Duality in convex minimum cost flow problems on infinite networks and hypernetworks

Sevnaz Nourollahi Archis Ghate

June 12, 2017

Abstract

Minimum cost flow problems on infinite networks arise, for example, in infinite-horizon sequential decision problems such as production planning. Strong duality for these problems was recently established for linear costs using an infinite-dimensional Simplex algorithm. Here, we use a different approach to derive duality results for convex costs. We formulate the primal and dual problems in appropriately paired sequence spaces such that weak duality and complementary slackness can be established using finite-dimensional proof techniques. We then prove, using a planning horizon proof technique, that the absence of a duality gap between carefully constructed finite-dimensional truncations of the primal problem and their duals is preserved in the limit. We then establish that strong duality holds when optimal solutions to the finite-dimensional duals are bounded. These theoretical results are illustrated via an infinite-horizon shortest path problem. We also extend our results to infinite hypernetworks and apply this generalization to an infinite-horizon stochastic shortest path problem.

1 Introduction

We study a class of convex minimum cost flow problems in networks with a countably infinite number of nodes and a countably infinite number of arcs. Prototypical applications of such minimum cost flow problems arise in infinite-horizon sequential decision-making [13, 32, 36, 37]. When flow costs are linear, these mathematical programs become a special case of countably infinite linear programs (CILPs). CILP theory from [33, 34] was thus applied to establish weak duality and complementary slackness, and a Simplex algorithm was developed and employed to establish strong duality for these linear problems under integer data by Sharkey and Romeijn [36]. Similar duality results are not available for the more general case of convex costs. We next outline the difficulties in obtaining duality results for minimum cost flow problems on countably infinite networks and then summarize how we circumvent these hurdles in this paper.

First recall that one standard approach for establishing strong duality in minimum cost flow problems in *finite* networks with *linear* costs is to show that the network Simplex method produces a pair of feasible complementary (and hence optimal) solutions with identical objective function values in the primal and the dual problems (see Proposition 5.8 in [5]). Extension of this Simplex algorithm to countably infinite networks with linear costs is far from straightforward. This was recently achieved by Sharkey and Romeijn [36]. This Simplex method uses a basic feasible solution characterization of extreme points, which relies on the assumption that the problem data are integer. It is known that if the data are not integer then there may exist extreme points that are not basic feasible solutions [12, 32]. Consequently, there is currently no Simplex method and hence no known strong duality results for countably infinite minimum cost network flow problems with linear costs and noninteger data. Sharkey and Romeijn also discussed the difficulties they encountered in applying transversality conditions that were known, at the time, to be sufficient for strong duality for more general CILPs [33, 34].

The Simplex-based strong duality proof for finite networks with linear costs does not generalize to finite networks with convex costs (see Section 9.3 in [5]). In fact, strong duality does not hold, in general, in finite networks with convex costs (see the counterexample in Figure 9.4 in [5]). Fortunately, there is a remarkable and celebrated result, which establishes the absence of a duality gap in these problems (see Proposition 9.4 in [5]). Briefly, this result states that if either the primal or the dual problem is feasible, then there is no duality gap. This result needs a sophisticated proof that was developed by Minty [26] and Rockafellar [29, 30, 31]. For instance, Rockafellar's proof is based on a conceptual algorithm that he called the *fortified descent algorithm*. An extension of this algorithm and hence the corresponding proof to countably infinite networks seems difficult if not impossible. For finite networks, there is a weaker result, which states that if the arc flows are bounded then there is no duality gap (See Exercise 9.1 in [5]). The proof of this result is perhaps simpler, but relies on the feasible differential theorem from Minty [26]. Again, an extension of this theorem to countably infinite networks is not available.

One may be tempted to apply interior point conditions that are known to be sufficient for a zero duality gap in infinite-dimensional convex programs formulated in abstract vector spaces (see, for example, Theorem 3.11.2 in [28]). However, Martin et al. [25] have recently noted a fundamental hurdle in concretely applying such conditions. They formalized this using the concept of a core point, which is an algebraic counterpart of the topological notion of an interior point. They called this hurdle the *Slater conundrum* and summarized it informally as "on the one hand, existence of a core point ensures a zero duality gap (a desirable property), but on the other hand, existence of a core point implies the existence of singular dual functionals (an undesirable property)." Singular dual functionals are difficult to characterize and interpret, and in particular, preclude a *price* interpretation of dual variables that has been central to the importance of duality theory in Operations Research and Economics. For instance, in the case of CILPs, the Slater conundrum means that the dual problem cannot, in general, be written using the transpose of the primal constraint matrix if one intends to use an interior point sufficient condition (see Example 1.2 in [25]).

An approach to circumvent the Slater conundrum in CILPs was recently proposed in [10]. The idea is to first embed the primal and dual variables in appropriately paired sequence spaces, in the spirit of Anderson and Nash [2], so that weak duality and complementary slackness are immediate; and then to revert to the so-called planning horizon method, as in [15, 17, 18, 19, 33, 34], to establish strong duality. The planning horizon method attempts to show that finite-dimensional strong duality is inherited in the limit by the infinite-dimensional problems; importantly, it avoids any reference to interior or core point conditions. This limiting argument derives a convergent subsequence of pairs of infinite-dimensional solutions to the primal and the dual problems from a sequence of pairs of optimal solutions to finite-dimensional primal-dual truncations of the original problems. It then shows that the limit of this subsequence is feasible and complementary in the original problems to conclude strong duality. A similar hybrid approach is currently not available for countably infinite *convex* programs or even for countably infinite separable convex programs. Consequently, duality results for convex minimum cost flow problems on countably infinite networks cannot be recovered as a special case of these problems. We therefore develop such results here by exploiting special structure as described next.

1.1 Our contributions and paper organization

We formally introduce our network flow problem and embed the dual prices in an appropriate sequence space in Section 2 (see our hypotheses H1, H2, and H3 there). We then define the

Lagrangian function. Specifically, our formulation ensures that three key properties hold: (i) the primal objective function is well-defined, finite, and continuous (see our proof of Lemma 2.1); (ii) the Lagrangian function is well-defined, finite, and continuous in flow for each fixed price (Lemmas 2.3 and 2.4); and (iii) the Lagrangian function can be equivalently rewritten in a form that is additively separable over arcs (Lemmas 2.3, 3.3, and 3.4). The dual problem is then formulated at the end of that section. In Section 3, we use the above properties of the Lagrangian function and hence of the dual problem to establish weak duality (Proposition 3.1) and complementary slackness (Proposition 3.9). As a direct consequence of our choice of hypotheses H1, H2, and H3 and of the properties of the Lagrangian function, we are able to replicate proofs of weak duality and complementary slackness from the theory of finite networks. In Section 4, we develop our planning horizon approach to prove a zero duality gap result. We begin that section by recalling a simple construction from [36] that allows us to view the countably infinite network as organized into layers indexed by positive integers n. We then introduce an n-layer truncation of our original network flow problem; this truncation includes a new, artificial node that absorbs any excess flow from or pumps any flow deficit into the truncated network. This artificial node enables us to prove that a truncation of any feasible solution to the original problem is feasible to the truncated problem. Ultimately, this helps in our proof of value convergence (Proposition 4.3): optimal values of the n-layer truncations converge to the optimal value of the original problem as $n \to \infty$. We then write the Lagrangian function for the n-layer network flow problem and prove that it is possible to set the dual price of the artificial node to zero without affecting the Lagrangian value. This allows us to rewrite this Lagrangian function in a manner such that the optimal value of the finite-dimensional dual problem is a lower bound on the optimal value of the original problem. This provides a crucial link between finite-dimensional zero duality gaps, value convergence, and infinite-dimensional weak duality that is needed in our proof that the absence of a duality gap is preserved in the limit (Theorem 4.6). As a corollary of our primal value convergence and zero duality gap results, we establish dual value convergence as well (Corollary 4.7). We then turn to strong duality in Section 5. We prove that when dual variables in the finite-dimensional truncations are bounded independently of the truncation's size, strong duality in the finite-dimensional problems is preserved in the limit. As in existing work on CILPs, our proof derives a convergent subsequence of optimal solutions to the finite-dimensional primal-dual pairs, and shows that the limit of these pairs is feasible and complementary in the original, infinite-dimensional problems. We illustrate our theoretical results on an infinite-horizon shortest path problem in Section 6. We extend these results to countably infinite hypernetworks in Section 7. To achieve this, we make two observations: a countably infinite hypernetwork can also be arranged in layers (Section 7.3), and a finite-layer truncation of the flow problem can be seen as a finite-dimensional monotropic program. Our zero duality gap (Theorem 7.7) and strong duality (Theorem 7.8) results establish that duality properties from finite-dimensional monotropic programs are inherited in the limit. We apply this theory to an infinite-horizon stochastic shortest path problem.

2 Problem formulation

We use the symbol $\triangleq$ to define new mathematical expressions and notations. The symbol # will denote set cardinalities. Let $\mathcal{N}$ denote the countable set of nodes. An arc from node $i \in \mathcal{N}$ to node $j \in \mathcal{N}$ is denoted by the ordered tuple (i, j); let $\mathcal{A}$ denote the countable set of all such arcs. We assume the standard regularity condition that each node has finite in- and out-degree (see [32, 36]). The resulting infinite directed network is denoted by $\mathcal{G} \triangleq (\mathcal{N}, \mathcal{A})$. A real number b_i, called source, is assigned to each node $i \in \mathcal{N}$. If $b_i > 0$, node i is called a supply node and a net flow of b_i needs to be pushed out from this node; if $b_i = 0$, node i is called a transshipment node; finally, if $b_i < 0$, node i is called a demand node and a net flow of $-b_i$ needs to be delivered to this node. The largest amount of flow that can be carried through arc $(i,j) \in \mathcal{A}$ is denoted by $0 \leq u_{ij} < \infty$; these are called flow capacities. We will use $u \triangleq \{u_{ij}\}_{(i,j)\in\mathcal{A}}$ to denote the sequence of capacities u_{ij} and $\mathcal{U} \triangleq \prod_{(i,j)\in\mathcal{A}} [0, u_{ij}]$ to denote the Cartesian product of the intervals $[0, u_{ij}]$ over arcs in $\mathcal{A}$. The assumption of finite flow capacities is standard in the literature on minimum cost flow problems in countably infinite networks (see [32, 36]). More generally, a primal variable-boundedness assumption is either implicitly or explicitly made in essentially all existing results on countably infinite mathematical programs [10, 11, 14, 33, 34]. This assumption can often be shown to hold either without loss of feasibility as in our shortest path problem in Section 6 or without loss of optimality. Arc flows are denoted by x_{ij}, for $(i,j) \in \mathcal{A}$. For each arc $(i,j) \in \mathcal{A}$, let $c_{ij} : [0, u_{ij}] \to \Re$ be a real-valued, continuous, and convex function. The cost of carrying a flow of x_{ij} through arc $(i,j) \in \mathcal{A}$ equals $c_{ij}(x_{ij})$. We make the natural assumptions that $c_{ij}(0) = 0$ and flow costs are nondecreasing over $[0, u_{ij}]$. The former ensures that a zero flow costs nothing and the latter means that larger flows are at least as expensive as smaller ones.

We will use $\Re^{\#\mathcal{A}}$ to denote the set of all sequences $x \triangleq \{x_{ij}\}_{(i,j)\in\mathcal{A}}$ of real numbers indexed by the arcs in $\mathcal{A}$. We assume that

H1. the series $\sum_{(i,j)\in\mathcal{A}} c_{ij}(u_{ij})$ of nonnegative terms is finite.

Hypothesis H1 was inspired by a similar assumption in recent work on duality in CILPs [10]. This type of hypotheses are standard in the literature on countably infinite mathematical programs as they ensure that the objective function in the primal problem is well-defined and finite ([11, 35, 36]). The goal is to find a flow $x \in \Re^{\#\mathcal{A}}$ that satisfies the supply, demand, and transshipment requirements at all nodes; abides by the arc flow capacities; and achieves this at minimum total cost. This minimum cost network flow problem can be formulated as

$$(P) \quad V \triangleq \inf \ C(x) \triangleq \sum_{(i,j)\in\mathcal{A}} c_{ij}(x_{ij}) \tag{1}$$

$$\text{subject to} \quad \sum_{\{j|(i,j)\in\mathcal{A}\}} x_{ij} - \sum_{\{j|(j,i)\in\mathcal{A}\}} x_{ji} = b_i, \ \forall i \in \mathcal{N}, \tag{2}$$

$$x_{ij} \leq u_{ij}, \ \forall (i,j) \in \mathcal{A}, \tag{3}$$

$$x_{ij} \geq 0, \ \forall (i,j) \in \mathcal{A}, \tag{4}$$

$$x \in \Re^{\#\mathcal{A}}. \tag{5}$$

In this paper, we use the product topology of componentwise convergence on $\Re^{\#\mathcal{A}}$. Thus, a sequence $x^n \in \Re^{\#\mathcal{A}}$ converges to a limit $\bar{x} \in \Re^{\#\mathcal{A}}$ if every component x^n_{ij} of this sequence converges, as a sequence of real numbers, to the corresponding component $\bar{x}_{ij}$ of $\bar{x}$. This topology, being a countable product of metric spaces $\Re$, is metrizable (see Theorem 3.36 on page 89 of [1]). Thus, the notions of compactness and sequential compactness coincide (see Theorem 3.28 on page 86 of [1]). We use $F \subset \Re^{\#\mathcal{A}}$ to denote the (possibly empty) feasible region of (P).

Lemma 2.1. *If (P) has a feasible solution, then it has an optimal solution.*

Proof. The feasible region F of (P) is a closed subset (because the linear functions on the left hand sides of the flow balance constraints (2) are continuous) of the product of intervals $\mathcal{U}$. By Tychonoff's theorem (see Theorem 2.61 on page 52 of [1]), $\mathcal{U}$ is compact in the product topology on

$\Re^{\#\mathcal{A}}$. Therefore, the feasible region F is also compact. Moreover, the objective function $C(x)$ in (P) is continuous over $\mathcal{U}$. To establish this, we prove that the series $C(x) = \sum_{(i,j)\in\mathcal{A}} c_{ij}(x_{ij})$ is uniformly convergent over $\mathcal{U}$ and then apply Theorem 9.7 in [3] to conclude continuity. To establish uniform convergence, we simply apply the Weierstrass test [3] to H1 after noting that c_{ij} are nonnegative and nondecreasing. The result then follows by Corollary 2.35 in [1], which implies that a continuous function attains its minimum over a nonempty compact set. □

We now choose a linear subspace to embed dual variables. The idea is to associate a real-valued dual (price) variable y_i with the flow balance constraint (2) for node $i \in \mathcal{N}$. To do this such that the resulting pair of primal-dual problems behaves essentially exactly like its finite-dimensional counterpart, we put some restrictions on the subspace of dual variables. We let $\Re^{\#\mathcal{N}}$ denote the set of all sequences $y \triangleq \{y_i\}_{i\in\mathcal{N}}$ of real numbers indexed by the nodes in $\mathcal{N}$. Moreover, let Y be the subset of $\Re^{\#\mathcal{N}}$ that includes all such sequences y for which

H2. the series $B(y) \triangleq \sum_{i\in\mathcal{N}} |b_i||y_i|$ of nonnegative terms is finite; and

H3. the series $\sum_{i\in\mathcal{N}} |y_i|\big(\sum_{\{j|(i,j)\in\mathcal{A}\}} |u_{ij}| + \sum_{\{j|(j,i)\in\mathcal{A}\}} |u_{ji}|\big)$ of nonnegative terms is also finite.

Again, hypotheses H2 and H3 were inspired by similar assumptions in recent work on duality in CILPs [10]. Hypothesis H2 ensures that a particular infinite series in the Lagrangian function that we later define to write our dual problem converges. Hypothesis H3 plays the role of a transversality condition that Romeijn et al. [33, 34] employed in their work on duality in CILPs. Specifically, H3 implies that a particular infinite series that appears in our Lagrangian function converges. In addition, it allows us to interchange the order of summations in a particular iterated series, which further enables us to equivalently write the Lagrangian function in a convenient form that is additively separable over arcs. Please see Lemmas 2.4 and 3.3 below for further details.

Lemma 2.2. *The subset Y is a linear subspace of $\Re^{\#\mathcal{N}}$.*

Proof. Let $y^1, y^2 \in Y$ and let λ_1, λ_2 be any two real numbers. To show that Y is a linear subspace of $\Re^{\#\mathcal{N}}$, we need to prove that $\lambda_1 y^1 + \lambda_2 y^2$ is in Y (see Section 2.3 on page 14 of [23]). That is, we need to show that $\lambda_1 y^1 + \lambda_2 y^2$ satisfies H2 and H3. To see that it satisfies H2, first note that $\sum_{i\in\mathcal{N}} |b_i||y_i^1|$ and $\sum_{i\in\mathcal{N}} |b_i||y_i^2|$ are both finite. This implies that

$$\sum_{i\in\mathcal{N}} |b_i||\lambda_1 y_i^1 + \lambda_2 y_i^2| \le \sum_{i\in\mathcal{N}} |b_i||\lambda_1||y_i^1| + \sum_{i\in\mathcal{N}} |b_i||\lambda_2||y_i^2| \le |\lambda_1| \sum_{i\in\mathcal{N}} |b_i||y_i^1| + |\lambda_2| \sum_{i\in\mathcal{N}} |b_i||y_i^2| < \infty. \quad (6)$$

To see that $\lambda_1 y^1 + \lambda_2 y^2$ satisfies H3, we proceed as follows. We have,

$$\sum_{i\in\mathcal{N}} |\lambda_1 y^1 + \lambda_2 y^2|\big(\sum_{\{j|(i,j)\in\mathcal{A}\}} |u_{ij}| + \sum_{\{j|(j,i)\in\mathcal{A}\}} |u_{ij}|\big) \quad (7)$$

$$\le \sum_{i\in\mathcal{N}} (|\lambda_1||y^1| + |\lambda_2||y^2|)\big(\sum_{\{j|(i,j)\in\mathcal{A}\}} |u_{ij}| + \sum_{\{j|(j,i)\in\mathcal{A}\}} |u_{ij}|\big) \quad (8)$$

$$\le |\lambda_1| \sum_{i\in\mathcal{N}} |y^1|\big(\sum_{\{j|(i,j)\in\mathcal{A}\}} |u_{ij}| + \sum_{\{j|(j,i)\in\mathcal{A}\}} |u_{ij}|\big)$$

$$+ |\lambda_2| \sum_{i\in\mathcal{N}} |y^2|\big(\sum_{\{j|(i,j)\in\mathcal{A}\}} |u_{ij}| + \sum_{\{j|(j,i)\in\mathcal{A}\}} |u_{ij}|\big) < \infty. \quad (9)$$

Here, the last inequality follows because y^1 and y^2 satisfy H3. □

We next prove that a particular series that we will later use in writing our Lagrangian function converges. We also establish that there is an alternative but equivalent way to write this series; this alternative expression is additively separable over the arcs in $\mathcal{A}$, which later helps in our analysis of complementary slackness.

Lemma 2.3. *For any $x \in \mathcal{U}$ and any $y \in Y$, the series $\sum_{i\in\mathcal{N}} y_i\Big(\sum_{\{j|(j,i)\in\mathcal{A}\}} x_{ji} - \sum_{\{j|(i,j)\in\mathcal{A}\}} x_{ij}\Big)$ converges (absolutely), and it also equals the series $\sum_{e\in\mathcal{A}} -(y_i - y_j)x_{ij}$.*

Proof. For the first claim, we have,

$$\sum_{i\in\mathcal{N}} \Big| y_i\Big(\sum_{\{j|(j,i)\in\mathcal{A}\}} x_{ji} - \sum_{\{j|(i,j)\in\mathcal{A}\}} x_{ij}\Big)\Big| \le \sum_{i\in\mathcal{N}} |y_i|\Big(\Big|\sum_{\{j|(j,i)\in\mathcal{A}\}} x_{ji} - \sum_{\{j|(i,j)\in\mathcal{A}\}} x_{ij}\Big|\Big) \tag{10}$$

$$\le \sum_{i\in\mathcal{N}} |y_i|\Big(\sum_{\{j|(j,i)\in\mathcal{A}\}} |x_{ji}| + \sum_{\{j|(i,j)\in\mathcal{A}\}} |x_{ij}|\Big) \le \sum_{i\in\mathcal{N}} |y_i|\Big(\sum_{\{j|(j,i)\in\mathcal{A}\}} |u_{ji}| + \sum_{\{j|(i,j)\in\mathcal{A}\}} |u_{ij}|\Big) < \infty, \tag{11}$$

where the last inequality follows by H3. This shows that $\sum_{i\in\mathcal{N}} y_i\Big(\sum_{\{j|(j,i)\in\mathcal{A}\}} x_{ji} - \sum_{\{j|(i,j)\in\mathcal{A}\}} x_{ij}\Big)$ converges (absolutely).

We now define the node-arc incidence matrix $A \triangleq \{a_{i,(j,k)}\}$ whose countably infinite set of rows corresponds to the set $\mathcal{N}$ of nodes in $\mathcal{G}$ and whose countably infinite set of columns corresponds to the set $\mathcal{A}$ of arcs in $\mathcal{G}$. Specifically, the entry $a_{i,(j,k)}$ corresponding to node $i \in \mathcal{N}$ and arc $(j,k) \in \mathcal{A}$ is -1 if $j = i$; it is $+1$ if $k = i$; and it is 0 if $j \neq i$ and $k \neq i$. Then, constraint (2) for node $i \in \mathcal{N}$ can be written as $\sum_{(j,k)\in\mathcal{A}} a_{i,(j,k)} x_{jk} = b_i$. Then, the convergent series in question can be written as the iterated series

$$\sum_{i\in\mathcal{N}} y_i\Big(\sum_{\{j|(j,i)\in\mathcal{A}\}} x_{ji} - \sum_{\{j|(i,j)\in\mathcal{A}\}} x_{ij}\Big) = \sum_{i\in\mathcal{N}}\Big(\sum_{\{j|(j,i)\in\mathcal{A}\}} x_{ji}y_i - \sum_{\{j|(i,j)\in\mathcal{A}\}} x_{ij}y_i\Big) \tag{12}$$

$$= \sum_{i\in\mathcal{N}} \sum_{(j,k)\in\mathcal{A}} a_{i,(j,k)} x_{jk} y_i = \sum_{(j,k)\in\mathcal{A}} \sum_{i\in\mathcal{N}} a_{i,(j,k)} x_{jk} y_i = \sum_{(j,k)\in\mathcal{A}} -(y_j - y_k)x_{jk}. \tag{13}$$

Here, the interchange of sums that is employed to write the third equality is allowed by Theorem 8.43 in [3] owing to H3 (also see Lemma 2.1 in [10]). The fourth equality follows from the structure of the node-arc incidence matrix A. This proves the second claim. □

We are now ready to define our Lagrangian function as

$$L(x;y) \triangleq \sum_{(i,j)\in\mathcal{A}} c_{ij}(x_{ij}) + \sum_{i\in\mathcal{N}} b_i y_i + \sum_{i\in\mathcal{N}} y_i\Big(\sum_{\{j|(j,i)\in\mathcal{A}\}} x_{ji} - \sum_{\{j|(i,j)\in\mathcal{A}\}} x_{ij}\Big),\ x \in \mathcal{U},\ y \in Y. \tag{14}$$

Lemma 2.4. *The Lagrangian function is well-defined and finite for every $x \in \mathcal{U}$ and $y \in Y$. Moreover, for each fixed $y \in Y$, the Lagrangian function is continuous over $\mathcal{U}$.*

Proof. For the first claim, note that the term $\sum_{(i,j)\in\mathcal{A}} c_{ij}(x_{ij})$ in the Lagrangian function is finite by H1 because functions c_{ij} are nonnegative and nondecreasing. The term $\sum_{i\in\mathcal{N}} b_i y_i$ is also finite by H2. The third term is also finite by the first claim in Lemma 2.3.

For the second claim, recall that the first series $\sum_{(i,j)\in\mathcal{A}} c_{ij}(x_{ij})$ in the Lagrangian function was shown to be continuous over $x \in \mathcal{U}$ in the proof of Lemma 2.1 in Section 2 using the Weierstrass test. The third series in the Lagrangian function can be similarly shown to be uniformly convergent and hence continuous over $\mathcal{U}$ by the Weierstrass test via H3. This completes the proof. □

We now let

$$\phi(y) \triangleq \min_{x\in\mathcal{U}} \ L(x;y), \ \ y \in Y, \tag{15}$$

and then write the dual of (P) as

$$(D) \ \ W \triangleq \sup_{y\in Y} \ \phi(y). \tag{16}$$

Note that we were able to define the function ϕ in (15) using a minimum instead of an infimum because the function $L(x;y)$ is continuous (by the second claim in Lemma 2.4) over the compact set $\mathcal{U}$ and hence attains its infimum. Also observe that

$$W \geq \phi(0) = \min_{x\in\mathcal{U}} \ L(x;0) = \min_{x\in\mathcal{U}} \sum_{(i,j)\in\mathcal{A}} c_{ij}(x_{ij}) = 0. \tag{17}$$

Here, the last equality follows since $c_{ij}(0) = 0$ and the functions c_{ij} are nondecreasing (and hence nonnegative) over $x_{ij} \in [0, u_{ij}]$. The discussion in this section helps in establishing weak duality and complementary slackness by following the proof techniques from finite networks [5]. We show this in the next section.

3 Weak duality and complementary slackness

Proposition 3.1 (Weak duality)**.** *Suppose $x \in F$ and $y \in Y$. Then, $C(x) \geq \phi(y)$. Note that this implies that $V \geq W$.*

Proof. Similar to the finite-dimensional case hence omitted. □

Corollary 3.2. *Suppose $x^* \in F$ and $y^* \in Y$ are such that $C(x^*) = \phi(y^*)$. Then x^* is optimal to (P) and y^* is optimal to (D).*

Proof. Similar to the finite dimensional case hence omitted. □

Our next goal is to establish a complementary slackness result by following a proof technique that is identical to finite networks. In order to achieve this, we need some simple intermediate lemmas. We first show that the Lagrangian function can be written using an expression that is additively separable over arcs.

Lemma 3.3. *The Lagrangian function originally defined in (14) can be equivalently rewritten as*

$$L(x;y) = \sum_{(i,j)\in\mathcal{A}} \Big(c_{ij}(x_{ij}) - (y_i - y_j)x_{ij}\Big) + \sum_{i\in\mathcal{N}} b_i y_i, \ \ x \in \mathcal{U}, \ \ y \in Y. \tag{18}$$

Proof. We have,

$$\begin{aligned} L(x;y) &= \sum_{(i,j)\in\mathcal{A}} c_{ij}(x_{ij}) + \sum_{i\in\mathcal{N}} b_i y_i + \sum_{i\in\mathcal{N}} y_i \Big(\sum_{\{j|(j,i)\in\mathcal{A}\}} x_{ji} - \sum_{\{j|(i,j)\in\mathcal{A}\}} x_{ij} \Big) && (19) \\ &= \sum_{(i,j)\in\mathcal{A}} c_{ij}(x_{ij}) + \sum_{i\in\mathcal{N}} b_i y_i - \sum_{(i,j)\in\mathcal{A}} (y_i - y_j)x_{ij} && (20) \\ &= \sum_{(i,j)\in\mathcal{A}} \Big(c_{ij}(x_{ij}) - (y_i - y_j)x_{ij}\Big) + \sum_{i\in\mathcal{N}} b_i y_i. \end{aligned}$$

Here, the first equality is simply the original definition of the Lagrangian function as in (14); the second equality follows by the second claim in Lemma 2.3; and the third equality holds by Theorem 8.8 in [3] because both $\sum_{(i,j)\in\mathcal{A}} c_{ij}(x_{ij})$ and $\sum_{(i,j)\in\mathcal{A}} (y_i - y_j)x_{ij}$ converge. □

This lemma also allows us to write the function ϕ using an expression that is additively separable over arcs.

Lemma 3.4. *Problem (15) can be rewritten as*

$$\phi(y) = \sum_{(i,j)\in\mathcal{A}} \phi_{ij}(y_i - y_j) + \sum_{i\in\mathcal{N}} b_i y_i, \tag{21}$$

where functions ϕ_{ij}, for $(i,j) \in \mathcal{A}$, are given by

$$\phi_{ij}(y_i - y_j) \triangleq \min_{x_{ij}\in[0,u_{ij}]} \Big(c_{ij}(x_{ij}) - (y_i - y_j)x_{ij}\Big). \tag{22}$$

Proof. Problem (15) can be rewritten using the equivalent definition of the Lagrangian function from Lemma 3.3 as

$$\phi(y) = \min_{x\in\mathcal{U}} L(x;y) = \min_{x\in\mathcal{U}} \left(\sum_{(i,j)\in\mathcal{A}} \Big(c_{ij}(x_{ij}) - (y_i - y_j)x_{ij}\Big) + \sum_{i\in\mathcal{N}} b_i y_i \right) \tag{23}$$

$$= \min_{\{0\le x_{ij}\le u_{ij} | (i,j)\in\mathcal{A}\}} \left(\sum_{(i,j)\in\mathcal{A}} \Big(c_{ij}(x_{ij}) - (y_i - y_j)x_{ij}\Big) + \sum_{i\in\mathcal{N}} b_i y_i \right) \tag{24}$$

$$= \sum_{(i,j)\in\mathcal{A}} \min_{\{0\le x_{ij}\le u_{ij}\}} \Big(c_{ij}(x_{ij}) - (y_i - y_j)x_{ij}\Big) + \sum_{i\in\mathcal{N}} b_i y_i = \sum_{(i,j)\in\mathcal{A}} \phi_{ij}(y_i - y_j) + \sum_{i\in\mathcal{N}} b_i y_i \tag{25}$$

as required. Here, the penultimate equality is obtained by interchanging the minimum with the series. □

Before defining complementary slackness, we next recall standard definitions and basic properties of the right- and the left-derivative.

Definition 3.5 (Right-derivative)**.** *(From Section 8A of [31].) The right-derivative of the function c_{ij} at $0 \le x_{ij} < u_{ij}$ is given by*

$$c^+_{ij}(x_{ij}) \triangleq \lim_{z\downarrow x_{ij}} \frac{c_{ij}(z) - c_{ij}(x_{ij})}{z - x_{ij}} \tag{26}$$

The right-derivative $c^+_{ij}(u_{ij})$ is defined to equal ∞.

Definition 3.6 (Left-derivative)**.** *(From Section 8A of [31].) The left-derivative of the function c_{ij} at $0 < x_{ij} \le u_{ij}$ is given by*

$$c^-_{ij}(x_{ij}) \triangleq \lim_{z\uparrow x_{ij}} \frac{c_{ij}(z) - c_{ij}(x_{ij})}{z - x_{ij}} \tag{27}$$

The left-derivative $c^-_{ij}(0)$ is defined to equal $-\infty$.

Since functions c_{ij} are convex over $[0, u_{ij}]$, both the right- and the left-derivatives are finite over the open interval $(0, u_{ij})$. In general, for convex functions defined over the interval $[0, u_{ij}]$, the right-derivative can be $-\infty$ at the left endpoint 0; similarly, the left-derivative can be $+\infty$ at the right-endpoint u_{ij} (see Figure 9.1(c) in [5] for an example)[1]. The right-derivative is right-continuous over $[0, u_{ij})$ and the left-derivative is left-continuous over $(0, u_{ij}]$. Both the right- and the left-derivatives are nondecreasing. We refer the reader to Section 8A in [31], Section 24 in [30], and Section 9.1 in [5] for detailed discussions and proofs of these properties.

[1] A flow $x \in \mathcal{U}$ is called regular if $c^-_{ij}(x_{ij}) < \infty$ and $c^+_{ij}(x_{ij}) > -\infty$ for all $(i,j) \in \mathcal{A}$ (see Definition 9.2 in Section 9.1 of [5]).

Definition 3.7 (Complementary flows and prices)**.** *(As in Definition 9.1 in [5].) A flow $x \in \mathcal{U}$ and a price $y \in Y$ are said to be complementary if, for all arcs $(i,j) \in \mathcal{A}$, we have that*

$$c_{ij}^{-}(x_{ij}) \leq y_i - y_j \leq c_{ij}^{+}(x_{ij}). \tag{28}$$

Remark 3.8. *Exactly as in Section 9.3 of [5], $x \in \mathcal{U}$ and $y \in Y$ are complementary if and only if x_{ij} is the minimizer in (22) for all arcs $(i,j) \in \mathcal{A}$.*

Proposition 3.9 (Complementary slackness)**.** *A feasible flow $x^* \in F$ and price $y^* \in Y$ are complementary if and only if x^* and y^* are optimal to (P) and (D), respectively, and the optimal objective values in (P) and (D) are equal.*

Proof. Similar to the finite-dimensional case hence omitted. □

Proposition 3.9 leaves open the question as to whether or not the optimal values V in (P) and W in (D) are equal. We answer it in the affirmative next using the planning horizon approach.

4 Absence of a duality gap

Following Section 2.1 from [36], we assume, without loss of generality, that network $\mathcal{G}$ is layered. That is, the nodes of the network are divided into mutually exclusive sets (called layers) $\mathcal{L}_n$ such that $\bigcup_{n=1}^{\infty} \mathcal{L}_n = \mathcal{N}$. Let $\bar{\mathcal{L}_n} \triangleq \bigcup_{m=1}^{n} \mathcal{L}_m$ denote the set of all nodes in the first n layers. Then, a layering can be constructed by starting with an arbitrary node $\hat{i} \in \mathcal{N}$, setting $\mathcal{L}_1 \triangleq \{\hat{i}\}$, and then recursively defining the other layers as

$$\mathcal{L}_{n+1} \triangleq \{j \in \mathcal{N} \setminus \bar{\mathcal{L}_n} | \exists i \in \bar{\mathcal{L}_n} \text{ with either } (i,j) \in \mathcal{A} \text{ or } (j,i) \in \mathcal{A}\}. \tag{29}$$

This construction implies that each layer $\mathcal{L}_n$ only includes a finite number of nodes (because each node has finite in- and out-degree). It also means that all arcs that connect nodes in $\bar{\mathcal{L}_n}$ with nodes in $\mathcal{N} \setminus \bar{\mathcal{L}_n}$ in fact connect nodes in $\bar{\mathcal{L}_n}$ with nodes in $\mathcal{L}_{n+1}$ — informally, there are no arcs that jump from one layer to a "distant" layer by skipping layer(s) in between. Let

$$\mathcal{A}_n \triangleq \{(i,j) \in \mathcal{A} | \text{ either } i \in \bar{\mathcal{L}_n} \text{ or } j \in \bar{\mathcal{L}_n}\}. \tag{30}$$

We also denote the set of forward arcs from nodes in $\mathcal{L}_n$ to nodes in $\mathcal{L}_{n+1}$ as

$$\mathcal{A}_n^F \triangleq \{(i,j) \in \mathcal{A} | i \in \mathcal{L}_n,\ j \in \mathcal{L}_{n+1}\}. \tag{31}$$

Similarly, we denote the set of backward arcs from nodes in $\mathcal{L}_{n+1}$ to nodes in $\mathcal{L}_n$ as

$$\mathcal{A}_n^B \triangleq \{(j,i) \in \mathcal{A} | j \in \mathcal{L}_{n+1},\ i \in \mathcal{L}_n\}. \tag{32}$$

Finally note that such a layering of $\mathcal{G}$ need not be unique because it depends on the choice of the initial node $\hat{i}$. A schematic of a layered network is shown in Figure 5 below.

We now consider a sequence (P_n) of finite-dimensional truncations of (P). Problem (P_n) includes the nodes in $\bar{\mathcal{L}_n}$ and the arcs in $\mathcal{A}_n$. In addition, it includes an artificial node connected to all forward and backward arcs between layers $\mathcal{L}_n$ and $\mathcal{L}_{n+1}$; a source of $-\sum_{i \in \bar{\mathcal{L}_n}} b_i$ is assigned to this artificial node. Then, problem (P_n) is a finite-dimesnional convex minimum cost network flow problem written as

$$(P_n) \quad V_n \triangleq \inf\ C_n(x) \triangleq \sum_{(i,j) \in \mathcal{A}_n} c_{ij}(x_{ij}) \tag{33}$$

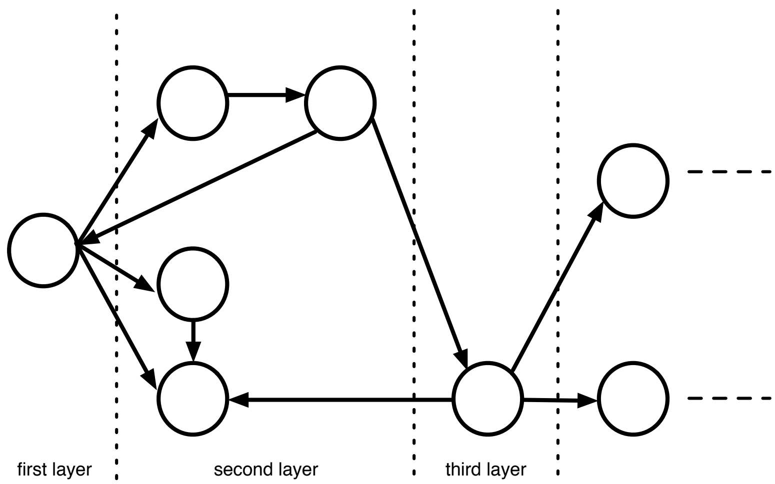


Figure 1: A layered network.

$$\text{subject to} \quad \sum_{\{j|(i,j)\in\mathcal{A}_n\}} x_{ij} - \sum_{\{j|(j,i)\in\mathcal{A}_n\}} x_{ji} = b_i, \ \forall i \in \bar{\mathcal{L}_n}, \tag{34}$$

$$\sum_{(i,j)\in\mathcal{A}_n^F} x_{ij} - \sum_{(j,i)\in\mathcal{A}_n^B} x_{ji} = \sum_{i\in\bar{\mathcal{L}_n}} b_i, \tag{35}$$

$$x_{ij} \le u_{ij}, \ \forall (i,j) \in \mathcal{A}_n, \tag{36}$$

$$x_{ij} \ge 0, \ \forall (i,j) \in \mathcal{A}_n, \tag{37}$$

$$x \in \Re^{\#\mathcal{A}_n}. \tag{38}$$

We first need two simple lemmas.

Lemma 4.1. *Suppose flow x is feasible to (P). Then its truncation to arcs in $\mathcal{A}_n$ is feasible to (P_n).*

Proof. It is clear that this truncation satisfies constraints (34), (36), and (37) in (P_n) as these constraints are identical to the corresponding constraints (2), (3), and (4) in (P). The solution x uses arcs in $\mathcal{A}_n^F$ and $\mathcal{A}_n^B$ to deliver a flow of $\sum_{i\in\bar{\mathcal{L}_n}} b_i$ from nodes in $\mathcal{N}\setminus\bar{\mathcal{L}_n}$ to nodes in $\bar{\mathcal{L}_n}$. In problem (P_n), the artificial node with a source of $-\sum_{i\in\bar{\mathcal{L}_n}} b_i$ is an aggregate representation of all nodes in $\mathcal{N}\setminus\bar{\mathcal{L}_n}$. As a result, the truncation of x must be feasible to constraint (35) as well. □

Lemma 4.2. *If (P_n) has a feasible solution then it has an optimal solution.*

Proof. Holds since the feasible region of (P_n) is compact and the objective is continuous. □

Proposition 4.3 (Primal value convergence)**.** *Suppose (P) has a feasible solution. Then, problems (P_n) have optimal solutions and their optimal values V_n converge to the optimal value V of (P); that is, $\lim_{n\to\infty} V_n = V$.*

Proof. Our proof employs a modification of the argument used in proving Berge's maximum principle [4]. Since (P) has a feasible solution, Lemma 4.1 means that (P_n) has a feasible solution. Lemma 4.2 thus implies that (P_n) has an optimal solution, which we denote by $x^*(n)$. It may, at times, be convenient to view $x^*(n)$ as belonging to $\Re^{\#\mathcal{A}}$ by appending it with a sequence of zeros. We note that an $x^*(n)$ constructed this way belongs to the set $\mathcal{U}$; to avoid introducing additional notation, we will continue to use the same notation for flows in $\Re^{\#\mathcal{A}_n}$ and their extensions formed by appending zeros; the meaning should be clear from context. It now remains to show that $\lim_{n\to\infty} V_n = V$.

Consider a sequence of optimal solutions $\{x^*(n)\} \in \mathcal{U}$ to problems (P_n). Since $\mathcal{U}$ is compact, $\{x^*(n)\}$ has a convergent subsequence; we denote this subsequence by $\{x^*(n_m)\}$ and use $\bar{x} \in \mathcal{U}$ to denote its limit. We show by contradiction that $\bar{x}$ is optimal to (P). So suppose not. This implies that

$$\sum_{(i,j)\in\mathcal{A}} c_{ij}(\hat{x}_{ij}) < \sum_{(i,j)\in\mathcal{A}} c_{ij}(\bar{x}_{ij}). \tag{39}$$

Now we construct another subsequence of flows $\{\tilde{x}(n_m)\} \in \mathcal{U}$ such that

$$\tilde{x}_{ij}(n_m) = \begin{cases} x^*_{ij}(n_m),\ (i,j) \in \mathcal{A}_{n_m},\ \text{and} \\ \hat{x}_{ij},\ (i,j) \in \mathcal{A} \setminus \mathcal{A}_{n_m}. \end{cases} \tag{40}$$

This new subsequence also converges to $\bar{x}_{ij}$. Consequently, inequality (41) can be rewritten as

$$\sum_{(i,j)\in\mathcal{A}} c_{ij}(\hat{x}_{ij}) < \sum_{(i,j)\in\mathcal{A}} c_{ij}(\lim_{m\to\infty} \tilde{x}_{ij}(n_m)). \tag{41}$$

Since function c_{ij} are continuous, this inequality can be further rewritten as

$$\sum_{(i,j)\in\mathcal{A}} c_{ij}(\hat{x}_{ij}) < \sum_{(i,j)\in\mathcal{A}} \lim_{m\to\infty} c_{ij}(\tilde{x}_{ij}(n_m)). \tag{42}$$

Recall that the series $C(x)$ of functions $c_{ij}(x_{ij})$ in the objective function in (P) converges uniformly and hence is continuous over $\mathcal{U}$. This implies that the limit and the infinite sum on the right hand side of (42) can be exchanged (see Theorem 9.7 in [3]). This yields,

$$\sum_{(i,j)\in\mathcal{A}} c_{ij}(\hat{x}_{ij}) < \lim_{m\to\infty} \sum_{(i,j)\in\mathcal{A}} c_{ij}(\tilde{x}_{ij}(n_m)). \tag{43}$$

Then splitting each of the two series into two terms, we obtain,

$$\sum_{(i,j)\in\mathcal{A}_{n_m}} c_{ij}(\hat{x}_{ij}) + \sum_{(i,j)\in\mathcal{A}\setminus\mathcal{A}_{n_m}} c_{ij}(\hat{x}_{ij}) \tag{44}$$

$$< \lim_{m\to\infty} \left[\sum_{(i,j)\in\mathcal{A}_{n_m}} c_{ij}(\tilde{x}_{ij}(n_m)) + \sum_{(i,j)\in\mathcal{A}\setminus\mathcal{A}_{n_m}} c_{ij}(\tilde{x}_{ij}(n_m)) \right]. \tag{45}$$

Then, substituting for $\tilde{x}(n_m)$ from its definition, we get,

$$\sum_{(i,j)\in\mathcal{A}_{n_m}} c_{ij}(\hat{x}_{ij}) + \sum_{(i,j)\in\mathcal{A}\setminus\mathcal{A}_{n_m}} c_{ij}(\hat{x}_{ij}) < \lim_{m\to\infty} \left[\sum_{(i,j)\in\mathcal{A}_{n_m}} c_{ij}(x^*_{ij}(n_m)) + \sum_{(i,j)\in\mathcal{A}\setminus\mathcal{A}_{n_m}} c_{ij}(\hat{x}_{ij}) \right]. \tag{46}$$

This implies that there exists a large enough m, which we call m^*, such that

$$\sum_{(i,j)\in\mathcal{A}_{n_{m^*}}} c_{ij}(\hat{x}_{ij}) + \sum_{(i,j)\in\mathcal{A}\setminus\mathcal{A}_{n_{m^*}}} c_{ij}(\hat{x}_{ij}) < \left[\sum_{(i,j)\in\mathcal{A}_{n_{m^*}}} c_{ij}(x^*_{ij}(n_{m^*})) + \sum_{(i,j)\in\mathcal{A}\setminus\mathcal{A}_{n_{m^*}}} c_{ij}(\hat{x}_{ij}) \right]. \tag{47}$$

In other words,

$$\sum_{(i,j)\in\mathcal{A}_{n_{m^*}}} c_{ij}(\hat{x}_{ij}) < \sum_{(i,j)\in\mathcal{A}_{n_{m^*}}} c_{ij}(x^*_{ij}(n_{m^*})). \tag{48}$$

This contradicts the optimality of $x^*(n_{m^*})$ to $(P_{n_{m^*}})$. This completes the proof that $\bar{x}$ is optimal to (P).

Finally, to show value convergence, we prove that every convergent subsequence of optimal values V_n converges to the same limit and that this limit is V. This would imply that the sequence V_n itself converges to V as required. So let V_{n_m} be such a convergent subsequence and let $\{x^*(n_m)\} \in \mathcal{U}$ be the corresponding subsequence of optimal solutions to (P_{n_m}). This subsequence of optimal solutions has a further convergent subsequence, which we denote by $\{x^*(n_{m_t})\} \in \mathcal{U}$. We denote the limit of this convergent subsequence by $\bar{x} \in \mathcal{U}$. Then, as shown above, $\bar{x}$ is optimal to (P). We thus have,

$$\begin{aligned}
\lim_{t\to\infty} V_{n_{m_t}} &= \lim_{t\to\infty} \sum_{(i,j)\in\mathcal{A}_{n_{m_t}}} c_{ij}(x^*_{ij}(n_{m_t})) && (49)\\
&= \lim_{t\to\infty} \sum_{(i,j)\in\mathcal{A}} c_{ij}(x^*_{ij}(n_{m_t})) \text{ (because the tail of } x^*(n_{m_t}) \text{ is zero and } c_{ij}(0)=0) && (50)\\
&= \sum_{(i,j)\in\mathcal{A}} \lim_{t\to\infty} c_{ij}(x^*_{ij}(n_{m_t})) \text{ (by continuity of the objective function over } \mathcal{U}) && (51)\\
&= \sum_{(i,j)\in\mathcal{A}} c_{ij}(\lim_{t\to\infty} x^*_{ij}(n_{m_t})) \text{ (by continuity of functions } c_{ij}) && (52)\\
&= \sum_{(i,j)\in\mathcal{A}} c_{ij}(\bar{x}_{ij}) = V. && (53)
\end{aligned}$$

That is, $V_{n_{m_t}}$ converges to V as $t \to \infty$. Since $V_{n_{m_t}}$ is a further subsequence of the convergent subsequence V_{n_m}, it must converge to the same limit as V_{n_m}. In other words, $\lim_{m\to\infty} V_{n_m} = V$. Thus, we have shown that every convergent subsequence of V_n converges to V. This completes the proof. □

We next want to write the dual of (P_n). Toward this end, we define $\mathcal{U}_n \triangleq \prod_{(i,j)\in\mathcal{A}_n} [0, u_{ij}]$ as the truncation of $\mathcal{U}$ to $\Re^{\#\mathcal{A}_n}$. For each $x \in \mathcal{U}_n$, $y \in \Re^{\#\bar{\mathcal{L}}_n}$, and $z^n \in \Re$, we first define the Lagrangian function of (P_n) as

$$\begin{aligned}
L_n(x^n;(y^n,z^n)) \triangleq& \sum_{(i,j)\in\mathcal{A}_n} c_{ij}(x^n_{ij}) + \sum_{i\in\bar{\mathcal{L}}_n} b_i y^n_i - z^n \sum_{i\in\bar{\mathcal{L}}_n} b_i + \sum_{i\in\bar{\mathcal{L}}_n} y^n_i \Big(\sum_{\{j|(j,i)\in\mathcal{A}_n\}} x^n_{ji} - \sum_{\{j|(i,j)\in\mathcal{A}_n\}} x^n_{ij} \Big)\\
&+ z^n \Big(\sum_{(j,i)\in\mathcal{A}^B_n} x^n_{ji} - \sum_{(i,j)\in\mathcal{A}^F_n} x^n_{ij} \Big) && (54)\\
=& \sum_{(i,j)\in\mathcal{A}_{n-1}} \Big(c_{ij}(x^n_{ij}) - (y^n_i - y^n_j) \Big) + \sum_{(i,j)\in\mathcal{A}^F_n} \Big(c_{ij}(x^n_{ij}) - (y^n_i - z^n) \Big)\\
&+ \sum_{(j,i)\in\mathcal{A}^B_n} \Big(c_{ij}(x^n_{ij}) - (z^n - y^n_i) \Big) + \sum_{i\in\bar{\mathcal{L}}_n} b_i y^n_i - z^n \sum_{i\in\bar{\mathcal{L}}_n} b_i. && (55)
\end{aligned}$$

We now make an observation about this Lagrangian function, which will simplify our planning horizon approach.

Lemma 4.4. *It is possible to set $z^n = 0$ without affecting the value of the above Lagrangian function.*

Proof. First suppose that $z^n = -\alpha < 0$ for some real number α. Then consider the alternative pair $(\hat{y}^n, \hat{z}^n)$, where $\hat{y}_i^n = y_i^n + \alpha$ for all $i \in \bar{\mathcal{L}}_n$ and $\hat{z}^n = 0$. Then, we have,

$$L_n(x^n; (\hat{y}^n, \hat{z}^n)) = L_n(x^n; (\hat{y}^n, 0)) \tag{56}$$

$$= \sum_{(i,j)\in\mathcal{A}_{n-1}} \Big(c_{ij}(x_{ij}^n) - (\hat{y}_i^n - \hat{y}_j^n)\Big) + \sum_{(i,j)\in\mathcal{A}_n^F} \Big(c_{ij}(x_{ij}^n) - (\hat{y}_i^n)\Big) + \sum_{(j,i)\in\mathcal{A}_n^B} \Big(c_{ij}(x_{ij}^n) - (-\hat{y}_i^n)\Big)$$
$$+ \sum_{i\in\bar{\mathcal{L}}_n} b_i \hat{y}_i^n. \tag{57}$$

By substituting $\hat{y}_i^n = y_i^n + \alpha$, the above simplifies to

$$= \sum_{(i,j)\in\mathcal{A}_{n-1}} \Big(c_{ij}(x_{ij}^n) - (y_i^n - y_j^n)\Big) + \sum_{(i,j)\in\mathcal{A}_n^F} \Big(c_{ij}(x_{ij}^n) - (y_i^n + \alpha)\Big) + \sum_{(j,i)\in\mathcal{A}_n^B} \Big(c_{ij}(x_{ij}^n) - (-\alpha - y_i^n)\Big)$$
$$+ \sum_{i\in\bar{\mathcal{L}}_n} b_i(\alpha + y_i^n). \tag{58}$$

Then, substituting $z^n = -\alpha$, this reduces to

$$= \sum_{(i,j)\in\mathcal{A}_{n-1}} \Big(c_{ij}(x_{ij}^n) - (y_i^n - y_j^n)\Big) + \sum_{(i,j)\in\mathcal{A}_n^F} \Big(c_{ij}(x_{ij}^n) - (y_i^n - z^n)\Big) + \sum_{(j,i)\in\mathcal{A}_n^B} \Big(c_{ij}(x_{ij}^n) - (z^n - y_i^n)\Big)$$
$$+ \sum_{i\in\bar{\mathcal{L}}_n} b_i y_i^n - z^n \sum_{i\in\bar{\mathcal{L}}_n} b_i = L_n(x^n; (y^n, z^n)), \tag{59}$$

as required.

Similarly, if $z^n = \alpha > 0$ for some real number α, we consider the alternative pair $(\hat{y}^n, \hat{z}^n)$, where $\hat{y}_i^n = y_i^n - \alpha$ for all $i \in \bar{\mathcal{L}}_n$ and $\hat{z}^n = 0$. We get,

$$L_n(x^n; (\hat{y}^n, \hat{z}^n)) = L_n(x^n; (\hat{y}^n, 0)) \tag{60}$$

$$= \sum_{(i,j)\in\mathcal{A}_{n-1}} \Big(c_{ij}(x_{ij}^n) - (\hat{y}_i^n - \hat{y}_j^n)\Big) + \sum_{(i,j)\in\mathcal{A}_n^F} \Big(c_{ij}(x_{ij}^n) - (\hat{y}_i^n)\Big) + \sum_{(j,i)\in\mathcal{A}_n^B} \Big(c_{ij}(x_{ij}^n) - (-\hat{y}_i^n)\Big)$$
$$+ \sum_{i\in\bar{\mathcal{L}}_n} b_i \hat{y}_i^n. \tag{61}$$

By substituting $\hat{y}_i^n = y_i^n - \alpha$, the above simplifies to

$$= \sum_{(i,j)\in\mathcal{A}_{n-1}} \Big(c_{ij}(x_{ij}^n) - (y_i^n - y_j^n)\Big) + \sum_{(i,j)\in\mathcal{A}_n^F} \Big(c_{ij}(x_{ij}^n) - (y_i^n - \alpha)\Big) + \sum_{(j,i)\in\mathcal{A}_n^B} \Big(c_{ij}(x_{ij}^n) - (\alpha - y_i^n)\Big)$$
$$+ \sum_{i\in\bar{\mathcal{L}}_n} b_i(y_i^n - \alpha). \tag{62}$$

Then, substituting $z^n = \alpha$, this reduces to

$$= \sum_{(i,j)\in\mathcal{A}_{n-1}} \Big(c_{ij}(x_{ij}^n) - (y_i^n - y_j^n)\Big) + \sum_{(i,j)\in\mathcal{A}_n^F} \Big(c_{ij}(x_{ij}^n) - (y_i^n - z^n)\Big) + \sum_{(j,i)\in\mathcal{A}_n^B} \Big(c_{ij}(x_{ij}^n) - (z^n - y_i^n)\Big)$$
$$+ \sum_{i\in\bar{\mathcal{L}}_n} b_i y_i^n - z^n \sum_{i\in\bar{\mathcal{L}}_n} b_i = L_n(x^n; (y^n, z^n)), \tag{63}$$

as required. □

In view of this lemma, it suffices to only focus on $L_n(x^n;(y^n,0))$, which we simply denote by $L_n(x^n;y^n)$ for brevity and write it here again for emphasis:

$$L_n(x^n;y^n) = \sum_{(i,j)\in\mathcal{A}_n} c_{ij}(x^n_{ij}) + \sum_{i\in\bar{\mathcal{L}_n}} b_i y^n_i + \sum_{i\in\bar{\mathcal{L}_n}} y^n_i \Big(\sum_{\{j|(j,i)\in\mathcal{A}_n\}} x^n_{ji} - \sum_{\{j|(i,j)\in\mathcal{A}_n\}} x^n_{ij} \Big). \tag{64}$$

We also introduce the problem

$$\phi_n(y^n) \triangleq \min_{x^n\in\mathcal{U}_n} \; L_n(x^n;y^n), \; y^n \in \Re^{\#\bar{\mathcal{L}_n}}, \tag{65}$$

and then write the dual problem as

$$(D_n) \; W_n \triangleq \sup_{y^n\in\Re^{\#\bar{\mathcal{L}_n}}} \phi_n(y^n). \tag{66}$$

It is easy to see, similar to (17), that $W_n \geq 0$. We now prove that the optimal values W_n provide lower bounds on W.

Lemma 4.5. *For each n, we have, $W \geq W_n$.*

Proof. For each $y^n \in \Re^{\#\bar{\mathcal{L}_n}}$, we define $\xi(y^n)$ as the corresponding price in Y constructed by appending y^n with a sequence of zeros. Specifically,

$$\xi_i(y^n) \triangleq \begin{cases} y^n_i, \; i \in \bar{\mathcal{L}_n} \\ 0, \; i \in \mathcal{N} \setminus \bar{\mathcal{L}_n}. \end{cases} \tag{67}$$

The set of all such price vectors is defined as

$$\Xi \triangleq \Big\{ \xi(y^n) | y^n \in \Re^{\#\bar{\mathcal{L}_n}} \Big\} \subseteq Y. \tag{68}$$

As a result, we get,

$$W_n = \sup_{y^n\in\Re^{\#\bar{\mathcal{L}_n}}} \phi_n(y^n) = \sup_{y^n\in\Re^{\#\bar{\mathcal{L}_n}}} \Big(\min_{x^n\in\mathcal{U}_n} \; L_n(x^n;y^n) \Big) = \sup_{y^n\in\Re^{\#\bar{\mathcal{L}_n}}} \Big(\min_{x\in\mathcal{U}} \; L(x;\xi(y^n)) \Big) \tag{69}$$

$$= \sup_{\xi\in\Xi} \Big(\min_{x\in\mathcal{U}} \; L(x;\xi) \Big) \leq \sup_{y\in Y} \Big(\min_{x\in\mathcal{U}} \; L(x;y) \Big) = W. \tag{70}$$

Here, the first equality is simply the definition of W_n. The second equality follows from the definition of $\phi_n(y^n)$. The third equality holds because, for every $y^n \in \Re^{\#\bar{\mathcal{L}_n}}$, we have, $\min_{x^n\in\mathcal{U}_n} \; L_n(x^n;y^n) = \min_{x\in\mathcal{U}} \; L(x;\xi(y^n))$ as the components x_{ij} for $(i,j) \in \mathcal{A}\setminus\mathcal{A}_n$ do not appear in this latter minimization. The fourth equality is simply an equivalent rewriting of the supremum in terms of the variable $\xi \in \Xi$. The subsequent inequality holds because $\Xi \subseteq Y$. The last equality is simply the definition of W. □

We are now ready to prove the main result of this section.

Theorem 4.6 (Zero duality gap)**.** *Suppose (P) has a feasible solution. Then $0 \leq W = V < \infty$.*

Proof. The fact that $0 \leq W$ was proven in (17). Since (P) has a feasible solution, it has an optimal solution by Lemma 2.1 and thus $V < \infty$. Now it remains to prove the main claim of this theorem, which is that $W = V$.

Since (P) has a feasible solution, Lemma 4.1 implies that problems (P_n) have feasible solutions for each n. Then Lemma 4.2 implies that problems (P_n) have optimal solutions with optimal values V_n. This also implies, from a well-known result about the absence of a duality gap for finite-dimensional convex minimum cost network flow problems (see Proposition 9.4 in [5]), that $W_n = V_n$. Combining this with our weak duality result that $V \geq W$ as in Proposition 3.1 and our bound that $W \geq W_n$ from Lemma 4.5, we get,

$$V \geq W \geq W_n = V_n, \tag{71}$$

for all n. Since $\lim_{n\to\infty} V_n = V$ by our value convergence in Proposition 4.3, taking a limit on the right hand side of (121) yields $V \geq W \geq V$. This shows that $W = V$. □

Corollary 4.7 (Dual value convergence)**.** *Suppose* (P) *has a feasible solution. Then,* $\lim_{n\to\infty} W_n = W$*.*

Proof. From the proof and the result of Theorem 4.6 and from the result of Proposition 4.3, we get,

$$\lim_{n\to\infty} W_n = \lim_{n\to\infty} V_n = V = W \tag{72}$$

as claimed. □

We view Theorem 4.6 as the countably infinite counterpart of Proposition 9.4 in [5] for finite networks, which implies that if the primal problem has a feasible solution, then there is no duality gap. This theorem, however, does not guarantee that the dual problem has an optimal solution. Specifically, at this point, it is not clear whether or not strong duality holds between (P) and (D). We next provide a condition under which it does.

5 Strong duality

We begin with a simple lemma.

Lemma 5.1. *Suppose* (P) *has a regular feasible solution. Then the finite-dimensional dual problems* (D_n) *have optimal solutions, which we denote by* $y^*(n)$ *in the sequel.*

Proof. Since (P) has a regular feasible solution, its truncation to $\Re^{\mathcal{A}_n}$ is feasible to (P_n) by Lemma 4.1 and it is also regular. Moreover, (P) has an optimal solution by Lemma 2.1. By Lemma 4.2, (P_n) has an optimal solution, say $x^*(n)$. Then, by the strong duality Proposition 9.3 in [5] for finite networks, (D_n) has an optimal solution $y^*(n)$ that is complementary to $x^*(n)$. □

Theorem 5.2 (Strong duality)**.** *Suppose* (P) *has a regular feasible solution (and hence has an optimal solution by Lemma 2.1). Let* $\{x^*(n), y^*(n)\}$*, for* $n = 1, 2, \ldots,$ *be a sequence of pairs of complementary optimal solutions to* (P_n) *and* (D_n) *as in Lemma 5.1. Suppose there exists a sequence* $v \triangleq \{v_i\}_{i\in\mathcal{N}}$ *in* Y *formed by positive real numbers such that* $|y_i^*(n)| \leq v_i$ *for all* $i \in \mathcal{N}$ *and all* n*. Then* (D) *has an optimal solution with* $W = V$*.*

Proof. We define the set $\mathcal{V} \triangleq \prod_{i\in\mathcal{N}} [-v_i, v_i]$, which is in Y because the sequence v is in Y. Since $\mathcal{U}$ and $\mathcal{V}$ are compact by Tychonoff's theorem, the sequence $\{x^*(n), y^*(n)\}$ has a convergent subsequence $\{x^*(n_k), y^*(n_k)\}$. We denote its limit by the pair $(\bar{x}, \bar{y}) \in \mathcal{U} \times \mathcal{V} \subset X \times Y$. We show that $\bar{x}$ and

$\bar{y}$ are complementary in (P) and (D). We also show that $\bar{x}$ is feasible to (P). This will imply, by Proposition 3.9, that $\bar{x}$ and $\bar{y}$ are optimal solutions to problems (P) and (D), respectively, and that $V = W$ as claimed.

To show that $\bar{x}$ and $\bar{y}$ are complementary, we need to prove that

$$c^-_{ij}(\bar{x}_{ij}) \le \bar{y}_i - \bar{y}_j \le c^+_{ij}(\bar{x}_{ij}), \tag{73}$$

for all arcs $(i,j) \in \mathcal{A}$. We do this by contradiction. So suppose not. Then there must exist either an arc $(i,j) \in \mathcal{A}$ such that

$$c^-_{ij}(\bar{x}_{ij}) > \bar{y}_i - \bar{y}_j \tag{74}$$

or an arc $(i,j) \in \mathcal{A}$ such that

$$\bar{y}_i - \bar{y}_j > c^+_{ij}(\bar{x}_{ij}). \tag{75}$$

First suppose that it is the former. We note that $\bar{x}_{ij} \neq 0$ because $c^-_{ij}(0) = -\infty$ and this contradicts (74). We consider two other subcases: (A) there exists a subsequence $x^*(n_{k_t})$ such that $x^*(n_{k_t}) \le \bar{x}_{ij}$ for all t; and (B) there is no such subsequence and hence $x^*(n_k) > \bar{x}_{ij}$ for all k large enough. In subcase (A), $x^*(n_{k_t}) \uparrow \bar{x}_{ij}$ as $t \to \infty$. Then, since the left-derivative is left-continuous over $(0,\infty)$, we have, $c^-_{ij}(\bar{x}_{ij}) = c^-_{ij}\big(\uparrow x^*(n_{k_t})\big) = \lim_{t\to\infty} c^-_{ij}(x^*(n_{k_t}))$. Moreover, $\bar{y}_i - \bar{y}_j = \lim_{t\to\infty}\big(y^*_i(n_{k_t}) - y^*_j(n_{k_t})\big)$. These two observations, when combined with (74) yield that $c^-_{ij}(x^*_{ij}(n_{k_t})) > y^*_i(n_{k_t}) - y^*_j(n_{k_t})$ for some t large enough. But this contradicts the fact that the pair $(x^*(n_{k_t}), y^*(n_{k_t}))$ is complementary for $(P_{n_{k_t}})$ and $(D_{n_{k_t}})$. In subcase (B), $c^-_{ij}(x^*_{ij}(n_k)) \ge c^-_{ij}(\bar{x}_{ij})$ for all k large enough because the left-derivative is nondecreasing. Moreover, $\bar{y}_i - \bar{y}_j = \lim_{k\to\infty}\big(y^*_i(n_k) - y^*_j(n_k)\big)$. So again combining these two observations with (74), we get, $c^-_{ij}(x^*_{ij}(n_k)) > y^*_i(n_k) - y^*_j(n_k)$ for some k large enough. This contradicts the fact that the pair $(x^*(n_k), y^*(n_k))$ is complementary for (P_{n_k}) and (D_{n_k}). Thus, (74) cannot occur.

Now suppose that it is the latter. We note that $\bar{x}_{ij} \neq u_{ij}$ because $c^+_{ij}(u_{ij}) = +\infty$ and this contradicts (75). We again consider two subcases: (A) there exists a subsequence $x^*(n_{k_t})$ such that $x^*(n_{k_t}) \ge \bar{x}_{ij}$ for all t; and (B) there is no such subsequence and hence $x^*(n_k) < \bar{x}_{ij}$ for all k large enough. In subcase (A), $x^*(n_{k_t}) \downarrow \bar{x}_{ij}$ as $t \to \infty$. Then, since the right-derivative is right-continuous over $[0,\infty)$, we have, $c^+_{ij}(\bar{x}_{ij}) = c^+_{ij}\big(\downarrow x^*(n_{k_t})\big) = \lim_{t\to\infty} c^+_{ij}(x^*(n_{k_t}))$. Moreover, $\bar{y}_i - \bar{y}_j = \lim_{t\to\infty}\big(y^*_i(n_{k_t}) - y^*_j(n_{k_t})\big)$. These two observations, when combined with (75) yield that $c^+_{ij}(x^*_{ij}(n_{k_t})) < y^*_i(n_{k_t}) - y^*_j(n_{k_t})$ for some t large enough. But this contradicts the fact that the pair $(x^*(n_{k_t}), y^*(n_{k_t}))$ is complementary for $(P_{n_{k_t}})$ and $(D_{n_{k_t}})$. In subcase (B), $c^+_{ij}(x^*_{ij}(n_k)) \le c^+_{ij}(\bar{x}_{ij})$ for all k large enough because the right-derivative is nondecreasing. Moreover, $\bar{y}_i - \bar{y}_j = \lim_{k\to\infty}\big(y^*_i(n_k) - y^*_j(n_k)\big)$. So again combining these two observations with (75), we get, $c^+_{ij}(x^*_{ij}(n_k)) < y^*_i(n_k) - y^*_j(n_k)$ for some k large enough. This contradicts the fact that the pair $(x^*(n_k), y^*(n_k))$ is complementary for (P_{n_k}) and (D_{n_k}). Thus, (75) cannot occur either.

We now need to show that $\bar{x} \ge 0$ and that it satisfies the flow balance constraints (2) in order to establish its feasibility to (P). The sequence $\bar{x}$ is nonnegative because it is the componentwise limit of nonnegative sequences $\{x^*(n_k)\}$. The sequence $\bar{x}$ satisfies the flow balance constraints (2) because, for any node $i \in \mathcal{N}$, we have,

$$\sum_{\{j|(i,j)\in\mathcal{A}\}} \bar{x}_{ij} - \sum_{\{j|(j,i)\in\mathcal{A}\}} \bar{x}_{ji} = \sum_{\{j|(i,j)\in\mathcal{A}\}} \lim_{k\to\infty} x^*_{ij}(n_k) - \sum_{\{j|(j,i)\in\mathcal{A}\}} \lim_{k\to\infty} x^*_{ji}(n_k) \tag{76}$$

$$= \lim_{k\to\infty} \sum_{\{j|(i,j)\in\mathcal{A}\}} x^*_{ij}(n_k) - \lim_{k\to\infty} \sum_{\{j|(j,i)\in\mathcal{A}\}} x^*_{ji}(n_k) \tag{77}$$

$$= \lim_{k\to\infty} \left(\sum_{\{j|(i,j)\in\mathcal{A}\}} x^*_{ij}(n_k) - \sum_{\{j|(j,i)\in\mathcal{A}\}} x^*_{ji}(n_k) \right) = b_i. \quad (78)$$

Here, the first equality is simply by the definition of $\bar{x}$. The second equality holds because the two sums include only a finite number of terms as node i has finite in- and out-degree. The third equality holds simply because the difference of limits equals the limit of the difference. The last equality holds because $x^*(n_k)$ satisfies (2) for node i for all k large enough. This completes our proof of strong duality. □

6 Application to an infinite-horizon shortest path problem

Here we apply the above results to a convex infinite-horizon shortest path problem whose special case with bounded linear costs is equivalent to infinite-horizon discounted cost dynamic programming (see [11, 14] for a detailed discussion). We will use $\mathbb{N} \triangleq \{1, 2, 3, \ldots\}$ to denote the set of positive integers. Consider a countable collection of nodes $\mathcal{N} \triangleq \bigcup_{n\in\mathbb{N}} S_n$, where $S_1 = \{i_1\}$ is a singleton and S_n are finite sets of nodes i_n. Cardinalities $\#S_n$ of these sets are uniformly bounded above by a positive integer K. In the special case of dynamic programming, n corresponds to discrete time-periods and S_n corresponds to the state-space in period n with $S_1 \triangleq \{i_1\}$ being the initial state. Let $\mathcal{A} \triangleq \bigcup_{n\in\mathbb{N}} \mathcal{A}^F_n$ denote the countable set of arcs, where $\mathcal{A}^F_n$ is the set of arcs that emerge from nodes in S_n and end at nodes in S_{n+1}. In dynamic programming, the arcs in $\mathcal{A}^F_n$ correspond to the set of actions available in various states in S_n. A supply of β_{i_n} is available at node i_n in set S_n, for $n \in \mathbb{N}$; here, β_{i_n} are strictly positive numbers such that $\beta^* \triangleq \sum_{n\in\mathbb{N}} \sum_{i_n\in S_n} \beta_{i_n} < 1$. We assume without loss of generality that each node has at least one emerging arc; in dynamic programming, this means that there is at least one feasible action in each state. Figure 2 illustrates the structure of one such network.

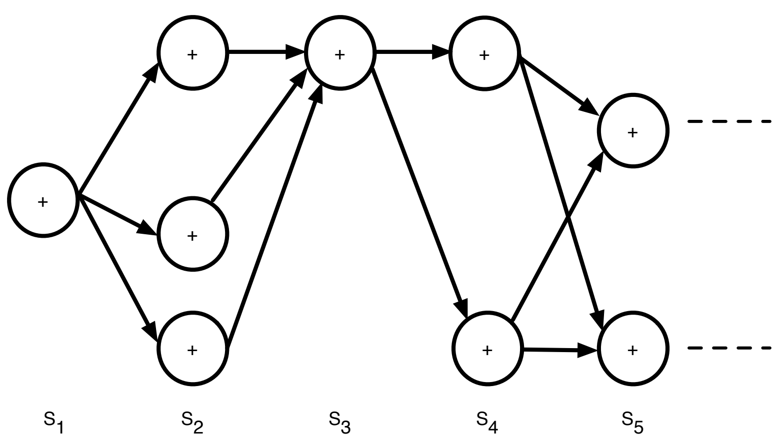


Figure 2: An infinite-horizon shortest path network. All nodes are supply nodes as suggested by the + signs on the nodes.

The goal is to push the flow out of all nodes all the way to a virtual node at infinity at minimum total discounted cost where the single-period discount factor is $0 < \lambda < 1$. The undiscounted cost function for arc $(i, j) \in \mathcal{A}$ is denoted by f_{ij}. Since the flow in any arc cannot be more than the total available supply in this network, arc flows are bounded above by $\beta^* < 1$. We thus choose

$u_{ij} = 1$ for each $(i,j) \in \mathcal{A}$ without loss of feasibility. This convex shortest path problem can be formulated as

$$\text{(shortP) } V \triangleq \inf\ C(x) \triangleq \sum_{n\in\mathbb{N}} \sum_{(i,j)\in\mathcal{A}_n^F} \lambda^{n-1} f_{ij}(x_{ij}) \tag{79}$$

$$\text{subject to} \quad \sum_{\{j|(i_1,j)\in\mathcal{A}_1^F\}} x_{i_1 j} = \beta_{i_1}, \tag{80}$$

$$\sum_{\{j|(i_n,j)\in\mathcal{A}_n^F\}} x_{i_n j} - \sum_{\{j|(j,i_n)\in\mathcal{A}_{n-1}^F\}} x_{j i_n} = \beta_{i_n},\ \forall i_n \in S_n,\ n \in \mathbb{N}\setminus 1, \tag{81}$$

$$x_{ij} \leq 1,\ \forall (i,j) \in \mathcal{A}, \tag{82}$$

$$x_{ij} \geq 0,\ \forall (i,j) \in \mathcal{A}, \tag{83}$$

$$x \in \Re^{\#\mathcal{A}}. \tag{84}$$

As in Section 2, we assume that, for each $(i,j) \in \mathcal{A}_n^F$, the flow cost function $c_{ij}(x_{ij}) \triangleq \lambda^{n-1} f_{ij}(x_{ij})$ is real-valued, continuous, convex, and nondecreasing over $[0,1]$ with $f_{ij}(0) = 0$. Also, hypothesis H1 reduces in this context to requiring that $\sum\limits_{n\in\mathbb{N}} \sum\limits_{(i,j)\in\mathcal{A}_n^F} \lambda^{n-1} f_{ij}(1) < \infty$. A sufficient condition for this is that $\sup\limits_{n\in\mathbb{N}} \left(\max\limits_{(i,j)\in\mathcal{A}_n^F} f_{ij}(1) \right) < \infty$. This shortest path problem has a feasible solution, say the one that splits equally, all incoming flow at each node, among all emerging arcs. In the special case of dynamic programming (where the costs are linear), solving this problem yields an infinite-horizon optimal policy.

We choose the space Y of dual prices as l_1, which is the space of absolutely summable sequences. H2 then holds because

$$\sum_{n\in\mathbb{N}} \sum_{i_n\in S_n} |y_{i_n}||\beta_{i_n}| \leq \sum_{n\in\mathbb{N}} \sum_{i_n\in S_n} |y_{i_n}| < \infty. \tag{85}$$

Here, the first inequality holds because $|\beta_{i_n}| \leq 1$. Let $\mathcal{A}_0^F \triangleq \emptyset$. H3 then holds because

$$\sum_{n\in\mathbb{N}} \sum_{i_n\in S_n} |y_{i_n}| \Bigg(\sum_{\{j|(i_n,j)\in\mathcal{A}_n^F\}} 1 + \sum_{\{j|(j,i_n)\in\mathcal{A}_{n-1}^F\}} 1 \Bigg) \leq 2K \sum_{n\in\mathbb{N}} \sum_{i_n\in S_n} |y_{i_n}| < \infty.$$

Thus the Lagrangian function is given by

$$L(x;y) = \sum_{n\in\mathbb{N}} \sum_{(i,j)\in\mathcal{A}_n^F} \left(\lambda^{n-1} f_{ij}(x_{ij}) - (y_i - y_j)x_{ij} \right) + \sum_{n\in\mathbb{N}} \sum_{i_n\in S_n} \beta_{i_n} y_{i_n},\ 0 \leq x \leq 1;\ y \in l_1. \tag{86}$$

Furthermore, we have,

$$\phi(y) = \min_{0\leq x\leq 1}\ L(x;y). \tag{87}$$

The dual problem then reduces to

$$\text{(shortD) } W = \sup_{y\in l_1}\ \phi(y). \tag{88}$$

Weak duality as in Proposition 3.1 and complementary slackness as in Proposition 3.9 thus hold for (shortP) and (shortD).

To establish the absence of a duality gap and strong duality, we first note that a natural layering of this infinite network is obtained by setting $\mathcal{L}_n = S_n$ for $n \in \mathbb{N}$. Then, an N-layer truncation of the countably infinite network is obtained by considering the nodes in $\bar{\mathcal{L}}_N = \bigcup_{n=1}^{N} S_n$, the arcs in $\mathcal{A}_N = \bigcup_{n=1}^{N} \mathcal{A}_n^F$, and an artificial node with a source of $-\sum_{n=1}^{N} \sum_{i_n \in S_n} \beta_{i_n}$, which is connected to the nodes in S_N via the arcs in $\mathcal{A}_N^F$. A schematic of such a truncation of the infinite-horizon network from Figure 2 is shown in Figure 3.

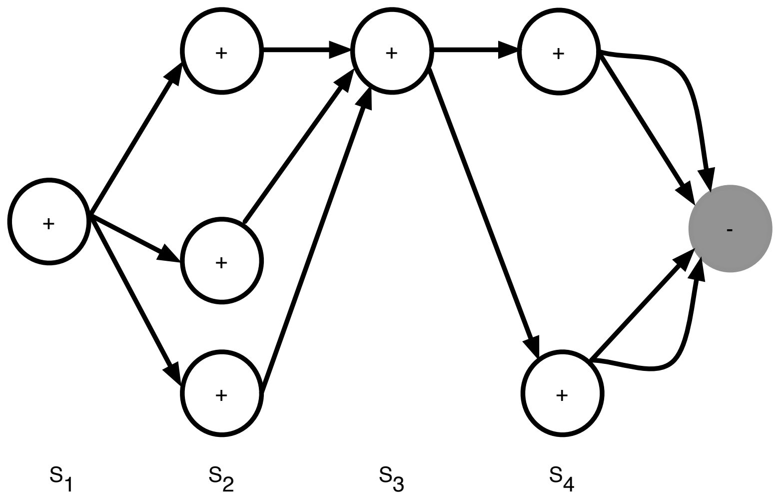


Figure 3: A four-layer truncation of the infinite-horizon shortest path network from Figure 2. The artificial "terminal" node (shown in solid grey) in this truncation is the only demand node as suggested by the $-$ sign on that node.

The resulting N-layer truncation of (shortP) is written as

$$\text{(shortPN)}\ V_N \triangleq \inf\ C_N(x) \triangleq \sum_{n=1}^{N} \sum_{(i,j)\in\mathcal{A}_n^F} \lambda^{n-1} f_{ij}(x_{ij}) \tag{89}$$

$$\text{subject to} \quad \sum_{\{j|(i_1,j)\in\mathcal{A}_1^F\}} x_{i_1 j} = \beta_{i_1}, \tag{90}$$

$$\sum_{\{j|(i_n,j)\in\mathcal{A}_n^F\}} x_{i_n j} - \sum_{\{j|(j,i_n)\in\mathcal{A}_{n-1}^F\}} x_{j i_n} = \beta_{i_n},\ \forall i_n \in S_n,\ n = 2,3,\ldots,N, \tag{91}$$

$$-\sum_{\{j|(j,i)\in\mathcal{A}_N^F\}} x_{ji} = -\sum_{n=1}^{N} \sum_{i_n \in S_n} \beta_{i_n}, \tag{92}$$

$$x_{ij} \leq 1,\ \forall (i,j) \in \mathcal{A}_N, \tag{93}$$

$$x_{ij} \geq 0,\ \forall (i,j) \in \mathcal{A}_N, \tag{94}$$

$$x \in \Re^{\#\mathcal{A}_N}. \tag{95}$$

Then, value convergence as in Proposition 4.3 holds: $\lim_{N\to\infty} V_N = V$.

The Lagrangian function for (shortPN) reduces to

$$L_N(x^N; y^N) = \sum_{n=1}^{N} \sum_{(i,j)\in\mathcal{A}_n} \left(\lambda^{n-1} f_{ij}(x_{ij}^N) - (y_i^N - y_j^N) x_{ij}^N\right) + y_1,\ 0 \leq x^N \leq 1;\ y^N \in \Re^{|\bar{\mathcal{L}}_N|}, \tag{96}$$

with

$$\phi_N(y^N) \triangleq \min_{0 \le x^N \le 1} \; L_N(x^N; y^N), \;\; y^N \in \Re^{|\bar{\mathcal{L}_N}|}. \tag{97}$$

The finite-dimensional dual problem is thus given by

$$\text{(shortDN)} \; W_N = \sup_{y \in \Re^{|\bar{\mathcal{L}}_N|}} \; \phi_N(y^N). \tag{98}$$

The bound $W \ge W_N$ as in Lemma 4.5 holds and there is no duality gap between (shortP) and (shortD) as in Theorem 4.6. Dual value convergence as in Corollary 4.7 also holds: $\lim_{N \to \infty} W_N = W$. A bit more work is needed in exploring if and when strong duality between (shortP) and (shortD) can be established via Theorem 5.2 as described next.

First fix some integer $N \ge 1$. Suppose $x^*(N)$ is an optimal solution to (shortPN). Suppose $y^*(N)$ is its complementary optimal solution to (shortDN) — such a solution exists by Lemma 5.1. Any feasible solution, and in particular the optimal solution $x^*(N)$, to (shortPN) has the property that for each node $i_n \in S_n$, for $n = 1, 2, \dots, N$, there is at least one arc $(i_n, i_{n+1}) \in \mathcal{A}_n^F$ such that $x^*_{i_n i_{n+1}}(N) > 0$. Then, by complementary slackness, we obtain,

$$0 \le c^-_{i_n i_{n+1}}(x^*_{i_n i_{n+1}}(N)) \le y^*_{i_n} - y^*_{i_{n+1}} \le c^+_{i_n i_{n+1}}(x^*_{i_n i_{n+1}}(N)) \le c^+_{i_n i_{n+1}}(\beta^*), \text{ for } n = 1, \dots, N. \tag{99}$$

Here, the lower bound of zero holds because $c^-_{i_n i_{n+1}}(x^*_{i_n i_{n+1}}(N)) \ge 0$ since $c_{i_n i_{n+1}}$ is nondecreasing. Also, the upper bound of $c^+_{i_n i_{n+1}}(\beta^*)$ holds because $x^*_{i_n i_{n+1}} \le \beta^*$ and $c^+_{i_n i_{n+1}}$ is nondecreasing. Then, after setting $y^*_{i_{N+1}} = 0$ as shown in Lemma 4.4 and using the above upper and lower bounds recursively in the order $n = N, N-1, \dots, 1$, we obtain,

$$0 \le y^*_{i_n} \le \sum_{m=n}^{N} c^+_{i_m i_{m+1}}(\beta^*) \le \sum_{m=n}^{\infty} c^+_{i_m i_{m+1}}(\beta^*) \text{ (because the right-derivative is nonnegative)} \tag{100}$$

$$= \lambda^{n-1} \sum_{m=n}^{\infty} \lambda^{m-n} f^+_{i_m i_{m+1}}(\beta^*), \;\; n = 1, 2, \dots, N. \tag{101}$$

By letting $\rho_n \triangleq \sum_{m=n}^{\infty} \lambda^{m-n} \max_{(i_m, i_{m+1}) \in \mathcal{A}_m^F} \left(f^+_{i_m i_{m+1}}(\beta^*) \right)$, we see that if $\sum_{n \in \mathbb{N}} \lambda^{n-1} \rho_n < \infty$, then the hypothesis in Theorem 5.2 would hold with $v_{i_n} = \lambda^{n-1} \rho_n$. This follows because

$$\sum_{n \in \mathbb{N}} \sum_{i_n \in S_n} v_{i_n} = \sum_{n \in \mathbb{N}} \sum_{i_n \in S_n} \lambda^{n-1} \rho_n \le K \sum_{n \in \mathbb{N}} \lambda^{n-1} \rho_n < \infty. \tag{102}$$

A simple sufficient condition for $\sum_{n \in \mathbb{N}} \lambda^{n-1} \rho_n < \infty$ is that ρ_n be uniformly bounded above independently of n, which in turn holds if $\sup_{m \in \mathbb{N}} \left(\max_{(i_m, i_{m+1}) \in \mathcal{A}_m^F} \left(f^+_{i_m i_{m+1}}(\beta^*) \right) \right) < \infty$. In short, strong duality holds if the right-derivatives at β^* are uniformly bounded. In the special case of dynamic programming (where the costs are linear), the right-derivative simply equals the cost per unit flow; thus, in that case, our conclusion reduces to the known result that strong duality holds when costs are bounded [14].

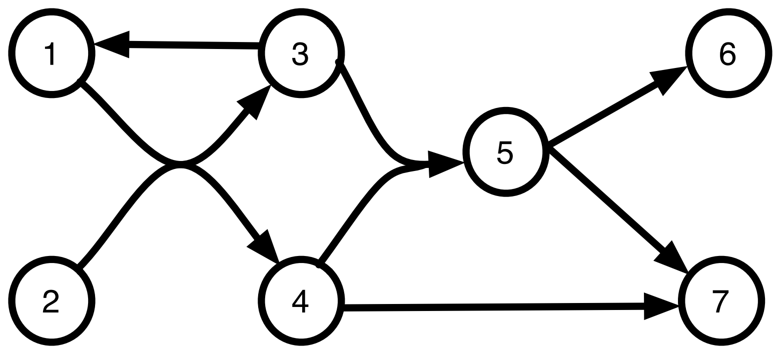


Figure 4: A hypernetwork with seven nodes and six hyperarcs.

7 Extension to infinite hypernetworks

Directed hypernetworks generalize directed networks by allowing for arcs with multiple tails/heads and for flow gains on tails/heads (see Figure 4). Minimum cost flow problems on directed hypernetworks with a finite number of nodes and hyperarcs arise in polyhedral combinatorics; in expert systems; in stationary Markov decision processes (MDPs) with finite state- and action-spaces; and in planning problems such as single- and multi-facility economic lot sizing, lot-size smoothing, warehousing, product-assortment, batch-queuing, capacity expansion, investment consumption, and reservoir-control [6, 20, 27, 38, 39]. Flow problems on hypernetworks with a *countably infinite* number of nodes and hyperarcs arise, for example, in infinite-horizon versions of such planning problems; in nonstationary and countable-state MDPs and generalized MDPs; and in infinite-horizon planning models of Leontief economies [7, 8, 9, 14, 16, 21, 22, 24, 36]. In this section, we extend the above results to countably infinite hypernetworks. We apply these results to a convex infinite-horizon stochastic shortest path problem on a hypernetwork. A special case of this analysis recovers known duality results for nonstationary MDPs.

7.1 Problem formulation

Here, $\mathcal{N}$ denotes the countably infinite set of nodes i and $\mathcal{A}$ denotes the countably infinite set of directed hyperarcs e. The resulting directed infinite hypernetwork is denoted by $\mathcal{G} \triangleq (\mathcal{N}, \mathcal{A})$. Define T_e to be the finite set of tail-nodes of e and H_e to be the finite set of head-nodes of e. We make the standard assumption that node-degrees are finite; that is, each node is connected to a finite number of hyperarcs. Corresponding to all head-nodes $i \in \mathrm{H}_e$ of each hyperarc $e \in \mathcal{A}$, there are nonnegative real numbers $g_i(e)$. These numbers are called head-gains. Similarly, associated with all tail-nodes $i \in \mathrm{T}_e$ of each hyperarc $e \in \mathcal{A}$, there are nonnegative real numbers $\mu_i(e)$. These numbers are called tail-gains. There is a real number b_i associated with each node $i \in \mathcal{N}$. When $b_i > 0$, i is called a supply node; when $b_i < 0$, i is called a demand node; and when $b_i = 0$, i is called a transshipment node. Let nonnegative real numbers u_e denote the flow capacities of hyperarcs $e \in \mathcal{A}$. Let x_e denote the amount of flow through hyperarc e. Let $c_e(x_e)$ be a nondecreasing, continuous and convex function on $[0, u_e]$ that denotes the cost of carrying flow x_e through hyperarc $e \in \mathcal{A}$. Assume that $c_e(0) = 0$; this also implies that flow costs are nonnegative since they are nondecreasing. Our first hypotheses generalizes in this case to **H1:** the series $\sum_{e \in \mathcal{A}} c_e(u_e)$ of nonnegative terms is finite.

The minimum cost flow problem on $\mathcal{G}$ can then be formulated as

$$(P) \quad V \triangleq \inf \; C(x) \triangleq \sum_{e \in \mathcal{A}} c_e(x_e) \tag{103}$$

$$\text{subject to} \sum_{\{e|i\in \mathrm{T}_e\}} \mu_i(e)x_e - \sum_{\{e|i\in \mathrm{H}_e\}} g_i(e)x_e = b_i,\ \forall i \in \mathcal{N}, \tag{104}$$

$$x_e \le u_e,\ \forall e \in \mathcal{A}, \tag{105}$$

$$x_e \ge 0,\ \forall e \in \mathcal{A}, \tag{106}$$

$$x \in \Re^{\#\mathcal{A}}. \tag{107}$$

Again, we use the product topology of componentwise convergence on $\Re^{\#\mathcal{A}}$. Similar to Lemma 2.1, if (P) has a feasible solution then it has an optimal solution.

Hypothesis H2 and H3 from above generalize in this case as follows. Let $\Re^{\#\mathcal{N}}$ denote the set of all sequences $\pi \triangleq \{\pi_i\}_{i\in\mathcal{N}}$ of real numbers indexed by the nodes in $\mathcal{N}$. Moreover, let Y be the subset of $\Re^{\#\mathcal{N}}$ that includes all such sequences π for which **H2:** the series $B(\pi) \triangleq \sum_{i\in\mathcal{N}} |b_i||\pi_i|$ of nonnegative terms is finite; and **H3:** the series $\sum_{i\in\mathcal{N}} |\pi_i|\big(\sum_{\{e|i\in \mathrm{T}_e\}} |\mu_i(e)u_e| + \sum_{\{e|i\in \mathrm{H}_e\}} |g_i(e)u_e|\big)$ of nonnegative terms is also finite. It can be shown similar to Lemma 2.2 that Y is a linear subspace of $\Re^{\#\mathcal{N}}$.

We use $\mathcal{U} \triangleq \prod_{e\in\mathcal{A}} [0, u_e]$ to denote the Cartesian product of the intervals $[0, u_e]$ over arcs in $\mathcal{A}$. Then, for any $x \in \mathcal{U}$ and any $\pi \in Y$, the series $\sum_{i\in\mathcal{N}} \pi_i \Big(\sum_{\{e|i\in \mathrm{H}_e\}} g_i(e)x_e - \sum_{\{e|i\in \mathrm{T}_e\}} \mu_i(e)x_e\Big)$ converges absolutely, and it also equals the series $-\sum_{e\in\mathcal{A}} \Big(\sum_{i\in \mathrm{T}_e} \pi_i\mu_i(e) - \sum_{i\in \mathrm{H}_e} \pi_i g_i(e)\Big)x_e$. This can be shown similar to Lemma 2.3.

Now define the Lagrangian function as

$$L(x;\pi) \triangleq \sum_{e\in\mathcal{A}} c_e(x_e) + \sum_{i\in\mathcal{N}} (b_i - \sum_{\{e|i\in \mathrm{T}_e\}} \mu_i(e)x_e + \sum_{\{e|i\in \mathrm{H}_e\}} g_i(e)x_e)\pi_i. \tag{108}$$

It is then easy to show, similar to Lemma 2.4, that this Lagrangian function is well-defined and finite for every $x \in \mathcal{U}$ and $\pi \in Y$, and that it is continuous over $\mathcal{U}$ for each fixed $\pi \in Y$. Since $\mathcal{U}$ is compact, this implies that $L(x;\pi)$ attains its minimum over $\mathcal{U}$ for each fixed π. Let

$$\phi(\pi) \triangleq \min_{x\in\mathcal{U}}\ L(x;\pi),\ \pi \in Y, \tag{109}$$

and then write the dual of (P) as

$$(D)\quad W \triangleq \sup_{\pi\in Y}\ \phi(\pi). \tag{110}$$

Also, since $W \ge \phi(\pi)$ for all $\pi \in Y$, observe that $W \ge \phi(0) = \min_{x\in\mathcal{U}}\ L(x;0) = \min_{x\in\mathcal{U}} \sum_{e\in\mathcal{A}} c_e(x_e) = 0$.

7.2 Weak duality and complementary slackness

Let F denote the set of flows feasible to (P).

Proposition 7.1 (Weak duality)**.** *Suppose $x \in F$ and $\pi \in Y$. Then, $C(x) \ge \phi(\pi)$. Note that this implies that $V \ge W$.*

Proof. Similar to Proposition 3.1 hence omitted. □

Corollary 7.2. *Suppose $x^* \in F$ and $\pi^* \in Y$ are such that $C(x^*) = \phi(\pi^*)$. Then x^* is optimal to (P) and π^* is optimal to (D).*

Proof. Similar to Corollary 3.2 hence omitted. □

It can be shown, similar to Lemma 3.3, that the Lagrangian function originally defined in (108) can be equivalently rewritten as

$$L(x;\pi) = \sum_{e\in\mathcal{A}} c_e(x_e) + \sum_{i\in\mathcal{N}} b_i\pi_i - \sum_{e\in\mathcal{A}} \Big(\sum_{i\in \mathrm{T}_e} \mu_i(e)\pi_i - \sum_{i\in \mathrm{H}_e} \pi_i g_i(e)\Big) x_e,\ x\in\mathcal{U}, \pi\in Y. \tag{111}$$

Thus, similar to Lemma 3.4, the function ϕ can be rewritten using an expression that is additively separable over hyperarcs and nodes as

$$\phi(\pi) = \sum_{e\in\mathcal{A}} \phi_e(\pi^e) + \sum_{i\in\mathcal{N}} b_i\pi_i. \tag{112}$$

Here, we have used the shorthand π^e to denote the vector of prices of nodes that are connected to hyperarc e, and functions ϕ_e are given by

$$\phi_e(\pi^e) \triangleq \min_{x_e\in[0,u_e]} \ c_e(x_e) - \Big(\sum_{i\in \mathrm{T}_e} \mu_i(e)\pi_i - \sum_{i\in \mathrm{H}_e} \pi_i g_i(e)\Big) x_e. \tag{113}$$

Let c_e^+ and c_e^- denote the right- and left-derivative of the flow cost function for hyperarc $e\in\mathcal{A}$.

Definition 7.3 (Complementary flows and prices)**.** *A flow $x\in\mathcal{U}$ and a price $\pi\in Y$ are said to be complementary if, for all arcs $e\in\mathcal{A}$, we have that*

$$c_e^-(x_e) \le \sum_{i\in T_e} \mu_i(e)\pi_i - \sum_{i\in H_e} g_i(e)\pi_i \le c_e^+(x_e). \tag{114}$$

Remark 7.4. *As in Section 9.3 of [5], $x\in\mathcal{U}$ and $\pi\in Y$ are complementary if and only if x_e is the minimizer in (113) for all arcs $e\in\mathcal{A}$.*

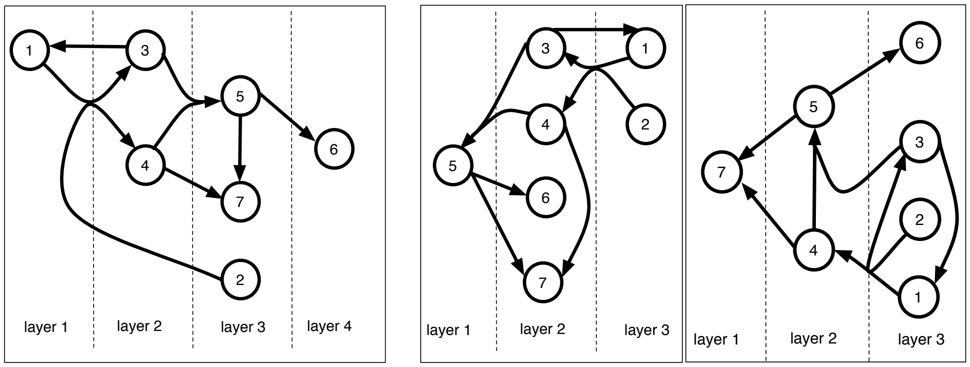
Proposition 7.5 (Complementary slackness)**.** *A feasible flow $x^*\in F$ and price $\pi^*\in Y$ are complementary if and only if x^* and π^* are optimal to (P) and (D), respectively, and the optimal objective values in (P) and (D) are equal.*

Proof. Similar to Proposition 3.9 hence omitted. □

7.3 Absence of a duality gap

We show that nodes of hypernetwork $\mathcal{G}$ can be divided into mutually exclusive layers $\mathcal{L}_n$ such that $\bigcup_{n=1}^{\infty} \mathcal{L}_n = \mathcal{N}$. Toward this end, let $\bar{\mathcal{L}_n} \triangleq \bigcup_{m=1}^{n} \mathcal{L}_m$ denote the set of all nodes in the first n layers. Such a layering can be constructed by starting with an arbitrary node $\hat{i}\in\mathcal{N}$, setting $\mathcal{L}_1 \triangleq \{\hat{i}\}$, and then recursively defining the other layers as $\mathcal{L}_{n+1} \triangleq \{j\in\mathcal{N}\setminus\bar{\mathcal{L}_n} | \exists i\in\bar{\mathcal{L}_n}$ with either $i\in \mathrm{H}_e$ and $j\in \mathrm{T}_e$, or $i\in \mathrm{T}_e$ and $j\in \mathrm{H}_e\}$. Let $\mathcal{A}_n \triangleq \{e\in\mathcal{A} | \exists i\in\bar{\mathcal{L}_n}$ such that either $i\in \mathrm{H}_e$ or $i\in \mathrm{T}_e\}$ denote the set of hyperarcs connected to nodes in the first n layers. A layering of $\mathcal{G}$ need not be unique because it depends on the choice of the initial node $\hat{i}$. Three possible layerings of the hypernetwork from Figure 4 are shown in Figure 5 below to build intuition behind this idea. Now consider a sequence (P_n) of finite-dimensional truncations of (P), where problem (P_n) includes flow balance constraints on nodes in $\bar{\mathcal{L}_n}$ using hyperarcs in $\mathcal{A}_n$. This problem is written as

$$(P_n)\ \ V_n \triangleq \inf\ C_n(x) \triangleq \sum_{e\in\mathcal{A}_n} c_e(x_e) \tag{115}$$



(a) (b) (c)

Figure 5: Three layerings of the hypernetwork from Figure 4. Although this picture depicts a finite hypernetwork, the ideas hold for countable hypernetworks as well.

$$\text{subject to} \sum_{\{e|i\in \mathrm{T}_e\}} \mu_i(e)x_e - \sum_{\{e|i\in \mathrm{H}_e\}} x_e g_i(e) = b_i,\ \forall i \in \bar{\mathcal{L}}_n, \tag{116}$$

$$x_e \le u_e,\ \forall e \in \mathcal{A}_n, \tag{117}$$

$$x_e \ge 0,\ \forall e \in \mathcal{A}_n, \tag{118}$$

$$x \in \Re^{\#\mathcal{A}_n}. \tag{119}$$

Because all constraints in (P_n) also appear in (P), a truncation to $\mathcal{A}_n$ of any flow feasible to (P) is feasible to (P_n). This idea is similar to Lemma 4.1. Moreover, via a standard continuity and compactness argument as in Lemma 4.2, it is easy to see that if (P_n) has a feasible solution then it has an optimal solution.

Proposition 7.6 (Primal value convergence)**.** *Suppose (P) has a feasible solution. Then, problems (P_n) have optimal solutions and their optimal values V_n converge to the optimal value V of (P); that is, $\lim_{n\to\infty} V_n = V$.*

Proof. Similar to Proposition 4.3 hence omitted. □

We next want to write the dual of (P_n). Toward this end, we define $\mathcal{U}_n \triangleq \prod_{e\in\mathcal{A}_n} [0, u_e]$ as the truncation of $\mathcal{U}$ to $\Re^{\#\mathcal{A}_n}$. For each $x \in \mathcal{U}_n$ and $\pi \in \Re^{\#\bar{\mathcal{L}}_n}$, we first define the Lagrangian function of (P_n) as

$$L_n(x^n;\pi^n) \triangleq \sum_{e\in\mathcal{A}_n} c_e(x^n_e) + \sum_{i\in\bar{\mathcal{L}}_n} b_i\pi^n_i + \sum_{i\in\bar{\mathcal{L}}_n} \pi^n_i\Big(\sum_{\{e\in\mathcal{A}_n|i\in\mathrm{H}_e\}} g_i(e)x^n(e) - \sum_{\{e\in\mathcal{A}_n|i\in\mathrm{T}_e\}} \mu_i(e)x^n(e)\Big),$$

and introduce the problem $\phi_n(\pi^n) \triangleq \min_{x^n\in\mathcal{U}_n} L_n(x^n;\pi^n),\ \pi^n \in \Re^{\#\bar{\mathcal{L}}_n}$. Then write the dual of (P_n) as

$$(D_n)\ W_n \triangleq \sup_{\pi^n\in\Re^{\#\bar{\mathcal{L}}_n}} \phi_n(\pi^n). \tag{120}$$

It is easy to see, similar to the earlier observation $W \ge 0$, that $W_n \ge 0$. Moreover, $W \ge W_n$, for each n. The proof of this fact is similar to Lemma 4.5.

Theorem 7.7 (Zero duality gap)**.** *Suppose (P) has a feasible solution. Then $0 \le W = V < \infty$.*

Proof. The fact that $0 \leq W$ was proven in Section 7.1. Since (P) has a feasible solution, it has an optimal solution and thus $V < \infty$. Now it remains to prove that $W = V$. Since (P) has a feasible solution, problems (P_n) have feasible solutions for each n. Thus problems (P_n) have optimal solutions with optimal values V_n. Problems (P_n) belong to the class of finite-dimensional monotropic programs (see Section 9.7 in [5], and [31]). This implies, from a well-known result about the absence of a duality gap for finite-dimensional monotropic programs (see Proposition 9.19 in [5]), that $W_n = V_n$. Combining this with our weak duality result that $V \geq W$ as in Proposition 7.1 and our bound that $W \geq W_n$ as discussed earlier, we get,

$$V \geq W \geq W_n = V_n, \tag{121}$$

for all n. Since $\lim_{n\to\infty} V_n = V$ by value convergence as in Proposition 4.3, taking a limit on the right hand side of (121) yields $V \geq W \geq V$. This shows that $W = V$. □

We end this section with the comment that if (P) has a feasible solution then $\lim_{n\to\infty} W_n = W$. That is, dual value convergence holds if the primal is feasible. The proof is similar to Corollary 4.7.

7.4 Strong duality

We begin with an observation. Suppose (P) has a regular feasible solution. Then the finite-dimensional dual problems (D_n) have optimal solutions. This can be seen as follows. Since (P) has a feasible solution, it has an optimal solution as explained in Section 7.1. The truncation of this solution to $\Re^{\#\mathcal{A}_n}$ is feasible to (P_n) as explained in Section 7.3 and also regular. Moreover, as explained in Section 7.3, (P_n) has an optimal solution, say $x^*(n)$. Then, by Proposition 9.18 in [5] for finite-dimensional monotropic programs, (D_n) has an optimal solution $\pi^*(n)$ that is complementary to $x^*(n)$. This is utilized in our next result.

Theorem 7.8 (Strong duality)**.** *Suppose (P) has a regular feasible solution (and hence has an optimal solution). Let $\{x^*(n), \pi^*(n)\}$, for $n = 1, 2, \ldots$, be a sequence of pairs of complementary optimal solutions to (P_n) and (D_n) as explained above. Suppose there exists a sequence $v \triangleq \{v_i\}_{i\in\mathcal{N}}$ in Y formed by positive real numbers such that $|\pi_i^*(n)| \leq v_i$ for all $i \in \mathcal{N}$ and all n. Then (D) has an optimal solution with $W = V$.*

Proof. Similar to Theorem 5.2 hence omitted. □

7.5 Application to an infinite-horizon stochastic shortest path problem

Here we apply the above results to a convex infinite-horizon stochastic shortest path problem. Its special case with bounded linear costs is equivalent to nonstationary infinite-horizon discounted cost MDPs (see [11, 14] for a detailed discussion). The deterministic special case of this problem yields the shortest path problem in Section 6.

Consider a countable collection of nodes $\mathcal{N} \triangleq \bigcup_{n\in\mathbb{N}} S_n$, where $S_1 = \{i_1\}$ is a singleton and S_n are finite sets of nodes i_n. Cardinalities $\#S_n$ of these sets are uniformly bounded above. In the special case of MDPs, n corresponds to discrete time-periods and S_n corresponds to the state-space in period n with $S_1 \triangleq \{i_1\}$ being the initial state. Each hyperarc has a single tail. Let $\mathcal{A} \triangleq \bigcup_{n\in\mathbb{N}} \mathcal{A}_n^F$ denote the countable set of hyperarcs, where $\mathcal{A}_n^F$ is the set of hyperarcs that emerge from nodes in S_n and end at nodes in S_{n+1}. In MDPs, hyperarcs in $\mathcal{A}_n^F$ correspond to actions in various states in S_n. For hyperarc $e \in \mathcal{A}_{n-1}^F$, for $n = 2, 3, \ldots$, head-gains $g_{i_n}(e)$ are assumed to be nonnegative fractions that add to 1. That is, $0 \leq g_{i_n}(e) \leq 1$ and $\sum_{i_n \in \mathrm{H}_e} g_{i_n}(e) = 1$. In MDPs, these correspond to state-transition probabilities. A supply of β_{i_n} is available at node i_n in set S_n, for $n \in \mathbb{N}$; here, β_{i_n} are strictly positive numbers such that $\beta^* \triangleq \sum_{n \in \mathbb{N}} \sum_{i_n \in S_n} \beta_{i_n} < 1$. We assume without loss of generality that each node has at least one emerging arc; in MDPs, this means that there is at least one feasible action in each state. The goal is to push the flow out of all nodes all the way to a virtual node at infinity at minimum total discounted cost where the single-period discount factor is $0 < \lambda < 1$. The undiscounted flow cost function for hyperarc $e \in \mathcal{A}$ is denoted by f_e. Since the flow in any arc cannot be more than the total available supply in this network, arc flows are bounded above by $\beta^* < 1$. We thus choose $u_e = 1$ for each $e \in \mathcal{A}$ without loss of feasibility. We assume that, for each $e \in \mathcal{A}_n$, the undiscounted flow cost functions $f_e(x_e)$ are continuous, convex, and nondecreasing over $[0, 1]$ with $f_e(0) = 0$.

The resulting convex stochastic shortest path problem can be formulated as

$$
\begin{aligned}
\text{(stochasticshortP)}\ V \triangleq \inf\ C(x) &\triangleq \sum_{n \in \mathbb{N}} \sum_{e \in \mathcal{A}_n} \lambda^{n-1} f_e(x_e) \\
\text{subject to} \quad & \sum_{\{e \in \mathcal{A}_1 | i_1 = \mathrm{T}_e\}} x_e = \beta_{i_1}, \\
& \sum_{\{e \in \mathcal{A}_n | i_n = \mathrm{T}_e\}} x_e - \sum_{\{e \in \mathcal{A}_{n-1} | i_n \in \mathrm{H}_e\}} g_{i_n}(e) x_e = \beta_{i_n},\ \forall i_n \in S_n,\ n \in \mathbb{N} \setminus 1, \\
& x_e \leq 1,\ \forall e \in \mathcal{A}, \\
& x_e \geq 0,\ \forall e \in \mathcal{A}, \\
& x \in \Re^{\#\mathcal{A}}.
\end{aligned}
$$

This problem is a special case of our formulation (P) from Section 7.1. In the special case of nonstationary MDPs (where costs are bounded and linear), solving this problem yields an infinite-horizon optimal policy.

A sufficient condition for hypothesis H1 in this case is that $\sup_{n \in \mathbb{N}} \left(\max_{e \in \mathcal{A}_n} f_e(1) \right) < \infty$; that is, flow costs are uniformly bounded. Choose the space Y of dual prices as l_1, which is the space of absolutely summable sequences. It is then easy to show, similar to the deterministic case, that H2 and H3 hold. Our weak duality and complementary slackness results thus hold.

Problem (stochasticshortP) has a regular feasible solution. For instance, the one that splits equally, all incoming flow at each node, among all outgoing hyperarcs. Absence of a duality gap can then be established by noting that a natural layering of this infinite hypernetwork is obtained by setting $\mathcal{L}_n = S_n$ for $n \in \mathbb{N}$. Finally, a more tedious variation of the algebraic derivation from the deterministic case can be done to show that strong duality holds if the right-derivatives at β^* are uniformly bounded. That is, strong duality holds if $\sup_{m \in \mathbb{N}} \left(\max_{e_m \in \mathcal{A}_m} \left(f^+_{e_m}(\beta^*) \right) \right) < \infty$. In the special case of nonstationary MDPs (where costs are linear), the right-derivative simply equals the cost per unit flow. Thus, in that case, our conclusion reduces to the known result that strong duality holds when costs are bounded [14].

8 Acknowledgments

Funded in part by the National Science Foundation through grants #CMMI 1333260 and 1561918.

References

[1] C D Aliprantis and K C Border. *Infinite-dimensional analysis: a hitchhiker's guide*. Springer-Verlag, Berlin, Germany, 1994.

[2] E J Anderson and P Nash. *Linear programming in infinite-dimensional spaces: theory and applications*. John Wiley and Sons, Chichester, UK, 1987.

[3] T M Apostol. *Mathematical analysis*. Addison Wesley, Reading, MA, USA, 1974.

[4] C Berge. *Topological Spaces*. McMillan, New York, NY, USA, 1963.

[5] D P Bertsekas. *Network Optimization: Continuous and Discrete Models*. Athena Scientific, Belmont, MA, USA, 1998.

[6] R Cambini, G Gallo, and M G Scutella. Flows on hypergraphs. *Mathematical Programming*, 78(2):195–217, 1997.

[7] T Cheevaprawatdomrong, I E Schochetman, R L Smith, and A Garcia. Solution and forecast horizons for infinite-horizon non-homogeneous Markov decision processes. *Mathematics of Operations Research*, 32(1):51–72, 2007.

[8] T Cheevaprawatdomrong and R L Smith. Infinite horizon production scheduling in time-varying systems under stochastic demand. *Operations Research*, 52(1), 2004.

[9] A Ghate. *Markov chains, game theory, and infinite programming: three paradigms for optimization of complex systems*. PhD thesis, University of Michigan, Ann Arbor, MI, USA, 2006.

[10] A Ghate. Circumventing the Slater conundrum in countably infinite linear programs. *European Journal of Operational Research*, 246(3):708–720, 2015.

[11] A Ghate, D Sharma, and R L Smith. A shadow simplex method for infinite linear programs. *Operations Research*, 58(4):865–877, 2010.

[12] A Ghate and R L Smith. Characterizing extreme points as basic feasible solutions in infinite linear programs. *Operations Research Letters*, 37(1):7–10, 2009.

[13] A Ghate and R L Smith. Optimal backlogging over an infinite horizon under time-varying convex production and inventory costs. *Manufacturing and Service Operations Management*, 11(2):362–368, 2009.

[14] A Ghate and R L Smith. A linear programming approach to nonstationary infinite horizon Markov decision processes. *Operations Research*, 61(2):413–425, 2013.

[15] R C Grinold. Infinite horizon programs. *Management Science*, 18(3):157–170, 1971.

[16] R C Grinold. A generalized discrete dynamic programming model. *Management Science*, 20(7):1092–1103, 1974.

[17] R C Grinold. Finite horizon approximations of infinite horizon linear programs. *Mathematical Programming*, 12(1):1–17, 1977.

[18] R C Grinold. Convex infinite horizon programs. *Mathematical Programming*, 25(1):64–82, 1983.

[19] R C Grinold and D S P Hopkins. Duality overlap in infinite linear programs. *Journal of Mathematical Analysis and Applications*, 41:333–335, 1973.

[20] R G Jeroslow, K Martin, R L Rardin, and J Wang. Gainfree Leontief substitution flow problems. *Mathematical Programming*, 57:69–75, 1992.

[21] G J Koehler. A generalized Markov decision process. *RAIRO - Operations Research - Recherche Operationnelle*, 14(4):349–354, 1980.

[22] I Lee, M A Epelman, H E Romeijn, and R L Smith. Simplex algorithm for countable-state discounted Markov decision processes. *Operations Research*, 2017.

[23] D G Luenberger. *Optimization by vector space methods*. John Wiley and Sons, New York, NY, USA, 1969.

[24] A S Manne. Sufficient conditions for optimality in an infinite horizon development plan. *Econometrica*, 38(1):18–38, 1970.

[25] K Martin, C T Ryan, and M Stern. The Slater conundrum: Duality and pricing in infinite dimensional optimization. *SIAM Journal of Optimization*, 26(1):111–138, 2016.

[26] G J Minty. Monotone networks. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 257(1289):194–212, 1960.

[27] P H Ng and D K Wagner. A simplex algorithm for a class of Leontief flow problems. *Mathematical and Computer Modelling*, 24(7):69–75, 1996.

[28] J Ponstein. *Approaches to the theory of optimization*, volume 77 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, England, 1980.

[29] R T Rockafellar. Convex programming and systems of elementary monotonic relations. *Journal of Mathematical Analysis and Applications*, 19(3):543–564, 1967.

[30] R T Rockafellar. *Convex Analysis*. Princeton University Press, Princeton, New Jersey, USA, 1970.

[31] R T Rockafellar. *Network Flows and Monotropic Optimization*. Athena Scientific, Belmont, MA, USA, 1998.

[32] H E Romeijn, D Sharma, and R L Smith. Extreme point solutions for infinite network flow problems. *Networks*, 48(4):209–222, 2006.

[33] H E Romeijn and R L Smith. Shadow prices in infinite dimensional linear programming. *Mathematics of Operations Research*, 23(1):239–256, 1998.

[34] H E Romeijn, R L Smith, and J C Bean. Duality in infinite dimensional linear programming. *Mathematical Programming*, 53(1-3):79–97, 1992.

[35] I E Schochetman and R L Smith. Infinite horizon optimization. *Mathematics of Operations Research*, 14(3):559–574, 1989.

[36] T C Sharkey and H E Romeijn. A simplex algorithm for minimum cost network flow problems in infinite networks. *Networks*, 52(1):14–31, 2008.

[37] R L Smith and R Q Zhang. Infinite horizon production planning in time-varying systems with convex production and inventory costs. *Management Science*, 44(9):1313–1320, 1998.

[38] A F Veinott. Extreme points of Leontief substitution systems. *Linear Algebra and its Applications*, 1(2):181–194, 1968.

[39] A F Veinott. Minimum concave-cost solution of Leontief substitution models of multi-facility inventory systems. *Operations Research*, 17(2):262 – 291, 1969.