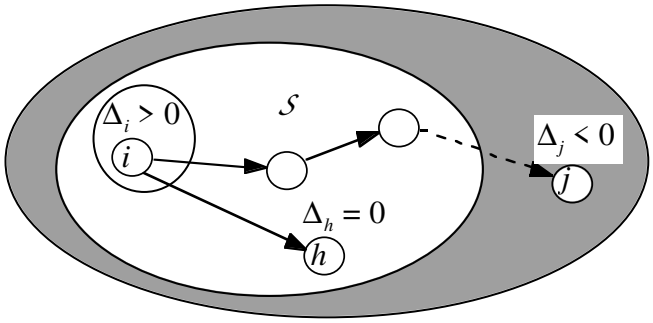

Figure 16: the "primal" phase of Relaxation methods

Unfortunately, generalization of the primal phase to (MMCF) is problematic, since Multicommodity Max-Flow problems lack of efficient combinatorial approaches corresponding to the labelling algorithm: conversely, the characterisation of $\psi'(\pi, \cdot)$ can be quite easily extended. In fact, for a pair $(\Pi, \boldsymbol{x})$ where $\boldsymbol{x}$ is a pseudoflow and the (CSC) are satisfied - i.e. $\boldsymbol{x}_{ij}$ is optimal for (K_Π^{ij}) - the extended definitions of active, balanced and inactive commodities on (i, j) have already been given in §II.4.2, so that it is only necessary to define what an *elementary* direction is in the (MMCF) case; several alternative definitions are - at least conceptually - possible, perhaps the most general one being

Definition II.4.3.1: *given a set of k cuts* $\Sigma = \{ \mathcal{S}^1, .., \mathcal{S}^k \}$ $(\mathcal{S}^h \subseteq \mathcal{N} \ \forall h \in \mathcal{K})$ *the direction* $\boldsymbol{d}_\Sigma = [\boldsymbol{e}_{\mathcal{S}^1}, .., \boldsymbol{e}_{\mathcal{S}^k}]$ *in the space of the dual multipliers* Π *is called an elementary direction; furthermore,* $\boldsymbol{d}_\Sigma$ *is called*

- *horizontal if* $\Sigma = \{ \varnothing, .., \mathcal{S}^h, .., \varnothing \}$ $(\mathcal{S}^h \subset \mathcal{N})$
- *vertical if* $\mathcal{S}^1 = \mathcal{S}^2 = .. = \mathcal{S}^k$

Horizontal and vertical directions (the names come from imagining each commodity living on a

separate copy of G, and the copies in a vertical stack) will be succinctly indicated by $\boldsymbol{d}_{\mathcal{S},h}$ and $\boldsymbol{d}_{\mathcal{S}}$ respectively: a characterisation of $\psi'()$ will be given for the horizontal directions case. By the definition of $\psi()$, its directional derivative in Π w.r.t. $\boldsymbol{d}_{\mathcal{S},h}$

$$\psi'(\Pi, \boldsymbol{d}_{\mathcal{S},h}) = \Sigma_{i \in \mathcal{S}} b_i^h + \Sigma_{(i,j) \in \mathcal{A}} \psi'_{ij}(\Pi, \boldsymbol{d}_{\mathcal{S},h})$$

is completely described by the m separate directional derivatives of the simpler functions $\psi_{ij}()$: since each $\psi_{ij}()$ is a polyhedral function, its restriction over a line is piecewise linear, so that for a "sufficiently small" $\alpha > 0$ (to be derived precisely later)

$$\delta_{ij} = \psi'_{ij}(\Pi, \boldsymbol{d}_{\mathcal{S},h}) = \frac{\psi_{ij}(\Pi + \alpha \boldsymbol{d}_{\mathcal{S},h}) - \psi_{ij}(\Pi)}{\alpha}.$$

Moving along the horizontal elementary direction $\boldsymbol{d}_{\mathcal{S},h}$, only the dual multipliers π^h ($\Rightarrow$ the reduced costs $\bar{\boldsymbol{c}}^h$) of commodity h are modified, so that the usual notions of *forward* and *backward* arcs in $\mathcal{S}$ can be used - the matter is clearly more complicated in the general case: hence, one has

$$\bar{c}^h_{ij}(\pi^h + \alpha \boldsymbol{e}_{\mathcal{S}}) = \bar{c}^h_{ij}(\pi^h) + \begin{cases} 0 & \text{if } (i,j) \notin \mathcal{S} \\ -\alpha & \text{if } (i,j) \in \mathcal{S}^+ \\ +\alpha & \text{if } (i,j) \in \mathcal{S}^- \end{cases}$$

i.e. $(i,j) \notin \mathcal{S} \Rightarrow \delta_{ij} = 0$. Moreover, since h is fixed the set A_Π (respectively I_Π, B_Π) of active (resp. inactive, balanced) arcs can be defined as

$$A_\Pi = \{ (i,j) : h \in A^{ij}_\Pi \text{ (resp. } I^{ij}_\Pi, B^{ij}_\Pi) \}$$

so that, recalling the definition of critical reduced cost η_{ij} from §II.4.2, for all

$$\alpha < \bar{\alpha} = \min_{(i,j) \in \mathcal{S}} \{ |\bar{c}^h_{ij} - \eta_{ij}| : ((i,j) \in \mathcal{S}^+ \cap I_\Pi) \text{ or } ((i,j) \in \mathcal{S}^- \cap A_\Pi) \}$$

no active or inactive arc becomes balanced: hence, for each $\alpha < \bar{\alpha}$, $\psi_{ij}(\Pi + \alpha \boldsymbol{d}_{\mathcal{S},h}) - \psi_{ij}(\Pi)$ is a linear function of α. If (i,j) is either active or inactive, $\boldsymbol{x}_{ij}$ remains optimal for (K^{ij}_Π), therefore

$$\psi_{ij}(\Pi + \alpha \boldsymbol{d}_{\mathcal{S},h}) - \psi_{ij}(\Pi) = \bar{c}^h_{ij}(\pi^h + \alpha \boldsymbol{e}_{\mathcal{S}}) - \bar{c}^h_{ij}(\pi^h) = \pm \alpha x^h_{ij}$$

so that $$\delta_{ij} = \begin{cases} 0 & \text{if } (i,j) \in I \\ -u^h_{ij} & \text{if } (i,j) \in \mathcal{S}^+ \cap A \\ +u_{ij}{}^h & \text{if } (i,j) \in \mathcal{S}^- \cap A \end{cases} ;$$

on the contraty, any $(i,j) \in B_\Pi$ becomes either active or inactive for any $\alpha > 0$, and the optimal solution of (K^{ij}_Π) (possibly) changes. In this case, the *free forward space*

$$\Delta_{ij} = u_{ij} - \sum_{p \in A^{ij}_\Pi} x^p_{ij} = u_{ij} - \sum_{p \in A^{ij}_\Pi} u^p_{ij}$$

and the *free backward space*

$$\nabla^h_{ij} = \begin{cases} x^h_{ij} & \text{if } \eta_{ij} = 0 \\ \sum_{p \in B^{ij}_\Pi / \{h\}} u^p_{ij} - x^p_{ij} & \text{if } \eta_{ij} < 0 \end{cases}$$

of (i,j) w.r.t. $(\Pi, \boldsymbol{x})$ and h play a role: the rationale of these definitions is clarified by the proofs of the following two Lemmas where, for notiational simplicity, the reference to (i,j) is dropped.

Lemma II.4.3.2: *if (i, j)* $\in \mathcal{S}^+ \cap B_\Pi$, *then* $\delta_{ij} = -\min\{\ \Delta_{ij}\ ,\ u_{ij}^h\ \}$.

Proof: The total flow and the contribution to the value of $\psi()$ of the balanced commodities are

$$y = \sum_{p \in B_\Pi} x^p \qquad \text{and} \qquad z = \sum_{p \in B_\Pi} \bar{c}^p x^p = \eta y$$

since $\bar{c}^p = \eta$ for all $p \in B_\Pi$: it is easy to see that $z = \eta\Delta$, as if $\eta < 0$ then $\bar{x} = u$ and $\Delta = y$, otherwise $\eta = 0$ and then $z = 0 = \eta\Delta$. Now, let $\Pi(\alpha) = \Pi + \alpha \boldsymbol{d}_{\mathcal{S},h}$ and consider any optimal solution of $(K_{\Pi(\alpha)})$: since $\bar{c}^h$ has decreased, flow from the (former) balanced commodities is "transferred" to the h-th until either $x^h = u^h$ or the maximum quantity of flow that can be transferred, i.e. $\Delta - x^h$, has been assigned to the h-th commodity; the flows of active and inactive commodities over (i, j) don't change, hence their contribution to δ is 0. Therefore, if $v = \min\{\ \Delta\ ,\ u^h\ \}$ is the new value of x^h, the total contribution of the commodities in (the former) B_Π to the (new) value of $\psi(\ \Pi(\ \alpha\ )\ )$ is $\eta(\ \Delta - v\ ) + (\ \eta - \alpha\ )v = \eta\Delta - \alpha v$, hence $\delta = -v$. ♦

Lemma II.4.3.3: *if (i, j)* $\in \mathcal{S}^- \cap B_\Pi$, *then* $\delta_{ij} = \max\{\ x_{ij}^h - \nabla_{ij}^h\ ,\ 0\ \}$.

Proof: if $\eta = 0$ ($\Rightarrow \nabla^h = x^h$ by definition $\Rightarrow \delta = 0$), x^h can be set to 0 leaving the rest of the optimal solution unchanged, so that the contribution of the h-th commodity to z is $\eta x^h = 0x^h = (\ \eta - \alpha\ )0 = 0$. If $\eta < 0$, then in the optimal solution of $(K_{\Pi(\alpha)})$ x^h will decrease as long as the other "nonfull" balanced commodities can "absorb" it: hence, ∇^h represents the maximum quantity of flow that can be "absorbed away" from the h-th commodity, and two different cases can occur

- either $\nabla^h \geq x^h$ ($x^h - \nabla^h \leq 0$), i.e. the h-th commodity can be "absorbed" by other (balanced) commodities, with the same reduced cost η: in this case, $\psi()$ doesn't change, therefore $\delta = 0$;
- or $\nabla^h < x^h$, i.e. *not all* the h-th commodity can be "absorbed": then, $v = x^h - \nabla^h > 0$ is the h-th commodity "left", and $\psi(\ \Pi(\ \alpha\ )\ ) = \eta(\ \Delta - v\ ) + (\ \eta + \alpha\ )v = \eta\Delta + \alpha v$, therefore $\delta = v$.

Hence, the definition of ∇^h only serves to unify the two cases $\eta = 0$ and $\eta < 0$. ♦

As a consequence of the above Lemmas, one has

$$\psi'(\ \Pi\ ,\ \boldsymbol{d}_{\mathcal{S},h}\ ) = b_{\mathcal{S}}^h - \sum_{(i,j) \in \mathcal{S}^+ \cap A} u_{ij}^h + \sum_{(i,j) \in \mathcal{S}^- \cap A} u_{ij}^h$$
$$- \sum_{(i,j) \in \mathcal{S}^+ \cap B} \min\{\ \Delta_{ij}\ ,\ u_{ij}^h\ \} + \sum_{(i,j) \in \mathcal{S}^- \cap B} \max\{\ x_{ij}^h - \nabla_{ij}^h\ ,\ 0\ \};$$

however, this can be turned in a more significative form by defining

$$\Delta_{\mathcal{S}}^h(\ \boldsymbol{x}\ ) = b_{\mathcal{S}}^h - \sum_{(i,j) \in \mathcal{S}^+} x_{ij}^h + \sum_{(i,j) \in \mathcal{S}^-} x_{ij}^h$$

(the total imbalance of $\mathcal{S}$ w.r.t. $\boldsymbol{x}^h$) and studying the quantity $\beta_{\mathcal{S}}^h(\ \boldsymbol{x}\ ) = \Delta_{\mathcal{S}}^h(\ \boldsymbol{x}\ ) - \psi'(\ \Pi\ ,\ \boldsymbol{d}_{\mathcal{S},h}\ )$.

Theorem II.4.3.4:

$$\beta_{\mathcal{S}}^h(\ \boldsymbol{x}\ ) = \sum_{(i,j) \in \mathcal{S}^+ \cap B} (\ \min\{\ \Delta_{ij}\ ,\ u_{ij}^h\ \} - x_{ij}^h\ ) + \sum_{(i,j) \in \mathcal{S}^- \cap B} \min\{\ \nabla_{ij}^h\ ,\ x_{ij}^h\ \}$$

Proof: For enumeration of the terms in the expression for $\Delta_{\mathcal{S}}^h$ and the directional derivative:

- $b_{\mathcal{S}}^h$ is present in both, hence it elides;

– an *inactive* arc (i, j) gives no contribution to $\psi'()$, and contribution $x_{ij}^h = 0$ to Δ_S^h;
– an *active* arc (i, j) gives the same contribution $\pm u_{ij}^h$ (depending if it is a backward or forward arc) to $\psi'()$ and to Δ_S^h;
– a *balanced forward* arc (i, j) gives contribution $-\min\{ \Delta_{ij} , u_{ij}^h \}$ to $\psi'()$ and $-x_{ij}^h$ to Δ_S^h, hence $\min\{ \Delta_{ij} , u_{ij}^h \} - x_{ij}^h$ to β_S^h;
– a *balanced backward* arc (i, j) gives contribution $\max\{ x_{ij}^h - \nabla_{ij}^h , 0 \}$ to $\psi'()$ and x_{ij}^h to Δ_S^h, hence $x_{ij}^h - \max\{ x_{ij}^h - \nabla_{ij}^h , 0 \} = x_{ij}^h + \min\{ \nabla_{ij}^h - x_{ij}^h , 0 \} = \min\{ \nabla_{ij}^h , x_{ij}^h \}$ to β_S^h. ♦

Theorem II.4.3.4 justifies the definitions of *nonempty* and *nonfull* balanced arcs (w.r.t. $(\Pi, \boldsymbol{x})$ and h) as the $(i, j) \in B_\Pi$ such that

$$\min\{ \nabla_{ij}^h , x_{ij}^h \} > 0 \qquad \text{and} \qquad \min\{ \Delta_{ij} , u_{ij}^h \} > x_{ij}^h$$

respectively: since $\Delta_{ij} \geq x_{ij}^h$, all the terms in the expression for β_S^h are nonnegative, hence $\beta_S^h \geq 0$. As in the single-commodity case, β_S^h is the flow that can be "pushed away" from S, so that the result

$$\Delta_S^h > 0 \;\Rightarrow\; \psi'(\Pi , \boldsymbol{d}_{S,h}) > 0 \Leftrightarrow \text{the total imbalance of } S \text{ is strictly larger than the flow that can be "pushed away"}$$

is readily obtained; an immediate corollary is that if all the balanced arcs in the cut are *forward full* or *backward empty*, then $\boldsymbol{d}_{S,h}$ is a direction of ascent.

Analogous results can be derived for vertical elementary directions and for general elementary directions, but they require a clumsier notation and a set of proper definitions: a simpler proof can be derived by considering the more general class of Min-Cost Flow problems on Hypergraphs. In a (directed) *hypergraph* $G(\mathcal{N} , \mathcal{A})$ [GLN93], *hyperarcs* can have multiple heads and tails: more formally, $\mathcal{A} \subseteq 2^{\mathcal{N}} \times 2^{\mathcal{N}}$, i.e. each $a \in \mathcal{A}$ is an (ordered) pair $(H(a), T(a))$ of (possibly empty) subsets of the nodes, $H(a)$ being the *head* nodes and $T(a)$ being the *tail* nodes.
One important class of hypergraphs is that of *B-hypergraphs*, where each hyperarc can have multiple tails but at most one tail: the B-Hypergraph Min-Cost Flow problem

(BHMCF) $\quad \min_{\boldsymbol{x}} \{ \boldsymbol{cx} : \boldsymbol{Ex} = \boldsymbol{b} , \boldsymbol{0} \leq \boldsymbol{x} \leq \boldsymbol{u} \}$

where $\boldsymbol{E}$ is the node-hyperarcs incidence matrix having $| H(a) | + | T(a) |$ nonzeroes in the column $\boldsymbol{E}^a$ relative to hyperarc a, is general enough to properly contain (MMCF) as a subclass, since any instance of (MMCF) can be transformed into one of (BHMCF) on a graph with $kn + m$ nodes, km hyperarcs with two tails and one head and m "tailless" hyperarcs, as depicted in Figure 17. Nevertheless, a nice characterisation can be given for the directional derivative (w.r.t. an elementary direction $\boldsymbol{d}_S$) of the Lagrangean function

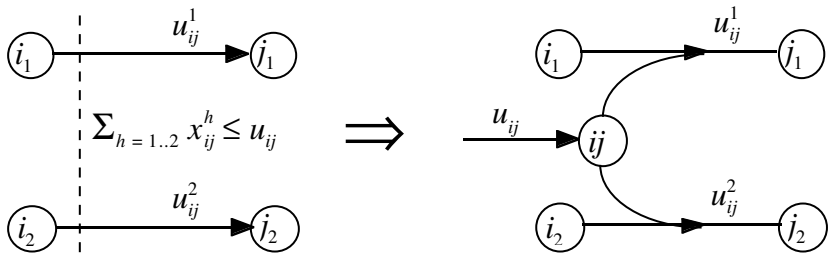

Figure 17: the reduction of (MMCF) to (BHMCF)

$$\psi(\pi) = \Sigma_{a \in \mathcal{A}} \min\{ \bar{c}_a(\pi) x_a : 0 \leq x_a \leq u_a \} + \pi \boldsymbol{b} ,$$

where the reduced costs are

$$\bar{c}_a(\pi) = c_a - \Sigma_{i \in T(a)} \pi_i + \Sigma_{j \in H(a)} \pi_j ,$$

by only properly generalizing the notion of *forward* and *backward* (hyper)arc. If the *directionality* of an hyperarc a with respect to a cutset $\mathcal{S}$ is

$$K(\mathcal{S}, a) = \left| T(a) \cap \mathcal{S} \right| - \left| H(a) \cap \mathcal{S} \right| = - \boldsymbol{e}_{\mathcal{S}} \boldsymbol{E}^a$$

a is forward if $K(\mathcal{S}, a) > 0$ and backward if $K(\mathcal{S}, a) < 0$ (Figure 18): the rationale of the definition is that a forward hyperarc has more tails than heads in $\mathcal{S}$, so that increasing x_a decreases the total imbalance $\Delta_{\mathcal{S}}(\boldsymbol{x})$ of $\mathcal{S}$, while the converse is true for a backward one.

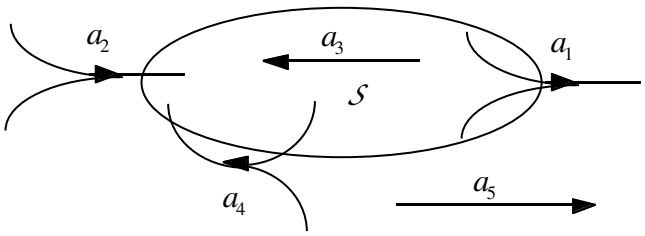

Figure 18: directionality - $K(\mathcal{S}, a_1) = 2$, $K(\mathcal{S}, a_1) = -1$, $K(\mathcal{S}, a_3) = K(\mathcal{S}, a_4) = K(\mathcal{S}, a_5) = 0$

Then, it can be easily shown that

$$\psi'(\pi, \boldsymbol{e}_{\mathcal{S}}) = \Delta_{\mathcal{S}}(\boldsymbol{x}) - \beta_{\mathcal{S}}(\boldsymbol{x})$$

where $\Delta_{\mathcal{S}}(\boldsymbol{x}) = \Sigma_{i \in \mathcal{S}} \left(\Sigma_{a : i \in H(a)} x_a - \Sigma_{a : i \in T(a)} x_a - b_i \right)$,

$$\beta_{\mathcal{S}}(\boldsymbol{x}) = \left(\sum_{a \in B : K(\mathcal{S}; a) > 0} K(\mathcal{S}, a)(u_a - x_a) \right) + \left(\sum_{a \in B : K(\mathcal{S}; a) < 0} K(\mathcal{S}, a) x_a \right)$$

and the set B of balanced hyperarcs is defined in the obvious way: note that hyperarcs having the same numbers of heads and tails in $\mathcal{S}$ (i.e. $K(\mathcal{S}, a) = 0$) do not contribute to $\beta_{\mathcal{S}}(\boldsymbol{x})$. Hence, $\psi'()$ is again the difference between the total imbalance of $\mathcal{S}$ and the flow that can be "pushed away" on the balanced hyperarcs: this can be easily extended to the case of "gains" in the hyperarcs, i.e. of not-all-unitary coefficients in the node-hyperarcs incidence matrix $\boldsymbol{E}$, and can be directly used to derive the results about general elementary directions in (MMCF).

However, no "pure" Relaxation method for (MMCF) or (HBMCF) has been obtained yet, because the primal phase - that finds either some sort of elementary direction or a feasible flow - seems to be much more complex than in the (MCF) case: in fact, the Max-Flow versions of the problems are not easily solvable, and it is not even clear if and how they would grant an ascent direction. Developing a "combinatorial" primal phase would be an interesting result, and could lead to efficient implementations: however, an alternative approach could be to provide *heuristics* for efficiently finding ascent directions. In fact, the convergence of a Primal-Dual method can be easily ensured by "immerging" it inside the Steepest Ascent approach of §II.4.2: the Steepest Ascent (QP) subproblem (SAD_{Π}) provides an "exact" primal phase to be used as "last resort", while the heuristics should provide low-cost ascent directions. Actually, the heuristics might also be capable of solving most instances, as indicated by some encouraging - although preliminary - computational results that have been independently obtained [DBR96] for Hypergraph (MCF) problems. The heuristics might exploit the special structure of the problem at hand, for instance searching for horizontal or vertical elementary directions for (MMCF), that are not likely to be identified when looking at the more general (BHMCF). Altogether, the lines of approach described in this Chapter seem to indicate that "hybrid" algorithms integrating smooth, nondifferentiable and combinatorial technique may in a future provide efficient means of solving (MMCF) or even more general classes of problems.