VNU Journal of Science: Mathematics - Physics

Journal homepage: https://js.vnu.edu.vn/MaP

Original Article

# Variational Inequality Gap-Based Equilibrium Metrics and a Simplified Multi-Agent Reinforcement Learning Scheme for Multi-Commodity Trade Networks

Pham Thanh Hieu*

*Posts and Telecommunications Institute of Technology, 96A Tran Phu, Ha Dong, Hanoi, Vietnam*

Received 13th February 2026
Revised 26th April 2026; Accepted 31st July 2026

**Abstract:** This paper formulates a multi-commodity trade network as a finite-dimensional monotone variational inequality (VI) and introduces a normalized equilibrium metric derived from the classical VI gap function. For the box-constrained feasible set considered here, the gap can be evaluated exactly from the network state, and the resulting metric vanishes if and only if the state is a VI solution. We then study a projected stochastic approximation scheme whose joint state updates can be interpreted as a simplified VI-aware learning dynamic. Under the convex-potential structure of the trade-network operator, unbiased noise with bounded conditional variance, and Robbins–Monro step sizes, we prove that the iterates converge almost surely to a VI solution and that the gap-based equilibrium metric tends to zero. Unlike heuristic equilibrium-proximity indices used in recent MARL–VI architectures, the proposed metric is derived directly from the VI formulation and is accompanied by explicit stochastic-approximation guarantees. The framework is illustrated on a 32-dimensional trade network constructed from 2022 CEPII BACI bilateral trade data for cereals and oilseeds among the United States, China, Germany, and Brazil. The experiments report the VI gap, normalized equilibrium metric, projected residual, and potential gap in addition to the distance to a reference solution. These diagnostics decrease consistently over the finite computational horizon. The results are largely insensitive to the tested noise levels but show a stronger dependence on the initial step size, illustrating numerical behavior consistent with the theoretical analysis.

*Keywords:* Variational inequality, gap function, equilibrium metric, stochastic approximation, trade network, BACI data.

* Corresponding author.
*E-mail address:* hieupt@ptit.edu.vn

https://doi.org/10.25073/2588-1124/vnumap.5117

## 1. Introduction

Variational inequalities (VIs) provide a unified framework for characterizing equilibrium in systems involving interacting decision makers, shared constraints, and nonlinear costs. In finite dimensions, many network and economic equilibrium problems can be formulated as VIs, allowing their existence, stability, and numerical solution to be studied using established analytical tools. Comprehensive treatments of VI theory and its connections with complementarity problems and constrained optimization can be found in the monographs of Kinderlehrer and Stampacchia and of Facchinei and Pang [1, 2]. Network economics is a natural application area because feasible flows subject to capacity and cost constraints often satisfy equilibrium conditions expressible as a VI [1–3].

This connection is particularly relevant to trade and spatial price networks. Nagurney and co-authors have developed VI formulations of multi-commodity trade models incorporating production and transport costs, tariffs, route choices, capacity constraints, and commodity losses [3–5]. In these models, supply regions determine production and routing decisions, demand regions respond through prices and consumption, and equilibrium is characterized by the absence of a feasible unilateral flow adjustment that improves the corresponding marginal condition. The resulting equilibrium conditions can be represented compactly by a finite-dimensional VI.

Multi-agent reinforcement learning (MARL), meanwhile, provides a flexible computational framework for modeling adaptive interactions among multiple agents. Instead of directly solving a static equilibrium problem, agents update their actions or policies through repeated observations of the environment. Related optimization and game-theoretic research has connected learning dynamics with monotone operators, saddle-point problems, stochastic VIs, expected VIs, and extragradient-type methods [7–9]. These connections suggest that VI theory may provide principled equilibrium measures and convergence tools for learning-based network models.

At the interface of these areas, De Natale, Fargetta, Scrimali, and Battiato proposed a hybrid MARL–VI architecture for multi-commodity international trade networks under crisis conditions, including abrupt supply disruptions and tariff shocks [6]. Their model extends earlier VI-based trade-network formulations by introducing learning agents representing supply and demand regions. The proposed actor–critic architecture combines a gradient-based learning-rate scheduler, prioritized replay, a dual individual–centralized reward, and an equilibrium-proximity index constructed from price differentials and temporal stability. Numerical experiments indicate improved performance relative to MAPPO and MADDPG in the crisis scenarios considered in [6].

This approach also motivates two mathematical questions. First, a heuristic proximity index is not necessarily zero precisely at solutions of the underlying VI and may depend on scaling or trajectory-specific parameters. A gap function derived directly from the VI formulation offers a more intrinsic measure of equilibrium violation. Second, simulation-based performance does not by itself provide convergence guarantees for the induced state dynamics. Stochastic approximation theory, together with gap functions and projection methods for VIs, provides a natural framework for addressing these issues [1, 2, 7–9].

The present paper develops this connection in a deliberately simplified setting. We consider a capacitated multi-commodity trade network in which the VI operator represents marginal production costs, marginal transport costs, and demand prices [3–5]. Under commodity-independent quadratic transport coefficients on each route, the operator is shown to be monotone and Lipschitz continuous and to admit a convex-potential representation. We then introduce a normalized equilibrium metric derived from the classical VI gap function and use it to assess the convergence of a projected stochastic approximation scheme. The joint projected update may be interpreted as a simplified VI-aware learning dynamic, although it does not attempt to reproduce the full actor–critic architecture of [6].

Our main contributions can be summarized as follows:

- We formulate a multi-commodity trade network as a finite-dimensional VI. Under the stated cost and demand assumptions, we prove that the associated operator is monotone and Lipschitz continuous, admits a convex potential, and has a nonempty solution set.
- We define the normalized equilibrium metric

$$G(x) = \frac{G(x)}{1 + G(x)}, \qquad G(x) = \sup_{y \in X} \langle F(x), x - y \rangle,$$

and establish its continuity and exactness: $E(x) = 0$ if and only if $x$ solves $\mathrm{VI}(F, X)$. For the box-constrained feasible set used in the model, $G(x)$ can be evaluated exactly componentwise.
- We analyze a projected stochastic approximation scheme under unbiased noise with bounded conditional variance and Robbins–Monro step sizes. Exploiting the convex-potential structure, we prove almost sure convergence of the iterates to a VI solution and show that $G(x_t)$ and $E(x_t)$ tend to zero.
- We illustrate the framework on a 32-dimensional network constructed from 2022 CEPII BACI bilateral trade data for cereals and oilseeds among the United States, China, Germany, and Brazil. In addition to the distance to a reference solution, the experiments report the VI gap, normalized equilibrium metric, projected residual, and potential gap. These diagnostics decrease over the finite computational horizon; the results are largely insensitive to the tested noise levels but show a stronger dependence on the initial step size.

The remainder of the paper is organized as follows. Section 2 presents the trade-network VI formulation and the gap-based equilibrium metric. Section 3 introduces the projected stochastic approximation scheme and establishes its almost sure convergence. Section 4 describes the BACI-based numerical experiments and reports the equilibrium diagnostics. Section 5 concludes and discusses possible extensions.

## 2. Variational Inequality Formulation and Gap-based Equilibrium Metric

This section formulates the multi-commodity trade-network equilibrium problem as a finite-dimensional variational inequality. We establish the monotonicity, Lipschitz continuity, and convex-potential structure of the associated operator and then introduce a gap-based equilibrium metric for assessing convergence.

### *2.1. Network Structure and Feasible Set*

Let $S, D, L,$ and $K$ be finite sets of supply nodes, demand nodes, routes, and commodities, respectively. For each $(i, j, \ell, k) \in S \times D \times L \times K$, let $x_{ij\ell}^k \geq 0$ the quantity of commodity $k$ shipped from supply node $i$ to demand node $j$ along route $\ell$. Collecting all these variables in a fixed order yields a vector $x \in \mathbb{R}^{\mathrm{d}}$, where

$$d = |S| \cdot |D| \cdot |L| \cdot |K|.$$

Each route–commodity flow is subject to a capacity bound

$$0 \leq x_{ij\ell}^k \leq u_{ij\ell}^k, \qquad u_{ij\ell}^k > 0, \qquad \forall\, (i, j, \ell, k) \in S \times D \times L \times K,$$

where $u_{ij\ell}^k > 0$ are given upper bounds that dominate any underlying node or arc capacity limits. The feasible set is therefore

$$X = \{\, x \in \mathbb{R}^{\mathrm{d}} : 0 \leq x_{ij\ell}^k \leq u_{ij\ell}^k \text{ for all } i, j, \ell, k \,\}.$$

The zero vector belongs to $X$; hence $X$ is nonempty. Moreover, $X$ is closed, convex, and bounded, and is therefore compact. The Euclidean projection $P_X$ is obtained by componentwise clipping:

$$\left(P_X(z)\right)_{ij\ell}^k = \min\{u_{ij\ell}^k, \max\{0, z_{ij\ell}^k\}\}.$$

For later use, define the following aggregate flows:

$$Q_i^k(x) = \sum_{j\in D}\sum_{\ell\in L} x_{ij\ell}^k,$$

representing the total outflow of commodity $k$ from supply node $i$;

$$q_{ij\ell}(x) = \sum_{k\in K} x_{ij\ell}^k,$$

representing the total flow of all commodities on route $(i, j, \ell)$; and

$$D_j^k(x) = \sum_{i\in S}\sum_{\ell\in L} x_{ij\ell}^k,$$

representing the total inflow of commodity kkk to demand node $j$.

*2.2. Cost, Price and Variational Inequality Operator*

The model follows the standard structure of spatial price and trade-network equilibrium models, in which production and transport costs are convex and increasing, whereas inverse demand is decreasing [3–5].

For each supply node i and commodity $k$, let the production cost be

$$C_i^k(x) = b_i^k\, Q_i^k(x) + \frac{1}{2}\, c_i^k \left(Q_i^k(x)\right)^2,$$

where $b_i^k > 0$ and $c_i^k > 0$. Its marginal cost with respect to to any flow $x_{ij\ell}^k$ is

$$\frac{\partial C_i^k}{\partial x_{ij\ell}^k}(x) = b_i^k + c_i^k Q_i^k(x).$$

For each route $(i, j, \ell)$ and commodity $k$, the aggregate transport cost is

$$T_{ij\ell}^k(x) = \frac{1}{2}\eta_{ij\ell}^k \left(q_{ij\ell}(x)\right)^2, \quad \eta_{ij\ell}^k > 0.$$

The coefficient $\eta_{ij\ell}^k$ is common to all commodities transported on the same route. Hence, the marginal transport cost is

$$\frac{\partial T_{ij\ell}^k}{\partial x_{ij\ell}^k}(x) = \eta_{ij\ell}^k\, q_{ij\ell}(x).$$

At each demand node $j$ and for each commodity $k$, the price is given by a linear inverse demand function

$$P_j^k(x) = p_{0,j}^k - p_{1,j}^k\, D_j^k(x),$$

where $p_{0,j}^k \in \mathbb{R}$ and $p_{1,j}^k > 0$. Thus, the demand price decreases as the total inflow increases.

The effective marginal-cost operator $F: X \to \mathbb{R}^d$ is defined componentwise by

$$F_{ij\ell}^k(x) = \left(b_i^k + c_i^k\, Q_i^k(x)\right) + \eta_{ij\ell}^k\, q_{ij\ell}(x) - P_j^k(x).$$

Equivalently,

$$F_{ij\ell}^k(x) = b_i^k - p_{0,j}^k + c_i^k\, Q_i^k(x) + \eta_{ij\ell}^k\, q_{ij\ell}(x) + p_{1,j}^k\, D_j^k(x),$$

for all $(i, j, \ell, k)$. We write $F(x) = (F_{ij\ell}^k(x)) \in \mathbb{R}^{\mathrm{d}}$.

**Proposition 2.1 (Monotonicity and Lipschitz continuity of F).** Assume that $c_i^k > 0,\ \eta_{ij\ell}^k > 0, p_{1,j}^k > 0$ for all relevant indices. Then the operator $F$ is monotone and Lipschitz continuous on $X$.

*Proof.* Let $x, y \in X$ and set $\Delta = x - y$. Since the aggregate-flow mappings are linear,

$$Q_i^k(x) - Q_i^k(y) = Q_i^k(\Delta),$$

with analogous identities for $q_{ij\ell}$ and $D_j^k$. Therefore,

$$\langle F(x) - F(y), x - y\rangle = \sum_{i,k} c_i^k \left(Q_i^k(\Delta)\right)^2 + \sum_{ij\ell} \eta_{ij\ell} \left(q_{ij\ell}(\Delta)\right)^2 + \sum_{j,k} p_{1,j}^k \left(D_j^k(\Delta)\right)^2 \geq 0.$$

Hence, $F$ is monotone. Moreover, $F$ is affine because each of its components is an affine function of the aggregate flows. Thus, there exist a constant matrix $J_F$ and a vector $h$ such that

$$F(x) = J_F(x) + h.$$

Consequently,

$$\|F(x) - F(y)\| \leq \|J_F(x - y)\| \leq \|J_F\|_2 \|x - y\|.$$

Therefore, $F$ is Lipschitz continuous with Lipschitz constant $L = \|J_F\|_2$. □

The trade network equilibrium problem is to find $x^* \in X$ such that

$$\langle F(x^*), x - x^*\rangle \geq 0 \text{ for all } x \in X.$$

We denote this problem by $\mathrm{VI}(F, X)$ and define its solution set by

$$X^* = \{x \in X : x \text{ solves } \mathrm{VI}(F, X)\}.$$

The operator also admits an explicit convex-potential representation. Define

$$\Phi(x) = \sum_{i,k} \left[b_i^k Q_i^k(x) + \frac{1}{2} c_i^k \left(Q_i^k(x)\right)^2\right] + \sum_{i,j,\ell} \frac{1}{2} \eta_{ij\ell} \left(q_{ij\ell}(x)\right)^2$$
$$+ \sum_{j,k} \left[p_{0,j}^k\, D_j^k(x) + \frac{1}{2}\, p_{1,j}^k\, \left(D_j^k(x)\right)^2\right].$$

Each quadratic term has a positive coefficient, so $\Phi$ is continuously differentiable and convex. Direct differentiation gives

$$\nabla\Phi(x) = F(x).$$

Therefore, VI(F,X) is equivalent to the convex optimization problem

$$\min_{x \in X} \Phi(x),$$

and

$$X^* = \operatorname*{argmin}_{x \in X} \Phi(x).$$

**Proposition 2.2. Existence of solutions.** Under the assumptions of Proposition 2.1, the solution set $X^*$ is nonempty.

*Proof.* The feasible set $X$ is nonempty, compact, and convex, and $F$ is continuous. The Hartman–Stampacchia existence theorem therefore guarantees that $\mathrm{VI}(F, X)$ has at least one solution; see [2, Theorem I.4.1] or [1, Corollary 2.2.5]. □

*2.3. Gap functions and Normalized Equilibrium Metric*

For $x \in X$, define the classical VI gap function [10] by:

$$G(x) = \max\langle F(x), x - y\rangle.$$

The maximum is well defined because $X$ is compact.

**Proposition 2.3 (Properties of the gap function).** The function $G : X \to \mathbb{R}$ satisfies:

i) $G(x) \geq 0$ for all $x \in X$;

ii) $G(x) = 0$ if and only if $x \in X^*$;

iii) $G$ is continuous on $X$.

*Proof.* i) Setting $y = x$ gives $G(x) \geq 0$.

ii) If $x \in X^*$, then $\langle F(x), x - y\rangle \leq 0$ for every $y \in X$, so $G(x) \leq 0$; together with (i), $G(x) = 0$. Conversely, if $G(x) = 0$, then $\langle F(x), y - x\rangle \geq 0$ for all $y \in X$, so $x$ solves $\mathrm{VI}(F, X)$ or $x \in X^*$.

iii) The function $\varphi(x, y) = \langle F(x), x - y\rangle$ is continuous on the compact set $X \times X$ and is therefore uniformly continuous. It follows that its maximum over $y \in X$ varies continuously with $x$. Hence $G$ is continuous. □

For the box-constrained feasible set used here, the gap can be evaluated exactly without solving an auxiliary optimization problem. Writing $r = (i, j, \ell, k)$ for a flow coordinate, we have

$$G(x) = \sum_r g_r(x),$$

where

$$g_r(x) = \begin{cases} F_r(x)x_r, & F_r(x) \geq 0, \\ F_r(x)(x_r - u_r), & F_r(x) < 0. \end{cases}$$

Indeed, when $F_r(x) \geq 0$, the maximizing coordinate $y_r = 0$, whereas when $F_r(x) < 0$, it is $y_r = u_r$. We define the normalized equilibrium metric by

$$E(x) = \frac{G(x)}{1 + G(x)} \in [0,1).$$

**Corollary 2.4.** The function $E: X \to [0, 1)$ is continuous and satisfies $E(x) = 0$ if and only if $x \in X^*$. Moreover, $E$ is strictly increasing as a function of the gap value $G(x)$.

*Proof.* The scalar function $\rho(t) = t/(1 + t)$ is continuous and strictly increasing on $[0, \infty)$ with $\rho(0) = 0$. The result follows from Proposition 2.3 and composition. □

## 3. Projected Stochastic Approximation Algorithm and Convergence

This section introduces a projected stochastic approximation scheme for solving $\mathrm{VI}(F, X)$. Exploiting the convex-potential structure established in Section 2, we prove almost sure convergence of the iterates and of the proposed gap-based equilibrium metric..

### *3.1. Projected Stochastic Approximation Scheme*

Let $(\Omega, \mathfrak{F}, \mathbb{P})$ be a probability space supporting a sequence of random variables $\{\xi_t\}$, and let $\mathfrak{F}_t = \sigma(x_0, \xi_0, \dots, \xi_{t-1})$ be the natural filtration. At iteration $t$, given $x_t \in X$, we observe a noisy estimate $\hat{F}_t = \hat{F}(x_t, \xi_t)$ of $F(x_t)$ and perform the update

$$x_{t+1} = P_X\big(x_t - \alpha_t \hat{F}_t\big), \qquad t = 0, 1, 2, \dots \tag{1}$$

where $x_0 \in X, P_X$ denotes the Euclidean projection onto $X$, and $\{\alpha_t\}$ is a sequence of positive step sizes.

We impose the following assumptions.

**- (A1) Convex-potential and Lipschitz structure.** There exists a continuously differentiable convex function $\Phi: \mathbb{R}^d \to \mathbb{R}$ such that $F = \nabla\Phi$ and $F$ is Lipschitz continuous on $X$ with constant $L > 0$.

- **(A2) Nonemptiness of the solution set.** The set $X^* = \{x \in X : x \text{ solves VI}(F, X)\}$ is nonempty.
- **(A3) Unbiased noise with bounded conditional variance.** For every $t \geq 0$,

$$\mathbb{E}[\hat{F}_t | \mathfrak{F}_t] = F(x_t),$$

and there exists $\sigma^2 < \infty$ such that

$$\mathbb{E}[\|\hat{F}_t - F(x_t)\|^2 \mid \mathfrak{F}_t] \leq \sigma^2.$$

- **(A4) Robbins–Monro step sizes.** The sequence $\{\alpha_t\}$ satisfies

$$\alpha_t > 0, \qquad \Sigma_t \, \alpha_t = \infty, \qquad \Sigma_t \, {\alpha_t}^2 < \infty.$$

Assumptions (A1)–(A2) hold for the trade-network operator introduced in Section 2. Proposition 2.1 establishes the Lipschitz continuity of $F$, the explicit potential satisfies $F = \nabla\Phi$, and Proposition 2.2 guarantees that $X^* \neq \emptyset$. Assumptions (A3)–(A4) are standard in stochastic approximation [7, 11].

*3.2. Almost Sure Convergence*

**Theorem 3.1 (Almost sure convergence of the projected stochastic approximation scheme).** Suppose that Assumptions (A1)–(A4) hold. Let $\{x_t\}$ be generated by (1), and define $\Phi^* = \min_{x \in X} \Phi(x)$. Then,

(a) for every $x^* \in X^*$, the sequence $\{\|x_t - x^*\|\}$ converges almost surely;
(b) almost surely, $\sum_{t=0}^{\infty} \alpha_t(\Phi(x_t) - \Phi^*) < \infty$, and consequently, $\liminf_{t \to \infty} \Phi(x_t) = \Phi^*$;
(c) there exists an $X^*$- valued random variable $x_\infty$ such that $x_t \to x_\infty$ almost surely;
(d) almost surely, $G(x_t) \to 0$, and consequently $E(x_t) \to 0$.

*Proof.* Define the noise term $\varepsilon_t = \hat{F}_t - F(x_t)$. By Assumption (A3),

$$\mathbb{E}[\varepsilon_t | \mathfrak{F}_t] = 0 \text{ and } \mathbb{E}[\|\varepsilon_t\|^2 | \mathfrak{F}_t] \leq \sigma^2.$$

Fix $x^* \in X^*$. Since $x^* \in X$, we have $P_X(x^*) = x^*$. The nonexpansiveness of the Euclidean projection [12, Theorem 3.14] gives

$$\begin{aligned} \|x_{t+1} - x^*\|^2 &\leq \|P_X(x_t - \alpha_t \hat{F}_t) - P_X(x^*)\|^2 \\ &\leq \left\|x_t - \alpha_t \hat{F}_t - x^*\right\|^2 \\ &\leq \|x_t - x^*\|^2 - 2\alpha_t\langle \hat{F}_t, x_t - x^*\rangle + \alpha_t^2 \left\|\hat{F}_t\right\|^2. \end{aligned}$$

Taking conditional expectation with respect to $\mathfrak{F}_t$ and using the unbiasedness assumption yield

$$\mathbb{E}[\|x_{t+1} - x^*\|^2 | \mathfrak{F}_t] \leq \|x_t - x^*\|^2 - 2\alpha_t\langle F(x_t), x_t - x^*\rangle + {\alpha_t}^2 \mathbb{E}[\|\hat{F}_t\|^2 | \mathfrak{F}_t].$$

Since $X$ is compact and $F$ is continuous, there exists $M > 0$ such that $\|F(x)\| \leq M$ for all $x \in X$. Furthermore,

$$\mathbb{E}[\|\hat{F}_t\|^2 | \mathfrak{F}_t] = \mathbb{E}[\|F(x_t) + \varepsilon_t\|^2 | \mathfrak{F}_t] = \|F(x_t)\|^2 + \mathbb{E}[\|\varepsilon_t\|^2 | \mathfrak{F}_t] \leq M^2 + \sigma^2.$$

Set C$= M^2 + \sigma^2$. Since $F = \nabla\Phi$ and $\Phi$ is convex

$$\Phi(x^*) \geq \Phi(x_t) + \langle \nabla\Phi(x_t), x^* - x_t\rangle.$$

Because $x^* \in X^* = \operatorname{argmin}_{x \in X} \Phi(x)$, we have $\Phi(x^*) = \Phi^*$. Therefore,

$$\langle \mathrm{F}(x_t), x_t - x^*\rangle \geq \Phi(x_t) - \Phi(x^*) = \Phi(x_t) - \Phi^*.$$

It follows that

$$\mathbb{E}[\|x_{t+1} - x^*\|^2 | \mathfrak{F}_t] \leq \|x_t - x^*\|^2 - 2\alpha_t(\Phi(x_t) - \Phi^*) + C\alpha_t^2. \tag{2}$$

By Assumption (A4), $\sum_t \alpha_t^2 < \infty$. The Robbins–Siegmund almost-supermartingale theorem [13] applied to (2) implies that, almost surely, $\|x_t - x^*\|^2$ converges to a finite random variable and

$$\sum_t \alpha_t(\Phi(x_t) - \Phi^*) < \infty.$$

This proves (a) and the first assertion in (b).

Since, $\Phi(x_t) - \Phi^* \geq 0$ and $\sum_t \alpha_t = \infty$, the finiteness of the weighted sum implies

$$\liminf_{t\to\infty}(\Phi(x_t) - \Phi^*) = 0.$$

Thus, $\liminf_{t\to\infty} \Phi(x_t) = \Phi^*$, which completes the proof of (b).

We next prove convergence of the whole sequence. Choose a subsequence $\{x_{t_n}\}$ such that $\Phi(x_{t_n}) \to \Phi^*$. Because $X$ is compact, this subsequence has a further subsequence, denoted again by $\{x_{t_n}\}$, such that $x_{t_n} \to \bar{x}$ for some $\bar{x} \in X$. By continuity of $\Phi$,

$$\Phi(\bar{x}) = \lim_{n\to\infty} \Phi(x_{t_n}) = \Phi^*,$$

so $\bar{x} \in X^*$.

To show that the whole sequence converges to $\bar{x}$, let

$$D = \{z_m : m \in \mathbb{N}\}$$

be a countable dense subset of $X^*$. By part (a), and by taking the intersection of countably many probability-one events, we may work on an event of probability one on which $\|x_t - z_m\|$ converges for every $m \in \mathbb{N}$.

For any $\epsilon > 0$, choose $z_m \in D$ such that $\|z_m - \bar{x}\| < \epsilon$. The reverse triangle inequality gives

$$|\|x_t - \bar{x}\| - \|x_t - z_m\|| \leq \|\bar{x} - z_m\| < \epsilon.$$

for every t. Since $\|x_t - z_m\|$ has a limit, it follows that

$$\limsup_{t\to\infty}\|x_t - \bar{x}\| - \liminf_{t\to\infty}\|x_t - \bar{x}\| \leq 2\epsilon.$$

Because $\epsilon > 0$ is arbitrary, the sequence $\|x_t - \bar{x}\|$ has a limit. Along the subsequence $x_{t_n} \to \bar{x}$, this distance tends to zero. Therefore, $\lim_{t\to\infty} \|x_t - \bar{x}\| = 0$. Thus, $x_t \to \bar{x} \in X^*$ almost surely. Setting $x_\infty = \bar{x}$ proves (c).

Finally, the gap function $G$ is continuous by Proposition 2.3. Since $x_t \to x_\infty \in X^*$,

$$G(x_{\mathrm{t}}) \to G(x_\infty) = 0 \text{ almost surely.}$$

By the definition and continuity of the normalized metric $E$,

$$E(x_{\mathrm{t}}) \to E(x_\infty) = 0 \text{ almost surely.}$$

This proves (d). □

**Remark 3.2. VI-aware learning interpretation.** The noisy vector $\hat{F}_{\mathrm{t}}$ may be interpreted as the aggregate marginal-cost and price information observed by supply and demand agents. The projected update moves the joint network state in an approximate negative-gradient direction of the convex potential $\Phi$.

The gap-based metric $E(x_t)$ provides a VI-grounded scalar measure for monitoring equilibrium progress. It could also be incorporated into reward shaping or adaptive learning-rate mechanisms in future learning architectures. However, scheme (1) is a joint projected state update rather than a full actor–critic MARL algorithm: it does not include policy or value-function approximation, experience replay, or neural parameterization.

Compared with stochastic Mirror-Prox and extragradient methods for more general monotone VIs [7, 9], the present scheme is simpler but relies on the additional convex-potential structure. Its role is therefore to provide a mathematically controlled bridge between trade-network VIs and learning-based equilibrium dynamics.

## 4. Numerical Experiments on a Real Multi-commodity Trade Network

This section illustrates the proposed stochastic approximation scheme on a multi-commodity trade network constructed from CEPII BACI data. In addition to the distance to a selected reference solution, we report the VI gap, the normalized equilibrium metric, the projected residual, and the potential gap. The latter quantities are directly aligned with the convergence results of Theorem 3.1.

### *4.1. Data and Model Construction*

We use the 2022 trade-flow file from the CEPII BACI HS17 V202601 database [14]. Four major trading economies, the United States, China, Germany, and Brazil, and two commodity groups, Cereals (HS Chapter 10) and Oilseeds (HS Chapter 12), are considered. The BACI trade values, originally reported in thousands of current USD, are aggregated over all six-digit products belonging to the two selected HS chapters.

With one aggregate route for each exporter–importer pair, the resulting VI has

$$d = 4 \times 4 \times 1 \times 2 = 32$$

decision variables. All trade flows are normalized by the largest observed bilateral flow, $x_{max} = 31.95929884$ billion USD. Thus, one normalized flow unit corresponds to approximately 31.9593 billion USD. The upper bound for each coordinate is set to

$$u_{ij}^k = \max\{2x_{ij,obs}^k, 0.05\},$$

where $x_{ij,obs}^k$ denotes the normalized observed flow. The model parameters are

$$b_i^k = 0.3, c_i^k = 0.5, \eta_{ij\ell}^k = 0.2, p_{0,j}^k = 3.0, p_{1,j}^k = 0.5.$$

The additive constants in the production and transport cost functions do not enter the marginal-cost operator and are therefore set to zero without loss of generality. Under these parameter choices, Proposition 2.1 ensures that $F$ is monotone and Lipschitz continuous, while Proposition 2.2 guarantees that the VI solution set is nonempty.

Figure 1 compares the largest observed bilateral flows with the flows at the computed reference equilibrium.

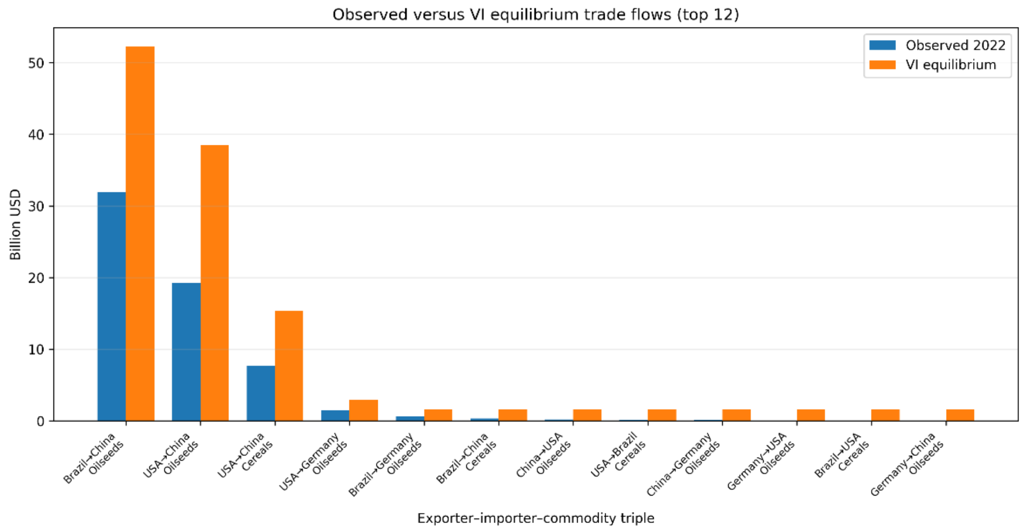


Figure 1. Observed BACI 2022 flows and VI reference-equilibrium flows for the twelve largest exporter–importer–commodity triples.

### *4.2. Experimental Protocol and Equilibrium Diagnostics*

A deterministic projected gradient method with step size $\beta = 0.01$ is first used to construct a reference solution. With the stopping criterion $\|x_{t+1} - x_t\|^2 < 10^{-7}$, the method terminates after 979 iterations. The resulting reference point has VI gap $G(x_{ref}) = 3.570689 \times 10^{-6}$. For evaluating the potential gap, the deterministic iteration is continued with a stricter tolerance, producing a refined reference point with gap $3.598443 \times 10^{-11}$.

The stochastic experiments start from $x_0 = P_X(0.3, 0.3, \cdots, 0.3)^T$ and use

$$\hat{F}(x_t, \xi_t) = F(x_t) + \xi_t, \quad \xi_t \sim \mathcal{N}(0, \sigma^2 I_{32}),$$

together with the Robbins–Monro schedule $\alpha_t = \alpha_0/(t+1)$. Each configuration is run for $T = 15{,}000$ iterations and repeated over five independent random seeds.

Because $X$ is a box, the classical gap can be evaluated exactly and componentwise as

$$G(x) = \sum_{r=1}^{32} \begin{cases} F_r(x)x_r, & F_r(x) \geq 0, \\ F_r(x)(x_r - u_r), & F_r(x) < 0. \end{cases}$$

We report the normalized equilibrium metric $E(x) = \frac{G(x)}{1+G(x)}$, the projected residual

$$R_\gamma(x) = \frac{1}{\gamma}\left\|x - P_X\big(x - \gamma F(x)\big)\right\|, \qquad \gamma = \frac{1}{L},$$

and the potential gap $\Phi(x) - \Phi^*$, where $\Phi^*$ is approximated using the refined deterministic solution.

### *4.3. Results*

Table 1 reports the final mean values over the five seeds. The distance to $x_{ref}$ is retained for comparison with the original experiments, whereas $G(x_T), E(x_T), R_\gamma(x_T)$, and $\Phi(x_T) - \Phi^*$ provide diagnostics that are directly connected with the VI formulation.

Table 1. Final convergence diagnostics after $T = 15{,}000$ iterations.

| $\alpha_0$ | $\sigma$ | $\|x_T - x_{ref}\|$ | $G(x_T)$ | $E(x_T)$ | $R_\gamma(x_T)$ | $\Phi(x_T) - \Phi^*$ |
|---|---|---|---|---|---|---|
| 0.02 | 0.10 | 1.1315 | 2.7333 | 0.7321 | 1.9975 | 1.1562 |
| 0.05 | 0.05 | 0.6714 | 1.0563 | 0.5137 | 1.1815 | 0.3926 |
| 0.05 | 0.10 | 0.6707 | 1.0543 | 0.5132 | 1.1803 | 0.3918 |
| 0.05 | 0.20 | 0.6692 | 1.0504 | 0.5122 | 1.1778 | 0.3901 |
| 0.15 | 0.10 | 0.1696 | 0.1086 | 0.0980 | 0.2035 | 0.0173 |

For $\alpha_0 = 0.05$, changing the noise level from 0.05 to 0.20 has only a minor effect. The final distance lies between 0.6692 and 0.6714, while the normalized equilibrium metric lies between 0.5122 and 0.5137. The differences among these configurations are below 0.6% for the principal diagnostics. Thus, within the tested range, the numerical behavior is largely insensitive to the noise level.

The initial step size has a considerably stronger effect. Since the common initial distance is 1.6231, the choice $\alpha_0 = 0.02$ produces a reduction of only 30.3%. The baseline value $\alpha_0 = 0.05$ yields a reduction of 58.7%. In contrast, $\alpha_0 = 0.15$ reduces the distance by 89.6% and gives

$$G(x_T) = 0.1086, E(x_T) = 0.0980, R_\gamma(x_T) = 0.2035.$$

Its potential gap is only 0.0173, compared with 0.3918 for the baseline configuration. Hence, among the tested values, $\alpha_0 = 0.15$ provides the most effective finite-horizon approximation to equilibrium.

Figure 2 presents the four VI-grounded diagnostics for the baseline configuration $(\alpha_0, \sigma) = (0{,}05, 0.10)$. All four quantities decrease throughout the experiment. Since Theorem 3.1 is asymptotic, their nonzero values at $T = 15{,}000$ do not contradict the convergence result; rather, they quantify the remaining equilibrium error after a finite computational budget.

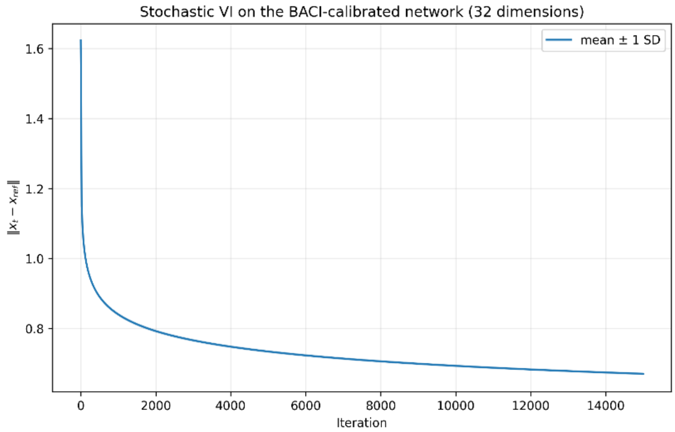


**Figure 2.** Mean VI-grounded convergence diagnostics, with $\pm 1$ standard-deviation bands, for $\alpha_0 = 0.05$, $\sigma = 0.10$, five independent seeds, and $T = 15{,}000$.

*4.4. Discussion*

The experiments show that the projected stochastic approximation scheme consistently reduces the VI gap, normalized equilibrium metric, projected residual, and potential gap. The results are robust to the tested noise levels but more sensitive to the initial step size. Because the VI solution need not be unique, the distance to the selected reference solution should be interpreted only as an empirical diagnostic. The gap-based quantities provide the more appropriate numerical evidence for the convergence behavior established in Theorem 3.1.

## 5. Conclusion

We formulated a finite-dimensional variational inequality model for a multi-commodity trade network and introduced a normalized equilibrium metric derived from the classical VI gap function. The metric vanishes exactly at VI solutions and provides a theoretically grounded measure of equilibrium violation. Under the convex-potential structure of the trade-network operator, unbiased noise with bounded conditional variance, and Robbins–Monro step sizes, the proposed projected stochastic approximation scheme converges almost surely to a VI solution, while both the VI gap and the normalized equilibrium metric tend to zero.

Experiments on a 32-dimensional network constructed from 2022 CEPII BACI data for cereals and oilseeds among the United States, China, Germany, and Brazil illustrate the finite-horizon behavior of the method. The VI gap, normalized equilibrium metric, projected residual, and potential gap decrease consistently during the stochastic iterations. The results are largely insensitive to the tested noise levels but show a stronger dependence on the initial step size, with $\alpha_0 = 0.15$ providing the best performance among the tested configurations. Because the VI solution need not be unique, the gap-based diagnostics provide more appropriate evidence of progress toward equilibrium than the distance to a single reference solution.

The analysis relies on the convex-potential structure of the operator and the compactness of the feasible set. Extensions to general monotone or pseudomonotone operators may require extragradient, Mirror-Prox, or related correction-based methods. Bridging the present idealized stochastic approximation scheme with the full actor–critic MARL architecture considered in [6], including function approximation, replay mechanisms, and neural parameterization, remains an important direction for future research. The proposed gap-based metric may also be used for reward shaping and adaptive learning-rate design in larger trade networks with more countries, commodities, and routes..

## Acknowledgments

This work was supported by the Posts and Telecommunications Institute of Technology (PTIT), Hanoi, Vietnam. The author gratefully acknowledges the institutional support and research facilities provided by PTIT during the completion of this study.

## References

[1] F. Facchinei, J.S. Pang, Finite-Dimensional Variational Inequalities and Complementarity Problems, Vols. I-II, Springer, New York, 2003.

[2] D. Kinderlehrer, G. Stampacchia, An Introduction to Variational Inequalities and Their Applications, SIAM, Philadelphia, 2000, https://doi.org/10.1137/1.9780898719451.

[3] A. Nagurney, I. Pour, S. Samadi, Network Economics: A Variational Inequality Approach, 2nd edition, Kluwer Academic Publishers, Dordrecht, 1999, https://doi.org/10.1007/978-1-4757-3005-0.

[4] A. Nagurney, A Variational Inequality Trade Network Model in Prices and Quantities Under Commodity Losses, Journal of Nonlinear and Variational Analysis, Vol. 8, 2024, pp. 935-952, https://doi.org/10.23952/jnva.8.2024.6.06.

[5] A. Nagurney, S. Samadi, Multicommodity trade, tariffs, and Rerouting, in: P. M. Pardalos, T. M. Rassias (Eds.), Convex and Variational Analysis with Applications: In Honor of Terry Rockafellar's 90th Birthday, Springer, 2026, pp. 319-345, https://doi.org/10.1007/978-3-032-07860-5_16.

[6] L. D. Natale, G. Fargetta, L. R. M. Scrimali, S. Battiato, Multi-agent Reinforcement Learning and Variational Inequality Models for International Trade Networks Under Crisis, Simulation Modelling Practice and Theory, Vol. 146, 2026, pp. 103219, https://doi.org/10.1016/j.simpat.2025.103219.

[7] A. Juditsky, A. Nemirovski, C. Tauvel, Solving Variational Inequalities with Stochastic Mirror-Prox Algorithm, Stochastic Systems, Vol. 1, No. 1, 2011, pp. 17-58, https://doi.org/10.1287/10-SSY011.

[8] B. H. Zhang, I. Anagnostides, E. Tewolde, R. E. Berker, G. Farina, V. Conitzer, T. Sandholm, Expected variational inequalities, Proceedings of the 42nd International Conference on Machine Learning, Proceedings of Machine Learning Research, Vol. 267, 2025, pp. 74422-74446, https://proceedings.mlr.press/v267/zhang25a.html (accessed on: January 10th, 2026).

[9] E. Gorbunov, N. Loizou, G. Gidel, Extragradient Method: O(1/K) Last-iterate Convergence For Monotone Variational Inequalities and Connections with Cocoercivity, in: Proceedings of the 25th International Conference on Artificial Intelligence and Statistics, Proceedings of Machine Learning Research, Vol. 151, 2022, pp. 366-402, https://proceedings.mlr.press/v151/gorbunov22a.html (accessed on: January 10th, 2026).

[10] A. Auslender, Optimisation: Methods Numeriques, Masson, Paris, 1976.

[11] H. Robbins, S. Monro, A stochastic approximation method, Ann. Math. Statist, Vol. 22, No. 3, 1951, pp. 400-407. https://doi.org/10.1214/aoms/1177729586.

[12] H. H. Bauschke, P. L. Combettes, Convex Analysis and Monotone Operator Theory in Hilbert Spaces, 2nd edition, Springer, Cham, 2017, https://doi.org/10.1007/978-3-319-48311-5.

[13] H. Robbins, D. Siegmund, A Convergence Theorem for Nonnegative Almost Supermartingales and Some Applications, in: J.S. Rustagi (Ed.), Optimizing Methods in Statistics, Academic Press, New York, 1971, pp. 233-257, https://doi.org/10.1016/B978-0-12-604550-5.50015-8.

[14] G. Gaulier, S. Zignago, BACI: International Trade Database at the Product-Level, the 1994-2007 Version, CEPII Working Paper No. 2010-23, 2010.