

Differential Stochastic Variational Inequalities with Parametric Optimization

Xiaojun Chen

Department of Applied Mathematics
The Hong Kong Polytechnic University

ICM2026, Section 18: Stochastic and Differential Modelling

25 July, 2026, Philadelphia

- 1 Differential Stochastic Variational Inequality (DSVI)
 - Two-stage stochastic VI
 - Differential VI
 - Differential system with stochastic optimization
- 2 DSVI with Parametric Optimization (DSVI-O) in AI
 - Embodied intelligence system for the elderly
 - Basic solution properties
- 3 Discrete Approximation
 - Euler scheme and sample average approximation
- 4 Conclusions

Variational Inequality (VI)

■ **Problem.** Given a nonempty closed convex set $X \subseteq \mathbb{R}^n$ and a continuous function $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$, the VI problem is to find $x \in X$ such that

$$0 \in F(x) + \mathcal{N}_X(x).$$

$\mathcal{N}_X(x) = \{d \mid d^\top(y - x) \leq 0, \forall y \in X\}$ is the **normal cone** to X at x , which is equivalent to

$$x = \Pi_X(x - F(x)),$$

where $\Pi_X : \mathbb{R}^n \rightarrow X$ is the Euclidean projection operator onto X .

■ **Two special cases.**

► **Complementarity problem:** $X = \mathbb{R}_+^n$

$$-F(x) \in \mathcal{N}_{\mathbb{R}_+^n}(x) \Leftrightarrow 0 \leq x \perp F(x) \geq 0$$

► **System of equations:** $X = \mathbb{R}^n$

$$F(x) = 0$$

★ **First order optimality condition:** $-\nabla f(x) \in \mathcal{N}_X(x)$ for

$$\min_{x \in X} f(x)$$

■ X. Chen & Jinglai Shen (Univ. Maryland, BC), SIAM J. Optim. 2022

► Consider the following ODE:

$$\dot{x}(t) = \Pi_X \left(x(t) - \mathbb{E}[\Phi(t, \xi, x(t), y(t, \xi))] \right) - x(t), \quad t \geq 0 \quad (1)$$

$$x(0) = x_0, \quad (2)$$

$$0 \in \Psi(t, \xi, x(t), y(t, \xi)) + \mathcal{N}_{C_\xi}(y(t, \xi)), \quad \text{for a.e. } \xi \in \Xi. \quad (3)$$

■ Notation.

- $X \subseteq \mathbb{R}^n$: nonempty closed convex set,
- $\xi : \Omega \rightarrow \mathbb{R}^\ell$: random vector on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ whose probability distribution $P = \mathbb{P} \circ \xi^{-1}$ is supported on the set $\Xi := \xi(\Omega) \subseteq \mathbb{R}^\ell$,
- $\Phi : \mathbb{R}_+ \times \mathbb{R}^\ell \times \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^n$,
- $\Psi : \mathbb{R}_+ \times \mathbb{R}^\ell \times \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^m$,
- C_ξ : closed convex set in $\mathbb{R}^m$ for each ξ and is $\mathcal{A}$ -measurable,
- $\mathcal{N}_{C_\xi}(y(t, \xi))$: the normal cone to C_ξ at $y(t, \xi)$.

■ Special case.

- If $X = \mathbb{R}_+^n$, $C_\xi = \mathbb{R}_+^m$ for all $\xi \in \Xi$,

$$\Phi(t, \xi, x, y(t, \xi)) = Ax + B(\xi)y(t, \xi) + q_1(t),$$

$$\Psi(t, \xi, x, y(t, \xi)) = M(\xi)y(t, \xi) + N(\xi)x + q_2(t, \xi),$$

then the **DSVI** becomes the **DSLCP** given by:

$$\begin{aligned} \dot{x} &= \Pi_{\mathbb{R}_+^n} \left(x - (Ax + \mathbb{E}[B(\xi)y(t, \xi)] + q_1(t)) \right) - x \\ x(0) &= x_0 \\ 0 &\leq y(t, \xi) \perp M(\xi)y(t, \xi) + N(\xi)x + q_2(t, \xi) \geq 0, \quad \text{a.e. } \xi \in \Xi. \end{aligned}$$

- $A \in \mathbb{R}^{n \times n}$, $B : \Xi \rightarrow \mathbb{R}^{n \times m}$, $q_1 : \mathbb{R}_+ \rightarrow \mathbb{R}^n$
- $M : \Xi \rightarrow \mathbb{R}^{m \times m}$, $N : \Xi \rightarrow \mathbb{R}^{m \times n}$, $q_2 : \mathbb{R}_+ \times \Xi \rightarrow \mathbb{R}^m$

Equilibria of the DSVI and two-stage SVI

► Let $\mathcal{Y}$ be the space of measurable functions from Ξ to $\mathbb{R}^m$. Then $(x^e, y^e(\xi)) \in \mathbb{R}^n \times \mathcal{Y}$ is an **equilibrium** of the DSVI if for a.e. $\xi \in \Xi$,

$$\begin{aligned} 0 &= \Pi_X(x^e - \mathbb{E}[\Phi(\xi, x^e, y^e(\xi))]) - x^e \\ 0 &\in \Psi(\xi, x^e, y^e(\xi)) + \mathcal{N}_{C_\xi}(y^e(\xi)), \quad \text{for a.e. } \xi \in \Xi, \end{aligned}$$

► or equivalently it is a solution to the following **two-stage SVI**:

$$\begin{aligned} 0 &\in \mathbb{E}[\Phi(\xi, x, y(\xi))] + \mathcal{N}_X(x), \\ 0 &\in \Psi(\xi, x, y(\xi)) + \mathcal{N}_{C_\xi}(y(\xi)), \quad \text{for a.e. } \xi \in \Xi. \end{aligned}$$

- X. Chen, T.K. Pong, R. Wets, Math. Program. (2017)
- R.T. Rockafellar, R. Wets, Math. Program. (2017)
- X. Chen, A. Shapiro, H. Sun, SIAM J. Optim. (2018)
- X. Chen, H. Sun, H. Xu, Math Program. (2019)
- X. Wang, X. Chen, SIAM J. Sci. Comp. (2022)

■ Differential Variational Inequality (DVI)

▶ When $X = \mathbb{R}^n$, and $y(t, \xi)$ and C_ξ are deterministic, the DSVI becomes the deterministic DVI:

$$\begin{aligned}
 \dot{x}(t) &= \Phi(t, x(t), y(t)), & t \geq 0 \\
 x(0) &= x_0, \\
 0 &\in \Psi(t, x(t), y(t)) + \mathcal{N}_C(y(t))
 \end{aligned}$$

- J.-S. Pang, D. Stewart, Math. Program. (2008)
- X. Chen, S. Xiang, Math. Program. (2013)
- X. Chen, Z. Wang, SIAM J. Optim. (2013)
- X. Chen, Z. Wang, Math. Program. (2014)
- S. Wu, X. Chen, SIAM J. Sci. Comp. (2018)
- S. Wu, T. Zhou, X. Chen, SIAM J. Control Optim. (2020)
- J. Luo, X. Chen, SIAM J. Control Optim. (2023)

■ Jianfeng Luo & X. Chen, SIAM J. Control & Optim. 2025

► Consider the following system:

$$\dot{x}(t) = \mathbb{E}[\Phi(t, \xi, x(t), y(t))], \quad t \geq 0$$

$$x(0) = x_0,$$

$$y(t) \in \operatorname{argmin}_{\mathbf{y}} \mathbb{E}[f(t, \xi, x(t), \mathbf{y})], \quad \text{s.t. } \mathbf{y} \in \mathcal{Y}(t, x(t)).$$

- Set $X = \mathbb{R}^n$ in (1).
- $\mathcal{Y}(t, x(t))$: nonempty closed convex set in $\mathbb{R}^m$ for every fixed $(t, x(t))$.
- $f(t, \xi, x(t), \cdot) : \mathbb{R}^m \rightarrow \mathbb{R}$ is a convex function.

If $f(t, \xi, x(t), \cdot)$ is differentiable, then $\Psi(t, \xi, x(t), \mathbf{y}) = \nabla_{\mathbf{y}} f(t, \xi, x(t), \mathbf{y})$

$$y(t) \in \operatorname{argmin}_{\mathbf{y}} \mathbb{E}[f(t, \xi, x(t), \mathbf{y})], \quad \text{s.t. } \mathbf{y} \in \mathcal{Y}(t, x(t))$$

$$\Leftrightarrow 0 \in \mathbb{E}[\Psi(t, \xi, x(t), y(t))] + \mathcal{N}_{\mathcal{Y}(t, x(t))}(y(t))$$

■ Joint work with Jian Guo and Guan Wang, SIOPT, second revision

$$\dot{x}(t) = \Pi_X \left(x(t) - \mathbb{E}_{P_t} [\Phi(t, \xi, x(t), y(t, \xi))] \right) - x(t), \quad t \geq 0, \quad (4)$$

$$x(0) = x_0, \quad (5)$$

$$y_i(t, \xi) \in \operatorname{argmin}_{y_i \in \mathcal{Y}_i(t, \xi)} f_i(t, \xi, x(t), y_i), \quad i = 1, \dots, k, \quad \text{a.e. } \xi \in \Xi, \quad (6)$$

$$y(t, \xi) = (y_1(t, \xi), \dots, y_k(t, \xi)), \quad \text{a.e. } \xi \in \Xi,$$

- $\mathbb{P}_{(\cdot)}(\cdot) : [0, T] \times \mathcal{F} \rightarrow [0, 1]$: a probability kernel
- $\mathbb{E}_{P_t}[\cdot]$ denotes expectation under P_t ,
- $\Phi(t, \xi, x(t), y(t, \xi)) = \Phi_1(t, \xi, x(t)) + \sum_{i=1}^k B_i(t, \xi, x(t)) y_i(t, \xi)$
- $\Phi_1 : \mathbb{R}_+ \times \mathbb{R}^\ell \times \mathbb{R}^n \rightarrow \mathbb{R}^n$
- $B_i : \mathbb{R}_+ \times \mathbb{R}^\ell \times \mathbb{R}^n \rightarrow \mathbb{R}^{n \times m_i}$ are continuous,
- $f_i : \mathbb{R}_+ \times \mathbb{R}^\ell \times \mathbb{R}^n \times \mathbb{R}^{m_i} \rightarrow \mathbb{R}$ is continuous and convex in y_i ,
- $\mathcal{Y}_i : [0, T] \times \Xi \rightrightarrows \mathbb{R}^{m_i}$ nonempty-closed-convex valued with $\mathcal{Y}_i(t, \xi) \subseteq \bar{\mathcal{Y}}_i$ (compact).

System model.

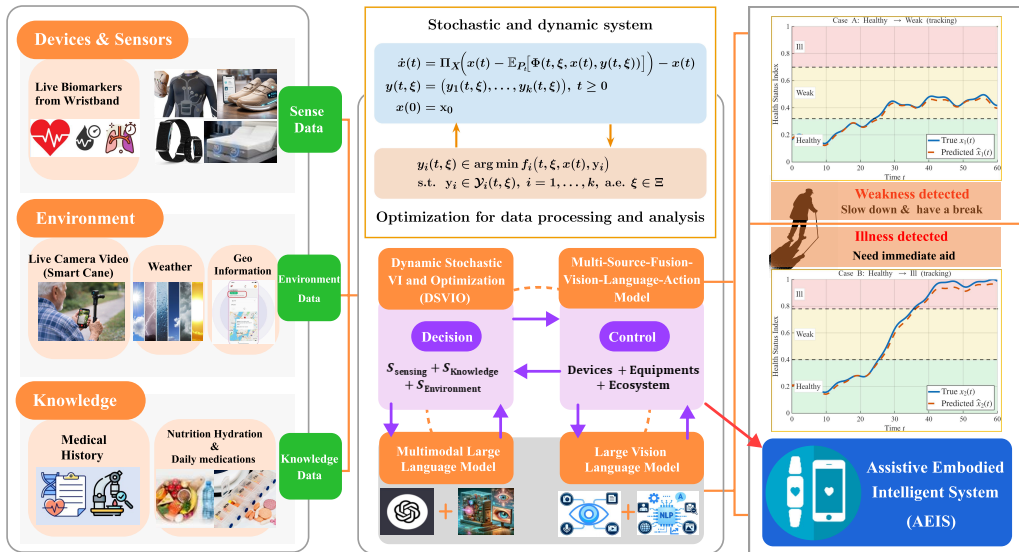
$$\begin{aligned}\dot{x}(t) &= \Pi_X \left(x(t) - \mathbb{E}_{P_t} [\Phi(t, \xi, x(t), y_1(t, \xi), \dots, y_k(t, \xi))] \right) - x(t), \quad t \geq 0, \\ x(0) &= x_0, \\ y_i(t, \xi) &\in \arg \min_{\mathbf{y}_i} f_i(t, \xi, x(t), \mathbf{y}_i), \quad \text{s.t. } \mathbf{y}_i \in \mathcal{Y}_i(t, \xi), \\ &\quad i = 1, \dots, k, \quad \text{a.e. } \xi \in \Xi.\end{aligned}$$

If $f_i(t, \xi, x(t), \cdot)$ is continuously differentiable, then

System model is equivalent to DSVI

$$\begin{aligned}\dot{x}(t) &= \Pi_X \left(x(t) - \mathbb{E}_{P_t} [\Phi(t, \xi, x(t), y_1(t, \xi), \dots, y_k(t, \xi))] \right) - x(t), \quad t \geq 0, \\ x(0) &= x_0, \\ 0 &\in \Psi_i(t, \xi, x(t), y_i(t, \xi)) + \mathcal{N}_{\mathcal{Y}_i(t, \xi)}(y_i(t, \xi)) \\ &\quad i = 1, \dots, k, \quad \text{a.e. } \xi \in \Xi.\end{aligned}$$

An Application of Embodied Intelligence System



Key Assumptions

A.1 For each $\mathcal{A} \in \mathcal{F}$, $t \mapsto \mathbb{P}_{(\cdot)}(\mathcal{A})$ is $\mathcal{B}_{[0,T]}$ -measurable; and for each $t \in [0, T]$, $\mathbb{P}_t(\cdot)$ is a probability measure on $(\Omega, \mathcal{F})$

A.2 There exist measurable functions $\kappa_{\Phi_1}, \kappa_{B_i} : [0, T] \times \Xi \rightarrow \mathbb{R}_+$, $i \in [k]$, whose expectations $\bar{\kappa}_{\Phi_1}(t) := \mathbb{E}_{P_t}[\kappa_{\Phi_1}(t, \xi)]$ and $\bar{\kappa}_{B_i}(t) := \mathbb{E}_{P_t}[\kappa_{B_i}(t, \xi)]$ are **integrable** on $[0, T]$, and for all $\mathbf{x}, \mathbf{x}' \in \mathbb{R}^n$,

$$\|\Phi_1(t, \xi, \mathbf{x}) - \Phi_1(t, \xi, \mathbf{x}')\| \leq \kappa_{\Phi_1}(t, \xi) \|\mathbf{x} - \mathbf{x}'\|$$

$$\|B_i(t, \xi, \mathbf{x}) - B_i(t, \xi, \mathbf{x}')\| \leq \kappa_{B_i}(t, \xi) \|\mathbf{x} - \mathbf{x}'\|.$$

A.3 There exists $\bar{\beta} > 0$ such that for $i \in [k]$,

$$\sup_{t \in [0, T]} \max \left\{ \|\mathbb{E}_{P_t}[\Phi_1(t, \xi, \mathbf{x}_0)]\|, \|\mathbb{E}_{P_t}[B_i(t, \xi, \mathbf{x}_0)]\| \right\} \leq \bar{\beta}.$$

A.4 The set-valued map $\mathcal{Y} : [0, T] \times \Xi \rightrightarrows \bar{\mathcal{Y}} := \bar{\mathcal{Y}}_1 \times \cdots \times \bar{\mathcal{Y}}_k \subset \mathbb{R}^m$ is **continuous**, where each $\bar{\mathcal{Y}}_i$ is compact and $\mathcal{Y}_i(t, \xi) \subseteq \bar{\mathcal{Y}}_i$ is nonempty closed-convex.

► Let $\mathcal{S}_i(t, \xi, \mathbf{x}) = \operatorname{argmin}_{\mathbf{y}_i \in \mathcal{Y}_i(t, \xi)} f_i(t, \xi, \mathbf{x}, \mathbf{y}_i)$.

► Define $\mathcal{S}(t, \xi, \mathbf{x}) := \mathcal{S}_1(t, \xi, \mathbf{x}) \times \cdots \times \mathcal{S}_k(t, \xi, \mathbf{x})$ and the set-valued map

$$F(t, \mathbf{x}) := \left\{ \mathbb{E}_{P_t} [\Phi(t, \xi, \mathbf{x}, \mathbf{y}(t, \xi))] : \mathbf{y}(t, \xi) \in \mathcal{S}(t, \xi, \mathbf{x}) \right\}.$$

► Then the DSVI-O can be viewed as a projected **differential inclusion**

$$\dot{\mathbf{x}}(t) \in H(t, \mathbf{x}(t)) := \{\Pi_X(\mathbf{x}(t) - \varphi) - \mathbf{x}(t) : \varphi \in F(t, \mathbf{x}(t))\}, \quad \mathbf{x}(0) = \mathbf{x}_0.$$

Lemma

*Under Assumptions **A.1–A.4**, the following statements hold.*

- (i) $\mathcal{S}(t, \xi, \mathbf{x})$ is **nonempty, convex and compact** for any fixed $t \geq 0$, $\xi \in \Xi$ and $\mathbf{x} \in \mathbb{R}^n$, and $\mathcal{S}$ is **u.s.c.**;
- (ii) $F(t, \mathbf{x})$ is **nonempty, convex and compact** for any fixed $t \geq 0$ and $\mathbf{x} \in \mathbb{R}^n$, $F(t, \cdot)$ is **u.s.c.** for any $t \geq 0$ and $F(\cdot, \mathbf{x})$ is **measurable** for any fixed $\mathbf{x} \in \mathbb{R}^n$.

► **Weak solution.** We call a pair $(x^* : [0, T] \rightarrow \mathbb{R}^n, y^* : [0, T] \times \Xi \rightarrow \mathbb{R}^m)$ a **weak solution** of (4)–(6) if x^* is **absolutely continuous** on $[0, T]$ and

$$x^*(t) = x_0 + \int_0^t \left(\Pi_X(x^*(\tau) - \mathbb{E}_{P_\tau}[\Phi(\tau, \xi, x^*(\tau), y^*(\tau, \xi))]) - x^*(\tau) \right) d\tau,$$

and y^* is **measurable** in ξ and **integrable** in t :

$$\int_0^T \mathbb{E}_{P_t}[\|y^*(t, \xi)\|] dt < \infty, \quad y_i^*(t, \xi) \in \mathcal{S}_i(t, \xi, x^*(t)), \quad i \in [k].$$

■ Theorem (**Existence of a weak solution**)

► Under Assumptions **A.1–A.4**, if $H(t, \mathbf{x})$ is a convex set for any $t > 0, \mathbf{x} \in \mathbb{R}^n$, then for any $T > 0$, the DSVI-O (4)–(6) has at least one **weak solution** (x^*, y^*) on $[0, T]$.

Discretization (Forward Euler) & Convergence

■ **Time grid.** ► Let $h = T/N$, $t_\nu = \nu h$, $\nu = 0, 1, \dots, N$.

■ **Forward Euler scheme.**

$$\begin{cases} \mathbf{x}_{\nu+1}^N = \mathbf{x}_\nu^N + h \left(\Pi_X(\mathbf{x}_\nu^N - \mathbb{E}_{P_{t_\nu}}[\Phi(t_\nu, \xi, \mathbf{x}_\nu^N, \mathbf{y}_\nu^N(\xi))]) - \mathbf{x}_\nu^N \right), \\ y_{\nu,i}^N(\xi) \in \underset{\mathbf{y}_i \in \mathcal{Y}_i(t_\nu, \xi)}{\operatorname{argmin}} f_i(t_\nu, \xi, \mathbf{x}_\nu^N, \mathbf{y}_i), \quad i \in [k], \quad \mathbf{y}_\nu^N(\xi) := (y_{\nu,1}^N(\xi), \dots, y_{\nu,k}^N(\xi)) \end{cases} \quad (7)$$

■ **Piecewise linear interpolation.**

$$x^N(t) = \frac{t_{\nu+1} - t}{h} \mathbf{x}_\nu^N + \frac{t - t_\nu}{h} \mathbf{x}_{\nu+1}^N, \quad t \in [t_\nu, t_{\nu+1}], \quad \nu = 0, \dots, N-1. \quad (8)$$

■ **Theorem (Euler subsequence convergence)**

► Under Assumptions **A.1–A.4**, every sequence $\{\mathbf{x}^N\}_{N \in \mathbb{N}}$ defined in (8) has a **uniformly convergent subsequence** $\{\mathbf{x}^{N_\tau}\}_{\tau \in \mathbb{N}}$ on $[0, T]$ with a limit $\mathbf{x}$ such that

$$\sup_{t \in [0, T]} \|\mathbf{x}^{N_\tau}(t) - \mathbf{x}(t)\| \rightarrow 0 \quad \text{as } \tau \rightarrow \infty,$$

and the limit $\mathbf{x}$ is the **x-part of a weak solution** of (4)–(6) on $[0, T]$.

SAA–Forward Euler Scheme and Convergence

■ **Sampling.** ► Let $\{\xi_\nu^j\}_{j=1}^J$ be iid samples from P_{t_ν} for each ν .

■ **SAA–Euler scheme.** ► For $i \in [k]$, $j \in [J]$,

$$\begin{cases} \mathbf{x}_{\nu+1}^{N,J} = \mathbf{x}_\nu^{N,J} + h \left(\Pi_X \left(\mathbf{x}_\nu^{N,J} - \frac{1}{J} \sum_{j=1}^J \Phi(t_\nu, \xi_\nu^j, \mathbf{x}_\nu^{N,J}, y_\nu^{N,J}(\xi_\nu^j)) \right) - \mathbf{x}_\nu^{N,J} \right), \\ y_{\nu,i}^{N,J}(\xi_\nu^j) \in \arg \min_{\mathbf{y}_i \in \mathcal{Y}_i(t_\nu, \xi_\nu^j)} f_i(t_\nu, \xi_\nu^j, \mathbf{x}_\nu^{N,J}, \mathbf{y}_i), \end{cases} \quad (9)$$

► where $y_\nu^{N,J}(\xi_\nu^j) := (y_{\nu,1}^{N,J}(\xi_\nu^j), \dots, y_{\nu,k}^{N,J}(\xi_\nu^j))$.

■ Theorem (SAA–Euler subsequence convergence)

► Under Assumptions **A.1–A.4**, every sequence $\{x^{N,J}\}_{N,J \in \mathbb{N}}$ in (9) has a **uniformly convergent subsequence** $\{x^{N_j, J_j}\}_{j \in \mathbb{N}}$ on $[0, T]$, with $N_j \rightarrow \infty$ and $J_j \rightarrow \infty$, to an absolutely continuous x such that for any $r > 0$,

$$\dot{x}(t) \in \text{conv} \left(H^r(t, x(t)) \right), \quad \text{a.e. } t \in [0, T],$$

where $H^r(t, x(t)) := \bigcup_{\mathbf{x}' \in \mathbb{B}(x(t), r)} H(t, \mathbf{x}')$.

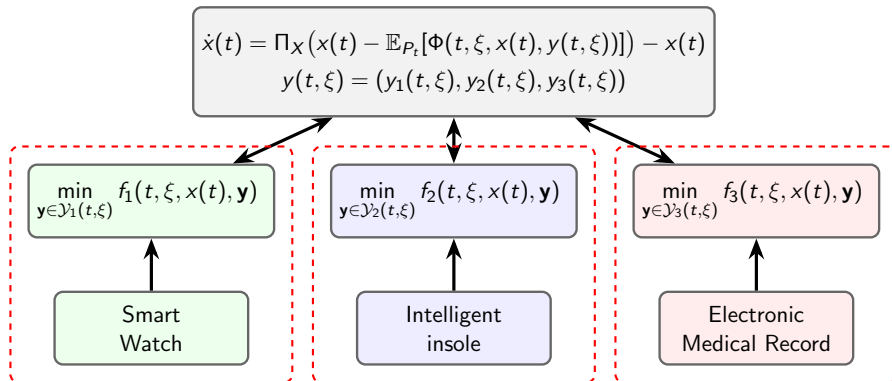


Figure: DSVI-O workflow for personalized elderly health recommendations

Real-time monitoring of elderly health quality

■ **Goal:** Real-time monitoring and forecasting the health-quality trajectory $x(t)$ of elderly individuals.

■ **Model:** A time-varying-parameter projected ODE

$$\dot{x}(t) = \Pi_X \left(x(t) - \underbrace{A(t)x(t) + \mathbb{E}_{P_t} \left[q(t, \xi) + \ell(t, \xi) \right]}_{\text{time-varying driving term}} \right) - x(t), \quad (10)$$

► Components.

- **Smooth clinical platform:** $\ell(t, \xi) = \ell_0 + \Delta \ell_1(t) m(t, \xi)$
- **Periodic rhythm:** $q(t, \xi) = \sum_{i=1}^3 A_i(t) \sin(2\pi \zeta_i t + \varphi_i(t)) + \varepsilon_1(t, \xi)$
- **Latent clinical score:** $m(t, \xi) = \sum_{i=1}^3 \alpha_i(t) B_i(t, \xi) y_i(t, \xi) + \varepsilon_2(t, \xi)$

► Notation.

- | | | | |
|-----------------|-----------------------------|---------------------------|-------------------------------|
| • $A(t)$ | physiological decay | • $x(t)$ | estimated health state |
| • $B_i(t, \xi)$ | observed physiological data | • $\alpha_i(t)$ | fusion weight for channel i |
| • $y_i(t, \xi)$ | physiological parameter | • $\varepsilon_i(t, \xi)$ | random clinical fluctuation |

- **Second-stage estimation (DSVI-O).** We find an optimal solution $y_i(t, \xi)$ in the second stage of DSVI-O. For $i = 1, 2, 3$,

$$y_i(t, \xi) \in \underset{\mathbf{y}_i \in [a_i(\xi), b_i(\xi)]}{\operatorname{argmin}} f_i(t, \xi, x(t), \mathbf{y}_i)$$

- **Objective.**

$$f_i(t, \xi, x(t), \mathbf{y}_i) = \|H_i(t, \xi)\mathbf{y}_i + N_i(t, \xi)x(t) - c_i(t, \xi)\|^2 + \rho_i \|B_i(t, \xi)\mathbf{y}_i - x(t)\|^2.$$

- **Data.**

$$\begin{aligned} B_i : \mathbb{R}_+ \times \Xi &\rightarrow \mathbb{R}^{n \times m_i}, & H_i : \mathbb{R}_+ \times \Xi &\rightarrow \mathbb{R}^{n_i \times m_i}, \\ c_i : \mathbb{R}_+ \times \Xi &\rightarrow \mathbb{R}^{n_i}, & N_i : \mathbb{R}_+ \times \Xi &\rightarrow \mathbb{R}^{n_i \times n}, \\ a_i, b_i : \Xi &\rightarrow \mathbb{R}^{m_i}, & \mathbf{a}_i \leq a_i(\xi) < b_i(\xi) \leq \mathbf{b}_i, & \rho_i > 0. \end{aligned}$$

Assumptions and solution properties

$$\begin{aligned}\dot{x}(t) &= \Pi_X\left(x(t) - A(t)x(t) + \mathbb{E}_{P_t}\left[q(t, \xi) + \ell(t, \xi)\right]\right) - x(t), \quad t \geq 0, \\ x(0) &= x_0, \\ y_i(t, \xi) &\in \operatorname{argmin}_{y_i \in [a_i(\xi), b_i(\xi)]} f_i(t, \xi, x(t), y_i), \quad \text{a.e. } \xi \in \Xi.\end{aligned}$$

► **Assumptions.** (i) A is continuous on $[0, T]$, and q and ℓ are continuous on $[0, T] \times \Xi$.

(ii) $\mathbb{E}_{P_t}[\|q(t, \xi)\|]$ and $\mathbb{E}_{P_t}[\|\ell(t, \xi)\|]$ are uniformly bounded in $t \in [0, T]$.

(iii) The functions $a_i, b_i : \Xi \rightarrow \mathbb{R}^{m_i}, i = 1, 2, 3$ are continuous.

Assumptions (i)-(iii) $\Rightarrow$ Assumptions (A.2)-(A.4)

★ Under Assumptions (A.1) and (i)–(iii), the DSVI-O has a weak solution (x, y) on $[0, T]$ for any $T > 0$, satisfying

$$\begin{aligned}x(t) &= x_0 + \int_0^t \left[\Pi_X\left(x(\tau) - A(\tau)x(\tau) + \mathbb{E}_{P_\tau}[q(\tau, \xi) + \ell(\tau, \xi)]\right) - x(\tau) \right] d\tau, \\ y_i(t, \xi) &\in \operatorname{argmin}_{y_i \in [a_i(\xi), b_i(\xi)]} f_i(t, \xi, x(t), y_i), \quad \text{a.e. } \xi \in \Xi, \quad \int_0^T \mathbb{E}_{P_t}[\|y(t, \xi)\|^2] dt < \infty.\end{aligned}$$

Preliminary numerical experiments

■ **Simulation setting.** Simulate the evolution of the health status of 10 persons over 10 days by using their health data in the past 90 days. Within every 24-hour cycle, measurements were sampled at 5-second intervals, yielding 17,280 time points per day.

► **Smart watch.** 14 features in $17280 \times 5s$:

$$Z_1(t_\nu, \xi^j) \in \mathbb{R}^{1 \times 14}, \quad \eta_1(t_\nu, \xi^j) \in \mathbb{R}, \quad \nu \in [17280], \quad j \in [90]$$
$$B_1(t_\nu, \xi^{90+j}) \in \mathbb{R}^{1 \times 14}, \quad \pi_1(t_\nu, \xi^j) \in \mathbb{R}, \quad \nu \in [17280], \quad j \in [10]$$

► **Intelligent insole.** 17 features in $17280 \times 5s$:

$$Z_2(t_\nu, \xi^j) \in \mathbb{R}^{1 \times 17}, \quad \eta_2(t_\nu, \xi^j) \in \mathbb{R}, \quad \nu \in [17280], \quad j \in [90]$$
$$B_2(t_\nu, \xi^{90+j}) \in \mathbb{R}^{1 \times 17}, \quad \pi_2(t_\nu, \xi^j) \in \mathbb{R}, \quad \nu \in [17280], \quad j \in [10]$$

► **Electronic Medical Records.** 30 features including white blood cell count, glucose, blood pressure, liver and kidney function, etc.

$$Z_3(\xi^j) \in \mathbb{R}^{1 \times 30}, \quad \eta_3(\xi^j) \in \mathbb{R}, \quad j \in [90]$$
$$B_3(\xi^{90+j}) \in \mathbb{R}^{1 \times 30}, \quad \pi_3(\xi^j) \in \mathbb{R}, \quad j \in [10]$$

■ Estimation of biological parameter.

► At each time t_ν , estimate $y_i(t_\nu, \xi)$ by sampling 50 days from $Z_i(t_\nu, \xi^j)$, $\eta_i(t_\nu, \xi^j)$, $j = 1, \dots, 90$, repeated 20 times:

$$H_i(t_\nu, \bar{\xi}^\ell) \in \mathbb{R}^{50 \times m_i}, \quad c_i(t_\nu, \bar{\xi}^\ell) \in \mathbb{R}^{50}, \quad \ell = 1, \dots, 20, \quad (m_1, m_2, m_3) = (14, 17, 30)$$

■ **Correction matrix.** Define $N_i(t_\nu, \bar{\xi}^\ell)$ using softmax-normalized Gaussian similarities over pairwise distances:

$$N_i^j(t_\nu, \bar{\xi}^\ell) = \lambda_i \left(\mathbf{1}_{50} - \text{softmax} \left(-\frac{1}{h_i^2} | \mathbf{1}_{50} B_i(t_\nu, \xi^{90+j}) - H_i(t_\nu, \bar{\xi}^\ell) | \mathbf{1}_{m_i} \right) \right)$$

► **Bandwidth:** h_i is the median pairwise distance in $H_i(t_\nu, \bar{\xi}^\ell)$:

$$h_i = \text{median} \{ \| (H_i)_{[r,:]} - (H_i)_{[s,:]} \| \mid 1 \leq r < s \leq 50 \}$$

► **Softmax:** The softmax of a vector $\mathbf{a} \in \mathbb{R}^{50}$ is defined as

$$(\text{softmax}(\mathbf{a}))_i = \frac{e^{a_i}}{\mathbf{1}_{50}^\top e^{\mathbf{a}}}, \quad i = 1, \dots, 50.$$

■ Correction label

- ▶ Historical labels c_i may not match today's state $\mathbf{x}_\nu$ due to variation.
- ▶ Adjust label:

$$N_i^j(t_\nu, \bar{\xi}^\ell) \mathbf{x}_\nu^j - c_i(t_\nu, \bar{\xi}^\ell)$$

- ▶ Similar days $\Rightarrow$ small correction; different days $\Rightarrow$ large correction.

■ Parameters

- ▶ $\theta_1 = 0.50$, $\theta_2 = 1.50$, $\sigma = 0.50$, $\zeta_1 = 0.025$, $\zeta_2 = 0.021$, $l_0 = 0.02$,
 $\lambda_i = 0.50$, $\rho_i = 5$, $i = 1, 2, 3$,
 $A_1(t) \equiv 0.15$, $A_2(t) \equiv 0.17$, $\Delta \ell_1(t) \equiv 1.5$.

■ Initial value

- ▶ $\mathbf{x}_0 = \pi_1(0, \xi^{90+j})$, $j \in [10]$.

■ Label: health=0; weak=1, ill=2 in c_i, π_i , $i = 1, 2, 3$.

Simulation results of DSVI

■ Consider a person in 10 days, for $j = 1, \dots, 10$, $\nu = 0, \dots, 17279$:

► **System dynamics.** The system evolves according to:

$$\mathbf{x}_{\nu+1}^j = \Pi_{[0,2]} \left(\mathbf{x}_{\nu}^j - A(t_{\nu}) \mathbf{x}_{\nu}^j + \frac{1}{20} \sum_{\ell=1}^{20} (q(t_{\nu}, \bar{\xi}^{\ell}) + \ell^j(t_{\nu}, \bar{\xi}^{\ell})) \right)$$

► **Optimization.** The optimization solutions $\mathbf{y}_i^j(t_{\nu}, \bar{\xi}^{\ell})$ for $i = 1, 2, 3$ are computed by:

$$\mathbf{y}_i^j(t_{\nu}, \bar{\xi}^{\ell}) \in \underset{\mathbf{y}_i \in [-10, 10]^{m_i}}{\operatorname{argmin}} \left\| H_i(t_{\nu}, \bar{\xi}^{\ell}) \mathbf{y}_i - N_i^j(t_{\nu}, \bar{\xi}^{\ell}) \mathbf{x}_{\nu}^j - c_i(t_{\nu}, \bar{\xi}^{\ell}) \right\|^2 \\ + \rho_i \| B_i(t_{\nu}, \bar{\xi}^{90+j}) \mathbf{y}_i - \mathbf{x}_{\nu}^j \|^2.$$

► **Prediction rule.** The health state predictions are obtained by:

$$\operatorname{Pred}^j_{\nu}(\ell) = \begin{cases} 0, & \text{if } \mathbf{x}_{\nu}^j(\ell) \leq 2/3, \\ 1, & \text{if } 2/3 < \mathbf{x}_{\nu}^j(\ell) < 4/3, \\ 2, & \text{if } \mathbf{x}_{\nu}^j(\ell) \geq 4/3. \end{cases}$$

Simulation results: Individual-wise Performance

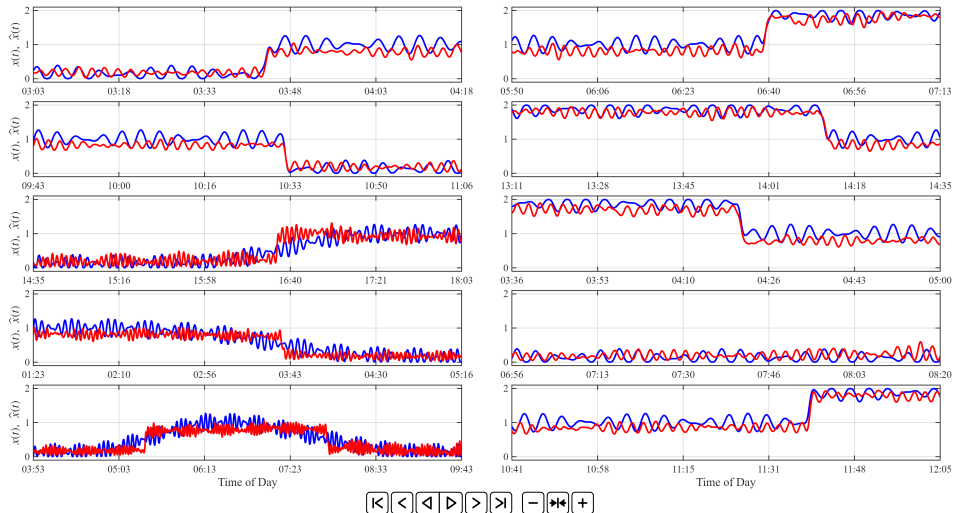


Figure: Tracked person's health curve: Estimated (red) $\hat{x}(t)$ vs True (blue) $x(t)$; 0: healthy, 1: weak, 2: ill.

Simulation results: Individual-wise Performance

■ Use $\text{Pred } j$ to estimate the label trajectory over the day, and compare it with ground-truth labels to generate the confusion matrix below.

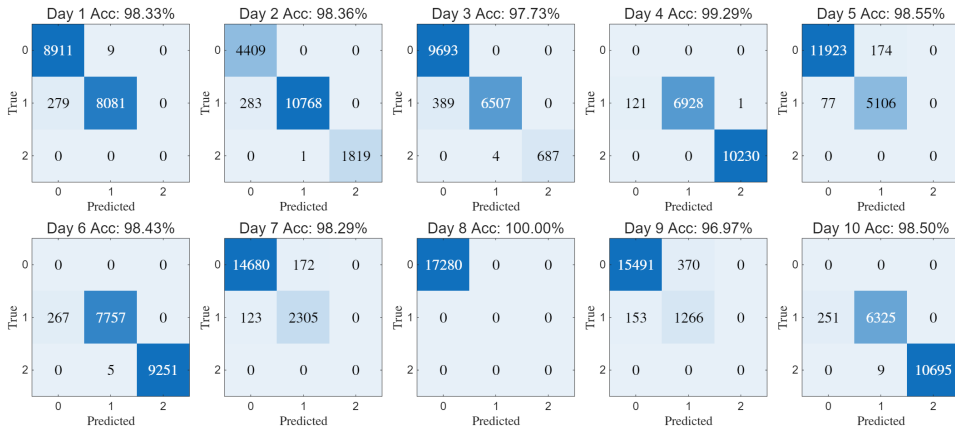


Figure: Confusion matrix comparing true and predicted labels for a person in 10 days.

Simulation results: Overall Performance

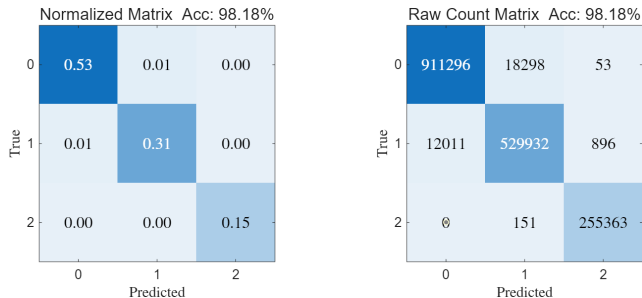


Figure: Confusion matrix of true and predicted labels for all ten persons.

Table: Class-wise Evaluation Results

Class	Precision	Recall	Specificity	F1 Score
healthy (0)	0.9870	0.9803	0.9850	0.9836
weak (1)	0.9664	0.9762	0.9844	0.9713
ill (2)	0.9963	0.9994	0.9994	0.9979

It Gathers Clues From Everywhere!

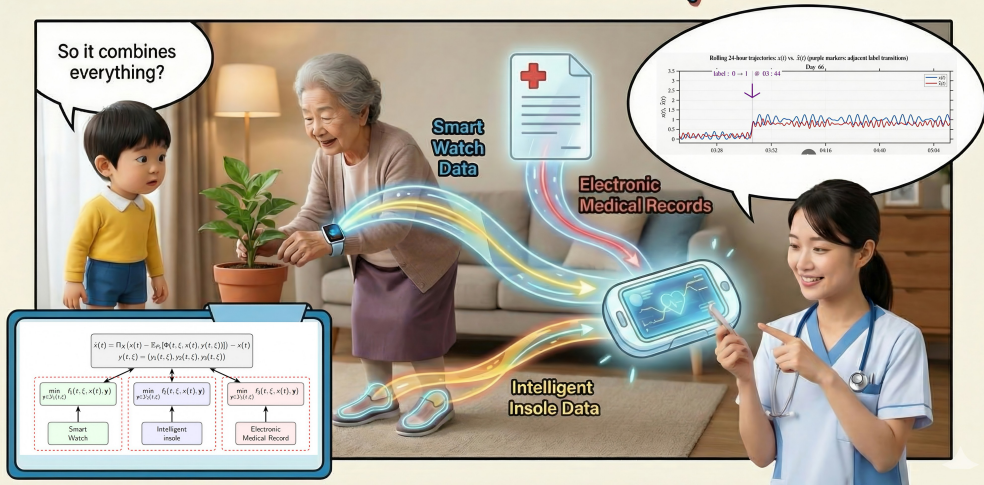


Figure: Dynamic stochastic VI with parametric optimization

Embodied Intelligence System based on DSVI-O for Elderly

Acknowledgement for the grants from

**Otto Poon Charitable Foundation
Smart Cities Research Institute**



Otto Poon Charitable Foundation
Smart Cities Research Institute
潘樂陶慈善基金智慧城市研究院

CITIC TELECOM CPC



CITIC TELECOM CPC

Summary

- ① Differential Stochastic Variational Inequalities with Parametric Optimization (DSVI-O) is introduced, which includes DVI and two-stage SVI as special cases.
- ② Fundamental solution properties of DSVI-O are studied.
- ③ Numerical approximation of DSVI-O and convergence analysis are discussed.
- ④ Applications of DSVI-O in AI are discussed.

Future Research Topics

- ① Efficient algorithms for solving DSVI-O
- ② More applications of DSVI-O in AI

Thank You