A causal decoupling approach to efficient planning for logistics problems with stateful stochastic demand

Diptanil Chaudhuri and Dylan A. Shell

***Abstract*— Future conceptions of agile, just-in-time fabrication, lean and "smart" manufacturing, and a host of allied processes that exploit advanced automation, depend in part on realizing improvements in logistics planning. The present paper hypothesizes that the key to improving flexibility will be the inclusion of sophisticated, time-correlated stochastic models of demand—whether that be demand by end-user consumers directly, or by other down-stream processes. Such dynamic models of demand, unfortunately, can greatly increase the space in which planning occurs when treated, as is common for planning under uncertainty, via the Markov Decision Processes formulation. To tackle this challenge, we identify three aspects that we postulate appear as commonalities in many logistics settings. They lead to an approach for approximate reduction of the planning problem via causal decoupling, which gives a spectrum of solutions where weakening time correlations affords faster optimization. Empirical results on small case studies—in lean manufacturing and commodity routing—show that retaining some limited (but non-zero) amount of temporal structure can provide a useful compromise between quality of the solution obtained and computation required.**

I. INTRODUCTION

Events of the past few years—the global pandemic, the 2021 Suez Canal obstruction—have disclosed the inter-reliance of many elements within large-scale networks that compose modern society. Indeed, supply-chain difficulties, problems with commodity fulfillment, and inventory back-logs are all concrete instances of the brittleness of these systems in the face of the unexpected. To improve future robustness, lean and "smart" manufacturing techniques are being studied to improve process flexibility, smoothen contingency handling, and make logistics more agile. Part of the work also includes designing logistic systems that better meet dynamic demand, when *who* (i.e., *how many* and *where*) wish to consume *what* changes over time.

Aspects such as dynamic-but-unknown future demand require models with some degree of sophistication. Basic models of stochastic demand fall short, being too simplistic: for instance, the addition of i.i.d. noise fails to capture correlations across time, so cannot model seasonal events, nor purchasing fads and fashions. This paper explores dynamic demands via stateful models, as these can help express some valuable time-extended and structural aspects of the process involved. However, the fundamental issue with stateful models is that they increase the size of the planning problem multiplicatively, increasing the computational complexity quickly. Furthermore, logistic problems get still more involved when there are multiple goods.

In this paper, we exploit three postulates to subdue the growth in complexity via what we call 'decoupling.' The first involves causality: state transitions within the demand model reflect aspects of the stochastic process which describe uncertainty. Oftentimes, these are driven entirely by external factors with dynamics being uninfluenced or only weakly affected by happenings within the network. Secondly, when there are multiple sites and different influences on demand at these sites, they can be factored by splitting into separate site-specific models. Thirdly, though one may not know future demand, one can usually determine current demand through suitable instrumentation (say, via market analysis). That is, we assume the current state(s) of the demand model(s) can be ascertained. The observation process is, thus, decoupled and distinct from the other aspects involved.

In many systems, one acts now to meet demand in the future (e.g., producing and transporting items to fill caches and inventories). This involves planning. The preceding postulates imply that the observable state of the system and the state of the demand models may be factored. This allows separate analysis and compression of the demand models, giving a reduced planning problem that is easier to solve. The compression is 'lossy', so the modification is only an approximation of the original instance, but can produce high quality plans in far less time than the full solution. Moreover, in this paper, the degree of compression is adjustable, so one can trade-off greater fidelity in expression of temporal correlations versus time to plan. The empirical results we report show that elimination of time-extended structure gives poor performance, but preserving even a little temporal information improves quality greatly.

The problem of meeting some stochastic demand fits with logistic, transport, and manufacturing problems at different scales, ranging from sparse but geographically extended markets, to internal activities within a single manufacturing facility. Because this paper's contribution and focus is on the underlying problem of planning, and the postulates we have identified accommodate many settings, the gains in performance have potential to apply broadly.

II. RELATED WORKS

At recent ICRAs, rapid developments with automated vehicles have spurred work on the routing of such vehicles in transportation networks [1], [2], [3]. When one thinks of these vehicles as enablers, they then form part of logistic networks within which the automated routing of goods and

Both authors are affiliated with Dept. of Computer Science & Engineering, Texas A&M University, College Station, TX, USA. {diptanil|dshell}@tamu.edu.
This material is based upon work supported by the National Science Foundation under Grants IIS-1849303, IIS-1849249, and IIS-1849291.

commodities becomes feasible. Taking inspiration from our prior work on planning under stocasticity [4], [5], this paper studies the planning problem for an autonomous operations agent capable of routing multiple commodities within a network under stateful stochastic demand.

The literature involving the flow of multiple commodities within a logistic network is vast and has been an important area of study since the first works of Ford and Fulkerson [6], and Hu [7] in the beginning of the 1960s, with a current review appear in [8]. Lately, work has sought to understand the multi-commodity flow problem in the presence of stochasticities. Given a variety of uncertainties present with the supply, demand, and transportation network, the problem of designing a multi-commodity distribution network has been tackled in the recent work including [9], [10], [11], [12]. Other work, like [13], considers a multi-commodity logistics problem with stochastic flow, taking into consideration the effects of transportation time, distance, and the steps involved in the transportation process along with stochastic supply and demand. Ding [14] investigates the multi-commodity flow problem in the presence of uncertain edge cost and edge capacity. Both the above studies pose the problem as a linear programming problem, with [13] employing a multi-objective genetic algorithm and [14] using the Dantzig–Wolfe decomposition method to solve it; the approach we propose here is a dynamic programming–based approach, providing a solution that can adapt/respond to the changes in the demand over time.

The work that most closely resembles the proposed problem is studied in [15], which utilizes a dynamic programming approach to solve the multi-commodity flow problem in the presence of stochastic demand. That study assumes the commodities as reusable and represents the stochastic demand via a random variable at each vertex, as opposed to the proposed work where the demands (for non-reusable commodities) at each vertex are assumed to be generated by a stateful stochastic process.

Finally, research has sought to understand the reliability of networks in the presence of stochastic damage and disruption [16], [17]. Those authors have studied the design of the network model and its reliability, while the work here considers a given network—those studies, thus, can complement our work, providing a way to select a robust network before computing plans which manage its operation.

III. DEFINITIONS AND PROBLEM STATEMENT

Consider the scenario where an autonomous operations agent oversees the supply chain logistics of a wholesale company. The paper formulates the planning problem for the operations agent responsible for routing multiple commodities within a logistic network. We use the generic term 'commodity' to refer to any item, where multiple such items are seen as equivalent to one another in the sense of being exchangeable (e.g., wholesale company selling rice). We study the problem of routing $m \in \mathbb{N}_+$ commodities within a logistic network, with one unit of any commodity being the smallest atomic quantity being considered for storage,

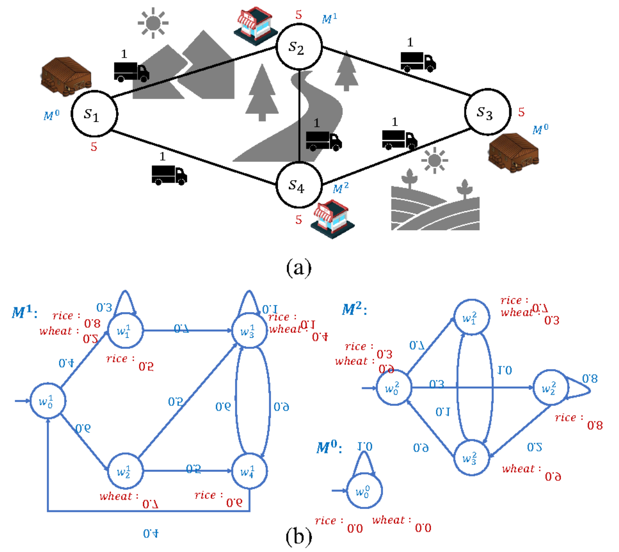

Fig. 1: (a) Logistic network example consisting of 2 storehouse and 2 retail units. The numbers in black represent the edge bandwidth function, while the ones in red represent the storage capacity. (b) The demand models M^1 and M^2. The storage vertices are associated with the trivial demand model M^0.

transportation, or use within the network. Each of the m commodities be indexed by $C = \{1, 2, \ldots, m\}$, called the commodity set. The problem assumes an initial quantity of each commodity within the network as given, and the objective of the problem is to devise a plan for the operations agent to consume the commodities, i.e., driving the quantity of each remaining commodity to zero, in the shortest time.

The planning problem presented here is composed of modular components. The logistic network, being influenced by the agent's decisions, is the causal part; the demand, being unaffected, is the non-causal part. We build up to the precise problem formalization, first identifying and defining the basic objects involved.

A. Logistic Network

The logistic network is modeled as a graph where the vertices of the graph either represent a storage vertex or utilization vertex (viz., locations where the commodities can be sold/consumed); edges of the network represent transportation routes along with their bandwidth. Each vertex of the graph is associated with a utilization model (see Section. III-B), that defines the dynamics of the utilization of commodities within that vertex. Storage vertices are associated with trivial zero utilization model. Hence:

Definition 1 (Logistic Network): A *logistic network* is a 4-tuple, $L = (V, E, S, U)$, where: (1) $V = \{1, 2, \ldots, n\}$ is the non-empty set of vertices; (2) $E \subseteq V \times V$ be the undirected edges of the network; (3) $S : V \to \mathbb{N}_+$ is the vertex storage capacity function, where $S(v)$ is the maximal quantity of all commodities that can be stored at the vertex v at any time; (4) $U : E \to \mathbb{N}_+$ is the edge bandwidth function, where $U(e)$ is the maximum quantity of commodities that can be transported via the edge at any time.

As an example, consider the network for the wholesale example be given by Figure 1a. The execution of the problem begins with a quantity of each commodity within each vertex, and the objective is to route the commodities until all the commodities are consumed. Therefore, an important variable that provides a snapshot of the network is the quantity of each commodity available in each vertex at any given time, we define $\mathcal{S}$ to be a $n \times m$ storage matrix, where s_{ij} is the quantity of commodity j at the i^{th} vertex. Let the set of possible storage matrices be $\mathbb{S}_L = \{\mathcal{S} | s_{ij} \geq 0 \text{ and } \sum_{j \in C} s_{ij} \leq S(i)\}$, for the logistic network L. By $\theta \in \mathbb{S}_L$ denote the zero storage matrix.

B. The Consumer: Demand Model

The dynamics of the consumption/utilization of commodities within each vertex of the logistic network is modeled in this paper via a stateful, discrete-time stochastic process we call the *demand model*. Formally, the demand model is:

Definition 2 (Demand Model): A *demand model* is a 5-tuple $M = (W, C, w_0, \tau, \delta)$, where: (1) W is the non-empty finite state space; (2) C is the set of commodities; (3) $w_0 \in W$, the initial state; (4) $\tau : W \times W \to [0, 1]$, the transition probability function, such that $\sum_{w' \in W} \tau(w, w') = 1$, for any w; (5) $\delta : W \times C \to [0, 1]$ is a demand function where $\delta(w, c)$ is the probability of demand for 1 unit of c in w.

The demand model starts from the state w_0 and progresses from one state to the other according to the probability $\tau(w_t, w_{t+1})$. When the system enters state w_t, demand for a unit of commodity $c \in C$ occurs with probability $\delta(w_t, c)$. Note that at any time step, it is possible that there might be a demand for more than a single type of commodity; however, each commodity will have demand for 1 unit only. When demand for a unit of some commodity arises within a vertex, if one or more units of commodity are present at that vertex, 1 unit is consumed; otherwise, the opportunity is lost and commodity is not consumed. For the wholesale company logistics example, the associated demand models are shown in Figure 1b.

Let the set of all demand models be denoted by $\mathbb{M}$. Define the special, trivial demand model where there are never demands for any commodity to be $M^0 = (W^0, C, w_0^0, \tau^0, \delta^0) \in \mathbb{M}$, with (a) $W^0 = \{w_0^0\}$; (b) $\tau^0(w_0^0, w_0^0) = 1$; and (c) $\delta^0(w_0^0, c) = 0$. The association between the logistic network's vertices and demand models is via a *vertex-demand mapping* function $\mathcal{F} : V \to \mathbb{M}$. The consumer model for storage vertices can be represented by M^0.

C. Routing Policies and Problem Statement

To define the problem, we first need to understand what problem parameters are observable to the operational agent. The quantity of each commodity in each vertex in the logistic network is observable to the agent. For the demand model, the exact demand (realization of the δ function) cannot be observed by the agent; however, the demand model states are observable. This is reasonable when suitable instrumentation (e.g., marketplace analytics, consumer surveys, etc.) is employed. Based on the current observable demand model states, the agent needs to anticipate where commodities might be utilized next and has to take preemptive routing actions for the commodities. Therefore, for a given vertex-demand mapping function $\mathcal{F}$, the states of each demand model corresponding to each vertex of the logistic network is a variable that the agent must keep track of. Let $\mathbb{W}^{\mathcal{F}} = \{\mathbf{w} | \mathbf{w} \in W^1 \times \cdots \times W^n, W^i = \mathcal{F}(i)^{(1)}\}$ be the set of all possible demand state configurations for the function $\mathcal{F}$.

At any time t, the agent's choice is governed by a policy $\pi(\cdot, \cdot)$ based on the states of the demand models corresponding to each vertex of the logistic network ($\mathbf{w}^t \in \mathbb{W}^{\mathcal{F}}$) and the quantity of each commodity available for each vertex ($\mathcal{S}^t \in \mathbb{S}_L$). The agent's policy governs the quantity of each commodity that is routed through each edge of the network. The action space for the agent is denoted in this paper as $\mathbf{A}$, where each action, $\mathbf{a} \in \mathbf{A}$, is a function, $\mathbf{a} : E \times C \to \mathbb{Z}$, such that for $\langle v_s, v_d \rangle \in E$ and $c \in C$, $\mathbf{a}(\langle v_s, v_d \rangle, c) = q$ moves quantity q of commodity c from v_s to v_d if $q \geq 0$; otherwise, if $q < 0$ quantity q of commodity c is moved from v_d to v_s. An action, $\mathbf{a}$, is said to satisfy the edge bandwidth if for all $\langle v_s, v_d \rangle \in E$, the quantity of all the commodities being routed through this edge is less than or equal to the edge bandwidth function value of that edge, that is, $\sum_{c \in C} |\mathbf{a}(\langle v_s, v_d \rangle, c)| \leq U(\langle v_s, v_d \rangle)$.

Let the storage matrix at time t be denoted as $\mathcal{S}^t$. At time t, action $\mathbf{a} \in \mathbf{A}$ is said to be *valid* for $\mathcal{S}$ if, for all $\langle u, v \rangle \in E$ and $c \in C$, we have: (a) the action satisfies the edge bandwidth; (b) the action does not move quantities of commodities more than available from vertex u, that is, $\forall c \in C$, $s_{u,c}^t - \mathbf{a}(\langle u, v \rangle, c) \geq 0$ (c) the action satisfies the vertex storage capacity of the vertex u, that is, $\sum_{c \in C} s_{u,c}^t - \mathbf{a}(\langle u, v \rangle, c) \leq S(u)$ (d) the action does not move quantities of commodities more than available from vertex v, that is, $\forall c \in C$, $s_{v,c}^t + \mathbf{a}(\langle u, v \rangle, c) \geq 0$ (e) the action satisfies the vertex storage capacity of the vertex v, $\sum_{c \in C} s_{v,c}^t + \mathbf{a}(\langle u, v \rangle, c) \leq S(v)$.

Let $\mathbf{1}_{\mathbf{a}}(\cdot)$ be an indicator function, such that, for $\mathbf{a} \in \mathbf{A}$ and $\mathcal{S} \in \mathbb{S}$, $\mathbf{1}_{\mathbf{a}}(\mathcal{S}) = 1$ if action $\mathbf{a}$ is valid for $\mathcal{S}$; otherwise, 0.

The goal is to have all the commodities consumed as quickly as possible. We state the problem formally:

Optimization Problem: Logistics with Demand (LWD)

Given: Commodity set C, a logistic network $L = (V, E, S, U)$, set of demand models $\mathbb{M}$, a vertex-demand mapping function $\mathcal{F}$, and an initial storage matrix $\mathcal{S}^0 \in \mathbb{S}_L$.

Output: A policy $\pi^* : \mathbb{W}^{\mathcal{F}} \times \mathbb{S}_L \to \mathbf{A}$ of valid actions that minimizes the expected time for all commodities to be consumed.

IV. FORMULATION OF LWD AS AN MDP

To solve the optimization problem, we construct a specific Markov Decision Problem, called the **LWD** MDP.

However, before we formally define the **LWD** MDP, we need to define some preliminary functions. We start by define the *vertex-demand transition* function, $T^{\mathcal{F}} : \mathbb{W}^{\mathcal{F}} \times \mathbb{W}^{\mathcal{F}} \to [0, 1]$, such that, for $\mathbf{w} = (w^1, \ldots, w^n)$, $\mathbf{y} = (y^1, \ldots, y^n) \in \mathbb{W}^{\mathcal{F}}$, we have $T^{\mathcal{F}}(\mathbf{w}, \mathbf{y}) = \prod_{v \in V} \tau^v(w^v, y^v)$, where $\tau^v =$

$\mathcal{F}(v)^{(4)}$, specifying the transition probability from one sequence of demand states in the logistic network to another sequence. Next we define the *transport* partial function $\xi : \mathbb{S}_L \times \mathbf{A} \hookrightarrow \mathbb{S}_L$, such that, for $\mathcal{S}, \mathcal{S}^- \in \mathbb{S}_L$ and $\mathbf{a} \in \mathbf{A}$, if $\mathbf{a}$ is valid: $\xi(\mathcal{S}, \mathbf{a}) = \mathcal{S}^-$ if $\forall s_{v,c} \in \mathcal{S}$ and $\forall s'_{v,c} \in \mathcal{S}^-$, we have $s'_{v,c} = s_{v,c} - \sum_{v' \in V} \mathbf{a}(\langle v, v' \rangle, c)$. That is, this function returns the storage matrix that results from the routing actions. Lastly, we define the *consumption* function, $\Delta^{\mathcal{F}} : \mathbb{S}_L \times \mathbb{W}^{\mathcal{F}} \times \mathbb{S}_L \to [0,1]$, to be a function of $\mathcal{S}^-, \mathcal{S}^+ \in \mathbb{S}_L$ and $\mathbf{w} = (w^1, \ldots, w^n) \in \mathbb{W}^{\mathcal{F}}$, such that $\Delta^{\mathcal{F}}(\mathcal{S}^-, \mathbf{w}, \mathcal{S}^+) = \prod_{(v,c) \in V \times C} \varphi(s^-_{v,c}, w^v, s^+_{v,c})$ where $\varphi(s^-_{v,c}, w^v, s^+_{v,c}) = \delta(w^v, c)$ if $s^+_{v,c} - s^-_{v,c} = 1$; $1 - \delta(w^v, c)$, if $s^+_{v,c} - s^-_{v,c} = 0$; and 0 otherwise. That is, this function determines the probability that some commodities are consumed from one storage matrix and result in the other.

With the requisite functions given, we are now ready:

Definition 3 (LWD MDP): Given an initial storage $\mathcal{S}^0 \in \mathbb{S}_L$, the logistic network $L = (V, E, S, U)$, a vertex-demand mapping function $\mathcal{F}_D$, the set of demand models $\mathcal{M} = \{M | M \in \mathbb{M}$ and $\exists v \in V, \mathcal{F}_D(v) = M\}$, the **LWD** MDP is constructed as $\mathbb{X}_{\mathbb{S}^0, \mathcal{F}_D, L} = (X, x_0, \mathbf{A}, P, X_G, J)$, where (1) $X \subseteq \mathbb{W}^{\mathcal{F}_D} \times \mathbb{S}_L$, the set of states; (2) $x_0 = (\mathbf{w_0}, \mathcal{S}^0)$, such that $\mathbf{w_0} = (w_0^1, w_0^2, \ldots, w_0^n)$, where $w_0^i \in \mathcal{F}_D(i)$, the initial state; (3) $\mathbf{A}$, the action space; (4) $P : X \times \mathbf{A} \times X \to [0,1]$, the transition probability function, such that, for $(\mathbf{w}, \mathcal{S}), (\mathbf{w}', \mathcal{S}') \in X$ and $\mathbf{a} \in \mathbf{A}$, $P((\mathbf{w}, \mathcal{S}), \mathbf{a}, (\mathbf{w}', \mathcal{S}')) = \mathbf{1_a}(\mathcal{S}) T^{\mathcal{F}_D}(\mathbf{w}, \mathbf{w}') \Delta^{\mathcal{F}_D}(\xi(\mathcal{S}, \mathbf{a}), \mathbf{w}', \mathcal{S}')$; (5) $X_G = \mathbb{W}^{\mathcal{F}} \times \{\theta\} \subseteq X$ is the set of goal states; (6) $J : X \times \mathbf{A} \to \mathbb{R}_{\geq 0}$ is the cost function, so, for $x \in X$ and $\mathbf{a} \in \mathbf{A}$, $J(x, \mathbf{a}) = 1$ if $x \notin X_G$; otherwise 0.

An optimal policy for product MDP, $\pi^* : X \to \mathbf{A}$ provides the routing policy. The full product would directly construct **LWD** MDP $\mathbb{X}_{\mathbb{S}^0, \mathcal{F}_D, L}$. Then, the policy can be obtained by using standard solution techniques (e.g. value iteration [18]).

V. SOLUTIONS VIA APPROXIMATION

As just formulated, the planning problem comprises of individual modular components, consisting of demand models (non-causal) and a logistic network (causal). The number of states in each demand model increases the size of the planning problem multiplicatively via the product in Definition 3. Rather than use the vertex-demand mapping function $\mathcal{F}_D$ directly, as the full solution does, the non-causality provides an opportunity to analyze and simplify the demand models independently. As the demand is not contingent on the actions, the dynamics of the demand do not influence the iterative policy update and thus can be analyzed beforehand. Two such approaches follow next.

A. Fundamental Matrix Analysis

Our second approach to the problem, which we call the *FMA approach*, takes inspiration from [15], and uses matrix analysis on each demand model to collapse all the states into a single state. This collapse of states destroys the temporal structure of the original demand models and reduces the dynamics of the consumer demand for every commodity into a Bernoulli random variable.

To solve the problem using this approach we need to first define two matrices. First, for demand model $M = (W, C, w_0, \tau, \delta) \in \mathbb{M}$, and commodity $c \in C$, we solve for the $|W| \times 1$ matrix, $\psi^{(M,c)}$, where $\psi_i^{(M,c)}$, the i^{th} element, is the expected number of steps before demand for commodity c will occur in the demand model if starting at the i^{th} state of the demand model. We can calculate the matrix $\psi^{(M,c)}$, by following the procedure described next.

Given a commodity $c \in C$ and a demand model $M = (W, C, w_0, \tau, \delta)$, we construct a new absorbing Markov chain (W', τ', w_0). To form this, first, we define a new set of states $W' = W \cup \{w^c_{\text{ABS}}\}$. For all $w, w' \in W$, such that, $\tau(w, w') > 0$ if $\delta(w', c) > 0$, we add the transitions $\tau'(w, w') = \tau(w, w')(1 - \delta(w', c))$ and $\tau'(w, w^c_{\text{ABS}}) = \tau(w, w')\delta(w', c)$ to the new absorbing Markov chain and if $\delta(w', c) = 0$, we add the transition $\tau'(w, w') = \tau(w, w')$. Performing fundamental matrix analysis [19] on the newly generated absorbing Markov chain yields matrix $\psi^{(M,c)}$.

The second matrix needed for this approach, ϕ^M, a $|W| \times 1$ matrix, is the stationary distribution. For a demand model $M = (W, C, w_0, \tau, \delta)$, if the Markov chain (W, w_0, τ) is non-absorbing the stationary distribution matrix can be calculated by solving the equation $\phi^M = \phi^M \tau$. Otherwise, if (W, w_0, τ) is absorbing the stationary distribution matrix $\phi^M = [q, q, \ldots, q]^T$, where $q = (|W|)^{-1}$.

The FMA approach replaces the original demand model with one having only a single state. Thus, define new set $\mathcal{M}^1 = \{M^1 = (W^1, C, w_0^1, \tau^1, \delta^1) \in \mathbb{M} \,|\, W^1 = \{w_0^1\}$ and $\tau^0(w_0^1, w_0^1) = 1\}$. Notice, $M^0 \in \mathcal{M}^1$.

By $\mathcal{F}_M : \mathbb{M} \to \mathcal{M}^1$ denote the *fundamental matrix reduction function*, where, for $M = (W, C, w_0, \tau, \delta) \in \mathbb{M}$, $\mathcal{F}_1(M) = (W^1, C, w_0^1, \tau^1, \delta^1)$, and for all $c \in C$, $\delta^1(w_0^1, c) = \left(\sum_{w \in W} \phi_w^M \psi_w^{M,c}\right)^{-1}$ Note, $\mathcal{F}_1(M^0) = M^0$.

The solution to **LWD** using this approach can be generated by constructing the **LWD** MDP $\mathbb{X}_{\mathbb{S}^0, \mathcal{F}_{M \circ D}, L}$, where for $v \in V$, $\mathcal{F}_{M \circ D} = \mathcal{F}_M(\mathcal{F}_D(v))$. And then solving that reduced MDP using some standard technique.

B. Model Reduction by Collapsing State Pairs

The two approaches just seen can be considered as two extremes: the first without any reductions, while the second reducing the whole demand model to a single state. In this section, we will devise a reduction function that gives a spectrum of approximations in-between, as it can be applied to the original demand functions iteratively to reduce the number of states one at a time.

For a given demand model $M = (W, C, w_0, \tau, \delta) \in \mathbb{M}$, an intuitive approach to reduce the number of states is to merge the two states within W that are most similar to each other. Each state w of the demand model is associated with two distributions: (a) distribution over the states of the demand model, given by the transition function $\tau(w, \cdot)$, and (b) joint probability distribution over every commodity derived from the demand function $\delta(w, \cdot)$. Therefore to quantify the similarity between two states, we would need to quantify the similarity between their distributions for both (a) and (b). We introduce a modified formulation of the Hellinger distance, with a parameter α, to quantify the similarity between two states of the same demand model. The parameter α acts as weights for the Hellinger distance of the two distributions, assigning preference of one over the other. For any two states $w, w' \in W$, the modified Hellinger distance, $\mathbf{H}_\alpha(w, w') = \alpha H_\tau(w, w') + (1-\alpha) H_\delta(w, w')$, where $H_\tau(w, w')$ is the Hellinger distance [20] between the two distributions $\tau(w, \cdot)$ and $\tau(w', \cdot)$, and $H_\delta(w, w')$ is the Hellinger distance between the two joint distribution derived from $\delta(w, \cdot)$ and $\delta(w', \cdot)$. Since the Hellinger distance is symmetric, we have, $\mathbf{H}_\alpha(w, w') = \mathbf{H}_\alpha(w', w)$. The two most similar states in terms of Hellinger distance are given as $W_\alpha^M = \arg\min_{w, w' \in W} \mathbf{H}_\alpha(w, w')$.

Using this, we can give the *Hellinger reduction* function $\mathcal{F}_H : \mathbb{M} \times [0,1] \to \mathbb{M}$, so that, for $M = (W, C, w_0, \tau, \delta) \in \mathbb{M}$ and $\alpha \in [0,1]$, $\mathcal{F}_H(M, \alpha) = M$, if $M \in \mathcal{M}^1$, otherwise, $\mathcal{F}_H(M, \alpha) = M' = (W', C, w_0', \tau', \delta')$, where: (1) $W' = \{w_{\text{new}}\} \cup W \setminus W_\alpha^M$ is the non-empty state space; (2) C is the set of all commodities; (3) $w_0' \in W'$ is the initial state, such that, if $w_0 \in W_\alpha^M$, $w_0' = w_{\text{new}}$, otherwise take $w_0' = w_0$; (4) $\tau' : W' \times W' \to [0,1]$ is the transition probability function, such that for $w, w', w_{\text{new}} \in W'$, and $\{w_a, w_b\} \in W_\alpha^M$, (a) if $w' \neq w \neq w_{\text{new}}$, $\tau'(w, w') = \eta_w \tau(w, w')$, (b) if $w = w_{\text{new}}$ and $w' \neq w_{\text{new}}$, $\tau'(w, w') = \eta_w \left(\tau(w_a, w') + \tau(w_b, w')\right)$, (c) if $w \neq w_{\text{new}}$ and $w' = w_{\text{new}}$, $\tau'(w, w') = \eta_w \left(\tau(w, w_a) + \tau(w, w_b)\right)$, (d) if $w = w' = w_{\text{new}}$, $\tau'(w, w') = \eta_w(\tau(w_a, w_a) + \tau(w_a, w_b)$ $+\tau(w_b, w_a) + \tau(w_b, w_b))$, where η_w is a normalizing factor such that $\sum_{w' \in W'} \tau'(w, w') = 1$; (5) $\delta' : W' \times C \to [0,1]$ is a demand function, such that, for $w \in W'$ $\delta'(w, c) = \delta(w, c)$ if $w \neq w_{\text{new}}$; otherwise $\delta'(w, c) = \frac{1}{\phi_{w_a}^M + \phi_{w_b}^M} \left(\phi_{w_a}^M \delta(w_a, c) + \phi_{w_b}^M \delta(w_b, c)\right)$, where $\{w_a, w_b\} \in W_\alpha^M$ and ϕ^M is the $|W| \times 1$ stationary distribution matrix.

For $\rho = (\rho_1, \ldots, \rho_n) \in \mathbb{N}_+^n$, and $v \in V$, let us write $\mathcal{F}_{\rho \circ D} = \mathcal{F}_H^{\rho_v}(\mathcal{F}_D(v))$, where $\mathcal{F}_H^{\rho_v}(\cdot)$ signifies application of the Hellinger distance function ρ_v times iteratively.

Then solution to **Lwd**, via this approach, is obtained by reducing the original demand model by ρ, and constructing the **Lwd** MDP $\mathbb{X}_{\mathbb{S}^0, \mathcal{F}_{\rho \circ D}, L}$. This MDP is then solved.

One thing to note here is that these are a few of the many approaches that can be used to pre-analyze the demand model in order to generate efficient approximate solutions.

VI. CASE STUDIES

We turn now to examine two example scenarios and present simulation results using a Python implementation of the algorithms, executed on a Windows 11 computer with a 2.90GHz CPU.

A. Routing grain: rice and wheat

For the first scenario, we revisit the wholesale company example with two commodities: rice, and wheat. The logistic network for this problem and their associated demand models are shown in Figure 1.

We solve the problem using approaches presented, each generating its own policy. First, we generate the policy by solving for the MDP considering the full demand models without any reduction. The second, third, and fourth solutions generate the MDP by first using the Hellinger distance reduction technique to collapse the number of states in the given demand models with $\rho = (0,2,1), (0,3,2), (0,4,3)$. Notice $\rho = (0,4,3)$ converts the demand model into a demand model with only a single state. Last, we use FMA to reduce the given demand models into a single state approximation, then used to generate the approximate policy. To keep the problem small enough for the full solution we study the problem with initial storage of 2 units of each commodity in the storage vertex s_1.

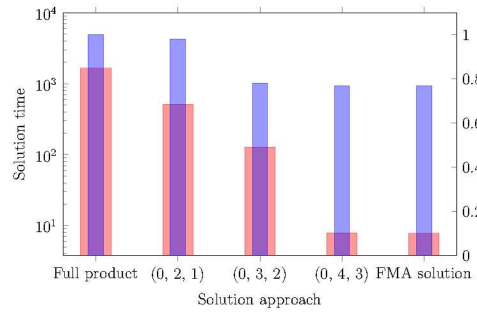

Fig. 2: A bar plot showing the solution time (red) on the left axis (logarithmic scale in seconds) and quality (blue) on the right axis for the different solution approaches.

To verify the correctness of the different solution approaches, we simulated the execution of the policies 1000 times. The results are shown in Figure 2. The graph presents the data in a way that allows comparison of performance of the solution method on two axes. The first is the time that it takes to generate a policy (including the time taken for reduction and constructing the **Lwd** MDP), shown with respect to the left-hand (log-scale) axis. The second is the quality of the resulting policy, which is measured as the average time taken by a policy to go from initial storage to θ in the 1000 simulations, shown with respect to the right-hand axis. The full solution (without any reduction applied to the demand models) provides the best policy selling all the commodities in less time, on average, than the others. However, it takes the longest to provide this solution. As is evident from Figure 2, as more reductions are applied, the time to generate the solutions decreases, while time to sell increases. These approach the demand model of a single state with both FMA and recursive Hellinger distance finally collapsing. Both 1-state models take significantly less time to generate the solution compared to the other solutions.

B. Lean manufacturing

Next, we consider a scenario where an agent in a lean manufacturing factory floor producing nails and screws must transport the raw materials (iron bars) to the different machines on the floor (Figure 3 (Top)). Assume that the demand for nails and screws is directly reflected in demand for iron bars for each machine. Thus, when there arises a demand for nails, one of the nail manufacturing machines produces a demand for iron bars, which are used to manufacture the nails. The agent's objective is to route the iron bars among the storage areas and the machines effectively, so as to reduce

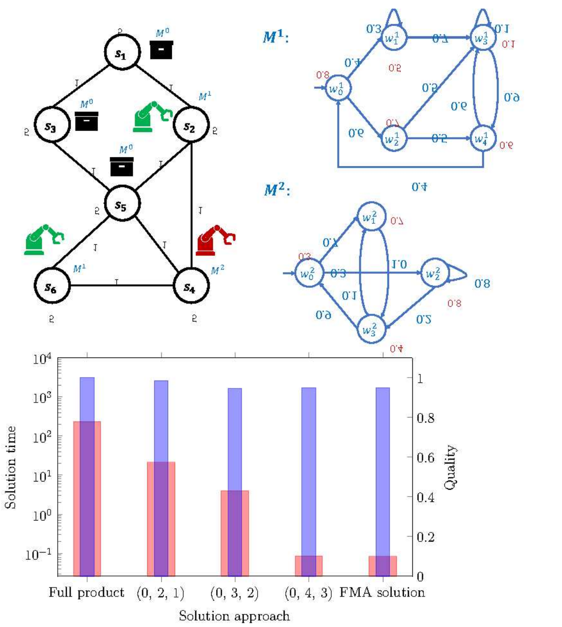

Fig. 3: (Top) Logistic network within a factory floor consisting of 3 storage vertices and 2 nail manufacturing machines and 1 screw manufacturing machine along with their associated demand models (M^1 and M^2). (Bottom) Bar plot showing the solution time (red) on the left axis (logarithmic scale in seconds) and quality (blue) on the right axis for the different solution approaches.

the overall time needed for some initial quantity of iron bars to be used completely. For this example, assume that the demand for iron bars arising from a machine lasts for one time step, i.e., if there are no available iron bars at that point of time, the machine stops (does not consume the iron bars).

To verify the correctness of the solutions, we take an approach similar to the other example by executing the policies a 1000 times. The results are shown in Figure 3 (Bottom). Similar to the other case study, the full solution provides the best solution but takes the longest time to generate the solution, while the fully collapsed (single state) demand models provide significantly faster solutions with slight decreases in quality. All the other reduction solutions lie between the full and 1-state solutions, with the solution quality and time decreasing as more reductions are applied.

VII. LIMITATIONS

Previous sections formulated the problem, gave methods for compression of demand models, and examined simulations showing how such compression improves solution time significantly, with minor decreases in solution quality. Now, we give a constructed example where the reduction based on Hellinger distance is detrimental: giving a poor policy and requiring a longer time to solve.

Consider Figure 4 (Top). The problem consists of two demand models that are deterministic regarding their state transitions. The problem is solved considering six different approximations, which are (1) the full product MDP; (2) a reduced MDP solution by reducing both the demand models into 7 states each via Hellinger distance approach ($\rho = (0, 2, 2)$); (3) a reduced MDP solution with $\rho = (0, 7, 7)$ (reduced to 2 states); (4) a reduced MDP solution with $\rho = (0, 8, 8)$ (reduced to 1 state); (5) the FMA approach;

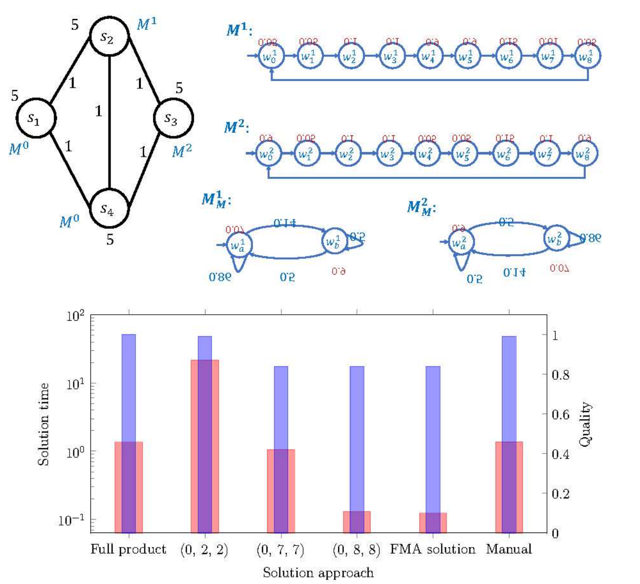

Fig. 4: (Top) Logistic network consisting of 2 storage vertices and consumer vertices with their associated demand models (M^1 and M^2). The demand models M^1_M and M^2_M corresponds to the manually reduced versions of M^1 and M^2. (Bottom) Bar plot showing the solution time (red) on the left axis (logarithmic scale in seconds) and quality (blue) on the right axis for the different solution approaches.

and (6) a reduced solution MDP produced by reducing the demand models into 2 states each manually. Quantitative results appear in the plot in Figure 4 (Bottom).

We can see from the results that even though the quality of the solution is similar for both approaches (1) and (2), the reduction solution here takes longer to generate the solution than the full-product MDP solution approach. The solution quality for the manually reduced solution approach (6) is better than the 2-state reduction via Hellinger distance (3). Not only that, the solution's quality is better compared to (4) or (5). The demand model here, being totally deterministic, was contrived to show that compression which operates by merging pairs of states incrementally can be too myopic. The variation in solution times arise from the fact that by reducing the deterministic demand model from 9-states, we make it stochastic, leading to an increase in the MDP's size.

The purpose of (6) is was to show that idiosyncratic models of demand may require techniques other than the Hellinger reduction one we propose, but that the casual decoupling ideas which underly the planning approach remain effective, no matter the source of the reduced demand model. The ideas we have explored are, thus, modular.

VIII. CONCLUSIONS

In conclusion, we have considered the problem of routing of multiple commodities within a logistic network in the presence of stateful stochastic demand. We present the problem in a modular form consisting of demand models and a logistic network, show three approaches to treat the problem, and present case studies to understand the effect of the approaches on the solution time and quality. Future work might consider other approaches to treating the individual components of the problem for efficient approximations.

References

[1] R. Zhang, F. Rossi, and M. Pavone, "Model predictive control of autonomous mobility-on-demand systems," in *Proceedings of IEEE International Conference on Robotics and Automation (ICRA)*, Stockholm, Sweden, May 2016.

[2] F. Rossi, R. Zhang, Y. Hindy, and M. Pavone, "Routing autonomous vehicles in congested transportation networks: structural properties and coordination algorithms," *Autonomous Robots*, vol. 42, no. 2, pp. 1427–1442, February 2018.

[3] M. Schaefer, M. Čáp, J. Mrkos, and J. Vokřínek, "Routing a fleet of automated vehicles in a capacitated transportation network," in *Proceedings of IEEE International Conference on Robotics and Automation (ICRA)*, Stockholm, Sweden, May 2016.

[4] D. Chaudhuri, R. Ike, H. Rahmani, D. A. Shell, A. T. Becker, and J. M. O'Kane, "Conditioning style on substance: Plans for narrative observation," in *Proceedings of IEEE International Conference on Robotics and Automation (ICRA)*, Xi'an, China/Online, May 2021.

[5] D. Chaudhuri, H. Rahmani, D. A. Shell, and J. M. O'Kane, "Tractable planning for coordinated story capture: Sequential stochastic decoupling," in *Proceedings of International Symposium on Distributed Autonomous Robotic Systems (DARS)*, Kyoto, Japan/Online, June 2021.

[6] D. R. Fulkerson and L. R. Ford, *Flows in Networks*. Princeton University Press, 1962.

[7] T. C. Hu, "Multi-commodity network flows," *Operations research*, vol. 11, no. 3, pp. 344–360, May 1963.

[8] K. Salimifard and S. Bigharaz, "The multicommodity network flow problem: state of the art classification, applications, and solution methods," *Operational Research*, vol. 22, no. 1, pp. 1–47, April 2020.

[9] Q. Jin, S. Feng, M. Li-xin, and T. Gui-jun, "Optimal model and algorithm for multicommodity logistics network design considering stochastic demand and inventory control," *Systems Engineering—Theory and Practice*, vol. 29, no. 4, pp. 176–183, April 2009.

[10] G. A. Tanonkou, L. Benyoucef, and X. Xie, "Design of multi-commodity distribution network with random demands and supply lead-times," in *Proceedings of IEEE International Conference on Automation Science and Engineering (CASE)*, Scottsdale, AZ, USA, September 2007.

[11] X. Gao and G. M. Lee, "A stochastic programming model for multicommodity redistribution planning in disaster response," in *Proceedings of International conference on Advances in Production Management Systems (APMS)*, Seoul, Korea, August 2018.

[12] I. Mejri, S. B. Layeb, M. Haouari, and F. Z. Mansour, "A simulation-optimization approach for the stochastic discrete cost multicommodity flow problem," *Engineering Optimization*, vol. 52, no. 3, pp. 507–526, May 2019.

[13] Q. Liu, Q. Zhao, and W. Zang, "Study on multi-objective optimization of flow allocation in a multicommodity stochastic-flow network with unreliable nodes," *Journal of Applied Mathematics and Computing*, vol. 28, no. 1, pp. 185–198, June 2008.

[14] S. Ding, "Uncertain minimum cost multicommodity flow problem," *Soft Computing*, vol. 21, no. 1, pp. 223–231, September 2017.

[15] H. Topaloglu and W. B. Powell, "Dynamic-programming approximations for stochastic time-staged integer multicommodity-flow problems," *INFORMS Journal on Computing*, vol. 18, no. 1, pp. 31–42, February 2006.

[16] Y.-K. Lin, "Study on the multicommodity reliability of a capacitated-flow network," *Computers and Mathematics with Applications*, vol. 42, no. 1–2, pp. 255–264, July 2001.

[17] G. Levitin, I. Gertsbakh, and Y. Shpungin, "Evaluating the damage associated with intentional supply deprivation in multicommodity network," *Reliability Engineering & System Safety*, vol. 119, pp. 11–17, November 2013.

[18] D. P. Bertsekas, *Dynamic Programming and Optimal Control*. Athena Scientific, 1995.

[19] C. M. Grinstead and J. L. Snell, *Introduction to Probability*. American Mathematical Society, 1997.

[20] E. Hellinger, "Neue begründung der theorie quadratischer formen von unendlichvielen veränderlichen." *Journal für die reine und angewandte Mathematik*, vol. 1909, no. 136, pp. 210–271, 1909.