Transversal Passivity: Dissipation Relative to Input and Output Manifolds

Feng-Yu Yue Daniel Zelazo
The Stephen B. Klein Faculty of Aerospace Engineering,
Technion–Israel Institute of Technology, Haifa, Israel
fengyu.yue@campus.technion.ac.il, dzelazo@technion.ac.il

arXiv:2611.06251v1 [eess.SY] 4 Nov 2026

Abstract

This paper introduces transversal passivity, a local dissipation property for regulating input and output signals to prescribed manifolds. Its supply rate is built from the transverse port errors induced by nearest-point projections, so that motion along the manifolds is not penalized. We give a necessary condition and a constructive sufficient condition for the property and relate it to existing passivity notions. For feedback interconnections through negative-adjoint linear channels, power conservation no longer cancels the transverse supply. The residual is governed by compatibility defects that depend only on the channels and the prescribed manifolds, so it can be assessed before any closed-loop analysis, and it falls into a hierarchy of exact compatibility, quadratic cancellation, and nonzero quadratic residuals. Since transversal storages are only positive semidefinite, port convergence is obtained by combining compact confinement, for instance by conserved quantities, with a dissipation margin, and we identify a geometric obstruction under which no quadratic margin is available near the feasible states. A federated learning case study with heterogeneous data shows that port convergence can be certified even though no server target is compatible with the interconnection, and that the obstruction arises for a natural choice of target.

Keywords: Differential geometric methods, interconnected systems, nonlinear systems, passivity theory.

1 Introduction

We consider control problems in which input and output signals are required to approach prescribed manifolds, rather than prescribed points. Such problems arise in many applications, including path-following [1], control allocation [2], formation control [3], synchronization [4], and the training of overparameterized learning models, whose interpolating solutions form a manifold [5]. This class of problems leaves a geometric degree of freedom in the control design – we need only penalize the transverse deviations of the signals from the target manifolds while leaving tangential motion along the manifolds unconstrained. This motivates an input-output formulation, where the deviations are measured without imposing a tangential reference.

Passivity theory provides a natural starting point for such a formulation. A passive system is characterized by a storage function whose derivative is bounded from above by a supply rate defined on the system inputs and outputs. This input-output characterization relates passivity to Lyapunov theory and finite-gain input–output stability [6]. A further advantage is compositionality, which guarantees that the feedback interconnection of two passive systems is passive, and the sum of the individual storage functions is a storage function of the interconnection [7].

Existing passivity variants reference the supply rate to points or trajectories; shifted passivity, for instance, uses a single input-output pair, which is not available a priori when the targets form a manifold and the limit depends on the initial condition. To adapt passivity to manifold-target problems, we introduce a new notion termed *transversal passivity*, which is a passivity property defined locally with respect to prescribed input and output manifolds. Its storage function is continuously differentiable and positive semidefinite, and its supply rate measures only the finite transverse errors induced by nearest-point projections of the inputs and outputs onto these manifolds.

Some related approaches retain this separation in the state space, but represent and constrain the transverse deviations differently. Transverse contraction [8] imposed contraction on infinitesimal state variations transverse to the system flow. Transverse exponential stability [9] required exponential decay of the transverse state coordinates about an invariant manifold. Local transverse feedback linearization [10] used local coordinates and regular static feedback so that the dynamics transverse to a controlled-invariant manifold became linear and controllable. Immersion and invariance [11] and nonlinear output regulation [12] also target invariant manifolds, but as design methods for a given closed loop: they construct a controller that renders a manifold invariant and attractive, rather than a port property of the subsystems that is pre-

served under interconnection. Closer in spirit, passivity of input–output relations with respect to submanifolds [13] measures the distances of the input and output signals to prescribed submanifolds, but only at the signal level, without a storage function, and for the standard negative feedback interconnection.

We study the symmetric feedback interconnection of two transversally passive systems through linear channels, in which the reverse channel is the negative adjoint of the forward channel; standard negative feedback is the special case in which the forward channel is the identity. The interconnection is power-conserving [14], i.e., the net power through the ports vanishes, and this cancellation underpins passivity-based closed-loop analysis [15, 16], including output synchronization [17] and consensus on directed graphs [18, 19]. In these analyses the supply rates are referenced to points or to the other agents, so the target set enters only through the storage or the interconnection. The first difficulty is that, for transversal passivity, the relevant inner product is instead that of the transverse port errors, which power conservation does not force to vanish. A second difficulty is that transversal storages are local and only positive semidefinite, so their sublevel sets need not be compact and dissipation alone does not yield convergence of the ports; related issues arise in the passivity-based stabilization of noncompact sets, studied through reduction principles and set detectability [20].

The paper makes three contributions. (i) We introduce transversal passivity, give a necessary condition and a constructive sufficient condition for it, and show that it recovers shifted and incremental passivity as special cases and connects to differential passivity at equilibria. (ii) We show that, under negative-adjoint feedback, transversal passivity composes up to a residual supply. The residual is governed by compatibility defects, which measure whether the channels preserve the transverse geometry and depend only on the channels and the prescribed manifolds, so it can be checked before any closed-loop analysis. We classify this residual into three cases: exact compatibility; quadratic cancellation, in which the channels need not preserve the transverse geometry but the residual vanishes to second order; and a nonzero quadratic residual. (iii) We give conditions for port convergence that combine compact confinement, for instance from a conserved quantity whose level set is fixed by the initial condition, with a dissipation margin, and we identify a geometric obstruction under which no quadratic margin is available near the feasible states. A federated learning case study with heterogeneous client data shows that these results certify port convergence even though no server target is compatible with the interconnection, and that the obstruction arises for a natural choice of target.

The remainder of this paper is organized as follows. Section 2 collects preliminaries, Section 3 develops transversal passivity, Section 4 treats feedback interconnection and the residual supply, Section 5 establishes port convergence, Section 6 presents the federated learning case study, and Section 7 concludes the paper.

Notation

$\mathbb{R}_{\geq 0}$ and $\mathbb{R}_{>0}$ denote the nonnegative and positive reals, $\langle \cdot, \cdot \rangle$ the Euclidean inner product, and $\| \cdot \|$ the induced norm; for linear and multilinear maps, $\| \cdot \|$ is the induced operator norm. A product space carries the norm $\|(a_1, a_2)\| = (\|a_1\|^2 + \|a_2\|^2)^{1/2}$. We write I_n for the $n \times n$ identity, $\mathbb{1}_N$ for the vector of N ones, $\otimes$ for the Kronecker product, $A^\top$ for the transpose, and $\ker A$ for the kernel of a matrix A; span denotes linear span. For a symmetric matrix P, $P \succ 0$ ($P \succeq 0$) means P is positive definite (semidefinite). The unit sphere in $\mathbb{R}^n$ is $\mathbb{S}^{n-1}$, and id is the identity map. For a linear subspace $M \subseteq \mathbb{R}^n$, P_M is the orthogonal projection onto M. For a set S, $\overline{S}$ is its closure, $\mathrm{dist}(x, S) := \inf_{y \in S} \|x - y\|$ is the distance from x to S, and $h^{-1}(S)$ is the preimage of S under a map h, which need not be invertible. C^k denotes the class of k times continuously differentiable maps. If ρ is differentiable at x, its differential is the linear map $\mathrm{d}_x\rho$, acting on v as $\mathrm{d}_x\rho[v]$; if ρ is twice differentiable, $\mathrm{d}_x^2\rho$ is the symmetric bilinear second differential, acting as $\mathrm{d}_x^2\rho[v, w]$. For a scalar function V, ∇V and $\nabla^2 V$ denote its gradient and Hessian. Landau symbols $o(\cdot)$ and $O(\cdot)$ refer to the limit of their argument tending to zero.

2 Preliminaries

This section introduces the definitions and background material used throughout the paper, including notions from passivity theory and differential geometry.

2.1 Standard Passivity

Consider $\dot{x} = f(x, u)$, $y = h(x, u)$, with $x \in \mathbb{R}^n$ and $u, y \in \mathbb{R}^m$. A passivity inequality is certified by a continuously differentiable storage $Q : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ and indices $\varepsilon, \gamma \geq 0$ satisfying

$$\mathrm{d}_x Q[f(x, u)] \leq u^\top y - \varepsilon \|y\|^2 - \gamma \|u\|^2 \tag{1}$$

for every (x, u) [6]. The system is *passive* if $\varepsilon = \gamma = 0$, *output-strictly passive* if $\varepsilon > 0$ and $\gamma = 0$, and *input-strictly passive* if $\varepsilon = 0$ and $\gamma > 0$. The right-hand side of (1) is called the supply rate [7]. We will discuss other passivity variants in Subsection 3.3.

2.2 Tubular Neighborhoods and Nearest-Point Projections

We employ basic notions from differential geometry in this work. Let $\mathcal{M}$ be a smoothly embedded submanifold of $\mathbb{R}^n$. Denote its tangent space at $p \in \mathcal{M}$ by $T_p\mathcal{M}$ and its normal space by $N_p\mathcal{M} := (T_p\mathcal{M})^\perp$.

For a continuous positive function $\epsilon : \mathcal{M} \to \mathbb{R}_{>0}$, define the *$\epsilon$-neighborhood* of $\mathcal{M}$ by [21]

$$\mathcal{N} := \{z \in \mathbb{R}^n : \|z - p\| < \epsilon(p) \text{ for some } p \in \mathcal{M}\}.$$

The ϵ-neighborhood theorem [21, Chapter 2] guarantees that ϵ can be chosen so that every $z \in \mathcal{N}$ has a unique nearest point on $\mathcal{M}$. If $\mathcal{M}$ is compact, then ϵ can be taken as a positive constant. Therefore, the associated nearest-point projection

$$r : \mathcal{N} \to \mathcal{M}, \quad r(z) := \arg\min_{p \in \mathcal{M}} \|z - p\|^2, \tag{2}$$

is well-defined and smooth. For every $z \in \mathcal{N}$, define the transverse error map by

$$e(z) := z - r(z). \tag{3}$$

Then $z = r(z) + e(z)$ and $e(z) \in N_{r(z)}\mathcal{M}$ is the normal displacement of z from its base point $r(z) \in \mathcal{M}$. Moreover, the nearest-point property gives

$$\|e(z)\| = \|z - r(z)\| = \operatorname{dist}(z, \mathcal{M}). \tag{4}$$

With ϵ chosen as above, $\mathcal{N}$ is a tubular neighborhood of $\mathcal{M}$ and r is the associated smooth orthogonal tubular retraction [22]; we use the two terms interchangeably.

Now, we summarize some standard properties of the nearest-point projection, which are fundamental to the subsequent discussion.

Lemma 1 ([23]). *Consider a submanifold $\mathcal{M}$, tubular neighborhood $\mathcal{N}$, nearest-point projection r, and transverse error map e. Define $V_{\mathcal{M}}(z) = \frac{1}{2}\|e(z)\|^2$. Then the following hold:*

(i) For every $z \in \mathcal{M}$, the projection fixes z and the transverse error vanishes,

$$r(z) = z, \quad e(z) = 0. \tag{5}$$

Moreover, the corresponding differentials of r and e satisfy

$$\mathrm{d}_z r = \mathrm{P}_{T_z\mathcal{M}}, \text{ and } \mathrm{d}_z e = \mathrm{P}_{N_z\mathcal{M}}. \tag{6}$$

(ii) For every $z \in \mathcal{N}$ and $v \in \mathbb{R}^n$, the derivative of $V_{\mathcal{M}}$ at z in the direction v is given by

$$\mathrm{d}_z V_{\mathcal{M}}[v] = \langle e(z),\, v\rangle, \tag{7}$$

since $\mathrm{d}_z r[v] \in T_{r(z)}\mathcal{M}$ is orthogonal to $e(z) \in N_{r(z)}\mathcal{M}$.

3 Transversal Passivity

In this section, we introduce transversal passivity, a state-space certificate for the input and output behavior relative to prescribed manifolds. We then establish a necessary condition and a constructive sufficient condition for this property, and relate it to other established passivity variants.

Consider the nonlinear system,

$$\Sigma : \quad \dot{x} = f(x, u), \quad y = h(x, u), \tag{8}$$

where $x \in \mathcal{X} \subseteq \mathbb{R}^n$, $u \in \mathcal{U} \subseteq \mathbb{R}^m$, and $y \in \mathcal{Y} \subseteq \mathbb{R}^m$ denote the state, input, and output, respectively, of Σ; we refer to u and y as the ports of Σ. Here, $\mathcal{X}, \mathcal{U}$ and $\mathcal{Y}$ are open sets, and $f : \mathcal{X} \times \mathcal{U} \to \mathbb{R}^n$ and $h : \mathcal{X} \times \mathcal{U} \to \mathbb{R}^m$ are continuously differentiable.

We are interested in control problems where the input and output signals are required to approach prescribed input manifolds $\mathcal{M}_u \subseteq \mathcal{U}$ and output manifolds $\mathcal{M}_y \subseteq \mathcal{Y}$. In such problems, the signals are allowed to move persistently along the manifolds, while their deviations from the manifolds are required to decay. This motivates a notion of passivity whose supply rate measures only the transverse errors of these signals.

In this direction, let $\mathcal{M}_u \subseteq \mathcal{U}$ and $\mathcal{M}_y \subseteq \mathcal{Y}$ be embedded submanifolds. Following Subsection 2.2, choose tubular neighborhoods $\mathcal{N}_u \subseteq \mathcal{U}$ and $\mathcal{N}_y \subseteq \mathcal{Y}$ on which the associated nearest-point projections r_u, r_y are uniquely defined and smooth. The corresponding transverse port error maps are

$$e_u(u) = u - r_u(u) \text{ and } e_y(y) = y - r_y(y).$$

By (4), these maps satisfy $\|e_u(u)\| = \operatorname{dist}(u, \mathcal{M}_u)$ and $\|e_y(y)\| = \operatorname{dist}(y, \mathcal{M}_y)$.

Since the transverse error maps are defined only on their respective tubular neighborhoods, a state-input pair (x, u) can be considered only when both $u \in \mathcal{N}_u$ and the corresponding output $h(x, u) \in \mathcal{N}_y$. This motivates the tubular domain

$$\mathcal{D} := \{(x, u) \in \mathcal{X} \times \mathcal{U} : u \in \mathcal{N}_u,\ h(x, u) \in \mathcal{N}_y\}. \tag{9}$$

The domain $\mathcal{D}$ is open because h is continuous and $\mathcal{N}_u, \mathcal{N}_y$ are open.

We now define transversal passivity on the tubular domain $\mathcal{D}$. This notion provides a state-space certificate for input-output behavior relative to the prescribed manifolds.

Definition 1 (Transversal Passivity). *Consider the system (8), and assume that the tubular domain $\mathcal{D}$ in (9) is nonempty. If there exist a continuously differentiable storage function $V : \mathcal{X} \to \mathbb{R}_{\geq 0}$ and constants $\varepsilon, \gamma \geq 0$ such that, for every $(x, u) \in \mathcal{D}$,*

$$\begin{aligned}\mathrm{d}_x V[f(x, u)] \leq \langle e_u(u), e_y(h(x, u))\rangle - \varepsilon\|e_y(h(x, u))\|^2 \\ - \gamma\|e_u(u)\|^2,\end{aligned} \tag{10}$$

then the system is called

- *$(\mathcal{M}_u, \mathcal{M}_y)$-passive if $\varepsilon = \gamma = 0$;*

- *output-strictly* $(\mathcal{M}_u, \mathcal{M}_y)$*-passive if* $\varepsilon > 0$ *and* $\gamma = 0$*;*
- *input-strictly* $(\mathcal{M}_u, \mathcal{M}_y)$*-passive if* $\varepsilon = 0$ *and* $\gamma > 0$*;*
- *input-output strictly* $(\mathcal{M}_u, \mathcal{M}_y)$*-passive if* $\varepsilon > 0$ *and* $\gamma > 0$*.*

The supply rate constrains only the normal displacements of u and y from $\mathcal{M}_u$ and $\mathcal{M}_y$, while motion tangent to the manifolds is not penalized; see Fig. 1. The term *transversal* refers to these normal displacements and this usage is distinct from the classical transversality in differential geometry [22], which concerns non-tangential and non-degenerate intersections of submanifolds.

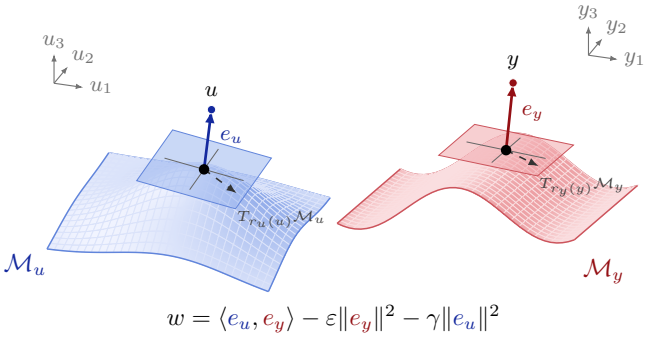


Figure 1: Transversal supply rate. Only the normal displacements e_u and e_y (solid) enter the supply rate; motion along the tangent directions (dashed) is not penalized.

Remark 1 (Locality)**.** *Inequality* (10) *and its integral form hold on any time interval over which* $(x(t), u(t))$ *remains in* $\mathcal{D}$*. The definition itself does not ensure that trajectories remain in this domain.*

Remark 2. *Decreasing* ε *or* γ *preserves* (10)*, so in the input-output strict case one may always assume* $\varepsilon\gamma < 1/4$*. The coefficient matrix* $\left[\begin{smallmatrix} -\gamma & 1/2 \\ 1/2 & -\varepsilon \end{smallmatrix}\right]$ *of the supply rate then has determinant* $\varepsilon\gamma - 1/4 < 0$*, so the supply rate is indefinite.*

Definition 1 gives a dissipation inequality that must hold on the entire tubular domain, leading to two immediate questions: what does this dissipation inequality imply about the entire system, and how can one find a storage function that satisfies it? The next two subsections address these questions by establishing a necessary condition and a constructive sufficient condition for transversal passivity.

3.1 A Necessary Condition

We now consider the infinitesimal consequence of transversal passivity at an equilibrium whose ports lie on the prescribed manifolds.

Let $(x^\star, u^\star) \in \mathcal{D}$ be an equilibrium pair of (8), so that

$$f(x^\star, u^\star) = 0, \quad y^\star = h(x^\star, u^\star), \tag{11}$$

where $u^\star \in \mathcal{M}_u$ and $y^\star \in \mathcal{M}_y$. Assume that f and h are differentiable at $(x^\star, u^\star)$. The linearization of (8) at $(x^\star, u^\star)$ is

$$\delta\dot{x} = \bar{A}\delta x + \bar{B}\delta u, \quad \delta y = \bar{C}\delta x + \bar{D}\delta u, \tag{12}$$

where $\bar{A} = \mathrm{d}_x f[x^\star, u^\star]$, $\bar{B} = \mathrm{d}_u f[x^\star, u^\star]$, $\bar{C} = \mathrm{d}_x h[x^\star, u^\star]$, and $\bar{D} = \mathrm{d}_u h[x^\star, u^\star]$. For brevity, write $\mathrm{P}_u := \mathrm{P}_{N_{u^\star}\mathcal{M}_u}$ and $\mathrm{P}_y := \mathrm{P}_{N_{y^\star}\mathcal{M}_y}$ for the orthogonal projections onto the normal spaces of $\mathcal{M}_u$ at $u^\star$ and of $\mathcal{M}_y$ at $y^\star$.

The following proposition gives a necessary local condition for transversal passivity. At the considered equilibrium, transversal passivity of (8) implies that the linearization (12) satisfies the standard passivity inequality in the transverse components of infinitesimal δu and δy.

Proposition 1. *Consider the equilibrium* $(x^\star, u^\star) \in \mathcal{D}$ *with* $u^\star \in \mathcal{M}_u$ *and* $y^\star \in \mathcal{M}_y$*, and the associated linearized system* (12)*. Suppose that* (10) *holds on* $\mathcal{D}$ *with a storage function* V *that is* C^2 *near* $x^\star$ *and has a local minimum at* $x^\star$*. Set* $Q = \nabla^2 V(x^\star)$ *and* $\delta V(\delta x) = \frac{1}{2}\delta x^\top Q \delta x$*. Then* $Q \succeq 0$*, and along* (12)*,*

$$\delta\dot{V} \le \langle \mathrm{P}_u \delta u,\, \mathrm{P}_y \delta y \rangle - \varepsilon \|\mathrm{P}_y \delta y\|^2 - \gamma \|\mathrm{P}_u \delta u\|^2. \tag{13}$$

Proof. Fix $(\delta x, \delta u)$ and set $x_\lambda = x^\star + \lambda\delta x$, $u_\lambda = u^\star + \lambda\delta u$; since $\mathcal{D}$ is open and contains $(x^\star, u^\star)$, these pairs lie in $\mathcal{D}$ for all sufficiently small $\lambda > 0$. The local minimum gives $\nabla V(x^\star) = 0$ and $Q \succeq 0$, so $\nabla V(x_\lambda) = \lambda Q \delta x + o(\lambda)$, while $f(x^\star, u^\star) = 0$ gives $f(x_\lambda, u_\lambda) = \lambda(\bar{A}\delta x + \bar{B}\delta u) + o(\lambda)$. Since $u^\star \in \mathcal{M}_u$ and $y^\star \in \mathcal{M}_y$, Lemma 1(i) gives $e_u(u^\star) = 0$, $e_y(y^\star) = 0$, $\mathrm{d}_{u^\star} e_u = \mathrm{P}_u$ and $\mathrm{d}_{y^\star} e_y = \mathrm{P}_y$, so that

$$e_u(u_\lambda) = \lambda \mathrm{P}_u \delta u + o(\lambda), \quad e_y\big(h(x_\lambda, u_\lambda)\big) = \lambda \mathrm{P}_y \delta y + o(\lambda).$$

Both sides of (10) evaluated at (x_λ, u_λ) are therefore $O(\lambda^2)$. Dividing by λ^2 and letting $\lambda \to 0$ leaves (13), since $\delta\dot{V} = \delta x^\top Q\, \delta\dot{x} = \delta x^\top Q(\bar{A}\delta x + \bar{B}\delta u)$ along (12). □

Remark 3 (Matrix form)**.** *Since* P_u *and* P_y *are orthogonal projections, the right-hand side of* (13) *is the quadratic supply* $\frac{1}{2}[\delta y^\top \ \delta u^\top]\Pi\,[\delta y^\top \ \delta u^\top]^\top$ *with*

$$\Pi := \begin{bmatrix} -2\varepsilon \mathrm{P}_y & \mathrm{P}_y \mathrm{P}_u \\ \mathrm{P}_u \mathrm{P}_y & -2\gamma \mathrm{P}_u \end{bmatrix}.$$

Because (13) *holds for every* $(\delta x, \delta u) \in \mathbb{R}^n \times \mathbb{R}^m$*, it is equivalent to the linear matrix inequality*

$$\begin{bmatrix} \bar{A}^\top Q + Q\bar{A} & Q\bar{B} \\ \bar{B}^\top Q & 0 \end{bmatrix} - \begin{bmatrix} \bar{C} & \bar{D} \\ 0 & I_m \end{bmatrix}^\top \Pi \begin{bmatrix} \bar{C} & \bar{D} \\ 0 & I_m \end{bmatrix} \preceq 0. \tag{14}$$

For $\mathrm{P}_u = \mathrm{P}_y = I_m$*,* (14) *is the Kalman–Yakubovich–Popov inequality for passivity of the linearization [24]; transversal passivity inserts the normal projections into the supply matrix. Proposition 1 then produces a convex test: if no* $Q \succeq 0$ *satisfies* (14)*, no storage function of the type considered certifies transversal passivity with indices* (ε, γ) *at this equilibrium. The test is necessary only, since* (14) *constrains the linearization alone.*

This proposition demonstrates that transversal passivity enforces the passivity constraint only along the normal directions of the prescribed manifolds. It also reveals a structural obstruction to input-strict transversal passivity.

Corollary 1 (Obstruction to input strictness)**.** *Under the assumptions of Proposition 1, suppose that the linearization has no direct input–output feedthrough, i.e., $\bar{D} = 0$, and that $\mathcal{M}_u$ has positive codimension at $u^\star$, i.e., $N_{u^\star}\mathcal{M}_u \neq \{0\}$. Then $\gamma = 0$.*

Proof. The inequality (13) holds for every pair of directions $(\delta x, \delta u) \in \mathbb{R}^n \times \mathbb{R}^m$. Since $N_{u^\star}\mathcal{M}_u \neq \{0\}$, we may choose $\delta u \in N_{u^\star}\mathcal{M}_u$ with $\|\delta u\| = 1$, and we set $\delta x = 0$. Then $\dot{\delta V} = 0$, while $\bar{D} = 0$ gives $\delta y = \bar{C}\delta x + \bar{D}\delta u = 0$ and hence $\mathrm{P}_y \delta y = 0$. Moreover, $\mathrm{P}_u \delta u = \delta u$, so (13) reduces to $0 \leq -\gamma$. Since $\gamma \geq 0$ by Definition 1, we conclude $\gamma = 0$. □

Equivalently, input strictness is available only when $N_{u^\star}\mathcal{M}_u = \{0\}$. In that case e_u vanishes identically near $u^\star$, so the term $\gamma\|e_u\|^2$ is identically zero and imposes no constraint. This is the transversal counterpart of the classical fact that a passive system with relative degree at least one cannot be input-strictly passive [25].

3.2 A Constructive Sufficient Condition

We next construct a storage candidate from the squared distance of the output to its prescribed manifold, obtaining a sufficient condition that can be checked directly from the system data.

Proposition 2 (Output-distance storage certificate)**.** *Consider the system Σ in (8) with $y = h(x)$ and $h \in C^1$. Restrict its state domain to $\mathcal{X}_y := h^{-1}(\mathcal{N}_y)$ and define*

$$V(x) := \frac{1}{2}\|e_y(h(x))\|^2. \tag{15}$$

If, for some $\varepsilon \geq 0$ and every $(x, u) \in \mathcal{D}$,

$$\begin{aligned}\langle e_y(h(x)), \mathrm{d}_x h[f(x,u)]\rangle \leq\ & \langle e_u(u), e_y(h(x))\rangle \\ & - \varepsilon\|e_y(h(x))\|^2,\end{aligned} \tag{16}$$

then V certifies transversal passivity on $\mathcal{D}$ with indices $(\varepsilon, \gamma) = (\varepsilon, 0)$. In particular, the system is output-strictly $(\mathcal{M}_u, \mathcal{M}_y)$-passive if $\varepsilon > 0$.

Proof. The function V is continuously differentiable and nonnegative on $\mathcal{X}_y$. By Lemma 1(ii) and the chain rule,

$$\mathrm{d}_x V[f(x,u)] = \langle e_y(h(x)), \mathrm{d}_x h[f(x,u)]\rangle .$$

Thus, (16) gives (10) with $\gamma = 0$. □

The choice $\gamma = 0$ is consistent with Corollary 1. Under the hypotheses of the corollary, the absence of direct feedthrough precludes input strictness when $\mathcal{M}_u$ has positive codimension.

For an input-dependent output $y = h(x, u)$, the squared output distance generally depends on u, and its time derivative involves $\dot{u}$ when the input is differentiable. The construction therefore requires additional assumptions or an augmented-state formulation in that setting.

The following example shows that transversal passivity can hold even when standard passivity fails.

Example 1 (A transversally passive oscillator that is not passive)**.** *Consider the nonlinear oscillator*

$$\dot{x} = \mathcal{J}x - \left(1 - \frac{1}{\|x\|}\right)x + u, \quad y = x,$$

where $x \in \mathcal{X} := \mathbb{R}^2 \setminus \{0\}$, $u \in \mathbb{R}^2$, and $\mathcal{J} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$.

We first show that the oscillator is not (standard) passive on $\mathcal{X} \times \mathbb{R}^2$. Choose any $\bar{x} \in \mathcal{X}$ with $1/2 < \bar{r} := \|\bar{x}\| < 1$, and set $\bar{u} = -\mathcal{J}\bar{x} + \left(1 - \frac{1}{\bar{r}}\right)\bar{x}$. Then $\dot{x} = f(\bar{x}, \bar{u}) = 0$ and $\bar{y} = \bar{x}$. Consequently, every continuously differentiable state function Q satisfies $\mathrm{d}_{\bar{x}} Q[f(\bar{x}, \bar{u})] = 0$. However, since $\mathcal{J}^\top = -\mathcal{J}$,

$$\bar{u}^\top \bar{y} = \left(1 - \frac{1}{\bar{r}}\right)\|\bar{x}\|^2 = \bar{r}(\bar{r} - 1) < 0.$$

The standard passivity inequality would therefore require $0 = \mathrm{d}_{\bar{x}} Q[f(\bar{x}, \bar{u})] \leq \bar{u}^\top \bar{y} < 0$, which is impossible. Hence, no continuously differentiable nonnegative storage can establish standard passivity on the stated domain.

Now, we prove that the system is transversally passive. Choose the target input and output manifolds $\mathcal{M}_u = \{0\}$ and $\mathcal{M}_y = \mathbb{S}^1 := \{y \in \mathbb{R}^2 : \|y\| = 1\}$. Take the tubular neighborhoods $\mathcal{N}_u = \mathbb{R}^2$ and $\mathcal{N}_y = \{y \in \mathbb{R}^2 : 1/2 < \|y\| < 3/2\}$. The associated nearest-point retractions and transverse error maps are $r_u(u) = 0$, $e_u(u) = u$, and $r_y(y) = \frac{y}{\|y\|}$, $e_y(y) = y - \frac{y}{\|y\|}$. The tubular domain is $\mathcal{D} = \{(x, u) \in \mathcal{X} \times \mathbb{R}^2 : 1/2 < \|x\| < 3/2\}$. Choose the continuously differentiable nonnegative storage function $V(x) = \frac{1}{2}(\|x\| - 1)^2 = \frac{1}{2}\|e_y(x)\|^2$. Its gradient satisfies $\nabla V(x) = e_y(x)$. Moreover, $e_y(x)$ is parallel to x, so $e_y(x)^\top \mathcal{J} x = 0$. Therefore, for every $(x, u) \in \mathcal{D}$,

$$\mathrm{d}_x V[f(x,u)] = \langle e_u(u), e_y(y)\rangle - \|e_y(y)\|^2.$$

Thus, the system is output-strictly $(\{0\}, \mathbb{S}^1)$-passive on $\mathcal{D}$; see Fig. 2.

The results of Propositions 1 and 2 provide basic necessary and sufficient tools for analyzing transversal passivity. We next explore how this notion relates to existing concepts in passivity theory.

3.3 Connections to Existing Passivity Notions

Existing passivity variants differ primarily in what their supply rate is referenced to — a point, another trajectory,

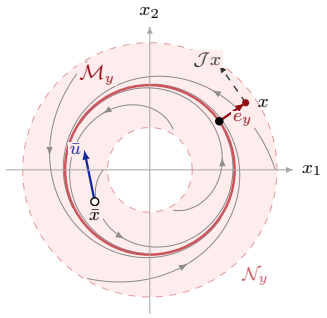


Figure 2: Example 1. Unforced trajectories (gray) converge to $\mathcal{M}_y$ inside the annulus $\mathcal{N}_y$; the radial error e_y is damped and the rotation $\mathcal{J}x$ is unpenalized. At $\bar{x}$, the input $\bar{u}$ holds the state at rest with $\bar{u}^\top \bar{y} < 0$.

or an infinitesimal variation. Transversal passivity references it to a manifold. This recovers the first two notions as special cases and connects to the third at equilibria. Throughout this subsection we take $\varepsilon = \gamma = 0$, the strict cases being analogous.

Equilibrium-referenced notions. For affine targets $\mathcal{M}_u = u^\star + S_u$ and $\mathcal{M}_y = y^\star + S_y$, with $S_u, S_y \subseteq \mathbb{R}^m$ linear subspaces, the nearest-point projections are $r_u(u) = u^\star + \mathrm{P}_{S_u}(u - u^\star)$ and $r_y(y) = y^\star + \mathrm{P}_{S_y}(y - y^\star)$, so that

$$e_u(u) = \mathrm{P}_{S_u^\perp}(u - u^\star), \quad e_y(y) = \mathrm{P}_{S_y^\perp}(y - y^\star), \tag{17}$$

where $\mathrm{P}_{S^\perp} := I - \mathrm{P}_S$. Transversal passivity is therefore shifted dissipativity expressed in the normal coordinates. With $S_u = S_y = \{0\}$ it reduces to shifted passivity about $(u^\star, y^\star)$ [26], and with $u^\star = y^\star = 0$ to standard passivity (1). The same specialization applies to each equilibrium pair appearing in the EIP and MEIP inequalities [16, 27]; those are families of singleton-target inequalities carrying a separate storage per equilibrium, and neither implies nor is implied by transversal passivity for a fixed pair of non-singleton targets.

Trajectory-referenced notions. Incremental passivity references the supply rate to a second trajectory, and is exactly transversal passivity of the product system relative to the diagonals. Consider two solutions of (8) with states (x_1, x_2), inputs (u_1, u_2) and outputs (y_1, y_2). The system is *incrementally passive* if there exists a continuously differentiable $V_\Delta : \mathcal{X} \times \mathcal{X} \to \mathbb{R}_{\geq 0}$ with

$$\dot{V}_\Delta(x_1, x_2) \leq (u_1 - u_2)^\top (y_1 - y_2) \tag{18}$$

along every pair of solutions.

Proposition 3. *Let $\mathcal{U} = \mathcal{Y} = \mathbb{R}^m$ and equip $\Sigma \times \Sigma$ with the diagonal targets $\Delta_u = \{(u, u) : u \in \mathbb{R}^m\}$ and $\Delta_y = \{(y, y) : y \in \mathbb{R}^m\}$. For a continuously differentiable nonnegative V_Δ, the following are equivalent.*

(i) V_Δ satisfies (18) along every pair of solutions.

(ii) $V_\Delta/2$ is a transversal storage for $\Sigma \times \Sigma$ with respect to (Δ_u, Δ_y) with indices $(\varepsilon, \gamma) = (0, 0)$.

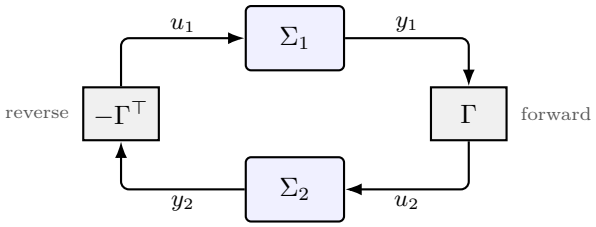


Figure 3: Feedback interconnection of Σ_1 and Σ_2 through the forward channel Γ and the reverse channel $-\Gamma^\top$.

Proof. The nearest-point projections onto the diagonals are the averages, so the transverse errors are the half-differences $e_{\Delta_u} = \frac{1}{2}(u_1 - u_2, u_2 - u_1)$ and $e_{\Delta_y} = \frac{1}{2}(y_1 - y_2, y_2 - y_1)$. Hence $\langle e_{\Delta_u}, e_{\Delta_y} \rangle = \frac{1}{2}(u_1 - u_2)^\top (y_1 - y_2)$, and the two dissipation inequalities coincide after scaling the storage by 2. □

Infinitesimal notions. Differential passivity references the supply rate to an infinitesimal variation and requires the inequality with supply $\delta u^\top \delta y$ along the variational system of every trajectory [28]. Transversal passivity meets this notion only at equilibria: at singleton targets the projections in (13) are the identity, and Proposition 1 gives the differential passivity inequality of the linearization at $(x^\star, u^\star)$. This is an equilibrium-level consequence and does not establish differential passivity along all trajectories.

4 Feedback Interconnection and Compatibility Residuals

Much of the power of passivity theory is its compositional property. The negative feedback interconnection of two passive systems is again passive, with the sum of the subsystem storage functions as the closed-loop storage function. This property rests on an exact cancellation of the two supply rates: the interconnection is power conserving, so that $\langle u_1, y_1 \rangle + \langle u_2, y_2 \rangle = 0$, and adding the subsystem inequalities $\dot{Q}_i \leq \langle u_i, y_i \rangle$ leaves $\dot{Q}_1 + \dot{Q}_2 \leq 0$.

This power-conserving property also holds for the feedback interconnection in Fig. 3, with inputs and outputs $u_i, y_i \in \mathbb{R}^{m_i}$ satisfying

$$u_1 = -\Gamma^\top y_2, \quad u_2 = \Gamma y_1, \quad \Gamma \in \mathbb{R}^{m_2 \times m_1}, \tag{19}$$

where Γ is the forward channel and $-\Gamma^\top$ the reverse channel. The port variables satisfying (19) form a separable Dirac structure [15, Proposition 2.5], so that $\langle u_1, y_1 \rangle + \langle u_2, y_2 \rangle = -\langle \Gamma^\top y_2, y_1 \rangle + \langle \Gamma y_1, y_2 \rangle = 0$.

This cancellation is exactly what can fail for transversal passivity, because its supply rates are built from the transverse port errors rather than the ports themselves, while power conservation constrains only the ports. Consequently, transversal passivity does not compose in general.

In this section, we first show that interconnecting two transversally passive systems through (19) leaves a non-cancellable residual supply in the closed-loop dissipation inequality. By analyzing the structure of the transversal supply rate and its residual, we then establish sufficient conditions under which transversal passivity composes, and show how closed-loop dissipation estimates can still be obtained when it does not.

4.1 Closed-loop Realization and Dissipation

Consider two nonlinear systems

$$\Sigma_i : \begin{cases} \dot{x}_i = f_i(x_i, u_i), \\ y_i = h_i(x_i, u_i), \end{cases} \quad i = 1, 2, \tag{20}$$

interconnected through (19) with $\Gamma \in \mathbb{R}^{m_2 \times m_1}$ as in Fig. 3, where $x_i \in \mathcal{X}_i \subseteq \mathbb{R}^{n_i}$, $u_i \in \mathcal{U}_i \subseteq \mathbb{R}^{m_i}$ and $y_i \in \mathcal{Y}_i \subseteq \mathbb{R}^{m_i}$, with $\mathcal{X}_i, \mathcal{U}_i, \mathcal{Y}_i$ open, h_i of class C^1 and f_i locally Lipschitz.

Adopting the notation of Section 3, equip each subsystem Σ_i with target manifolds $(\mathcal{M}_{u_i}, \mathcal{M}_{y_i})$, tubular neighborhoods $(\mathcal{N}_{u_i}, \mathcal{N}_{y_i})$, nearest-point retractions r_{u_i} and r_{y_i}, and transverse port error maps e_{u_i} and e_{y_i}. Assume that its transversal-passivity inequality holds with storage V_i and indices $\varepsilon_i, \gamma_i \geq 0$ on the nonempty domain

$$\mathcal{D}_i := \{(x_i, u_i) \in \mathcal{X}_i \times \mathcal{U}_i : u_i \in \mathcal{N}_{u_i},\ h_i(x_i, u_i) \in \mathcal{N}_{y_i}\}.$$

Both channels can be treated by the same construction. Let B be a linear map from an output space carrying a target manifold $\mathcal{M}_y$, with tubular neighborhood $\mathcal{N}_y$ and transverse error map e_y, to an input space carrying $\mathcal{M}_u$, $\mathcal{N}_u$ and e_u. On the channel domain

$$\mathcal{D}_B := \{y \in \mathcal{N}_y :\ By \in \mathcal{N}_u\}, \tag{21}$$

both error maps are well defined.

The channel transports transverse errors when $e_u(By) = B\, e_y(y)$. Since a linear channel need not commute with nonlinear nearest-point projections, this relation fails in general, and its failure is measured by the *compatibility defect*, defined for $y \in \mathcal{D}_B$ by

$$\delta_B(y) := e_u(By) - B\, e_y(y) = B\, r_y(y) - r_u(By). \tag{22}$$

The channel also constrains which target points are attainable: the *feasible target set*

$$\mathcal{F}_B := \mathcal{M}_y \cap B^{-1}(\mathcal{M}_u) \subseteq \mathcal{D}_B \tag{23}$$

collects the points of $\mathcal{M}_y$ that B maps onto $\mathcal{M}_u$. Both objects are illustrated in Fig. 4.

The forward and reverse channels of (19) are the cases $B = \Gamma$ and $B = -\Gamma^\top$, respectively. The signal-space defect maps are δ_Γ and $\delta_{-\Gamma^\top}$. We denote the corresponding channel domains and feasible sets by $\mathcal{D}_{fc}, \mathcal{D}_{rc}, \mathcal{F}_{fc}, \mathcal{F}_{rc}$,

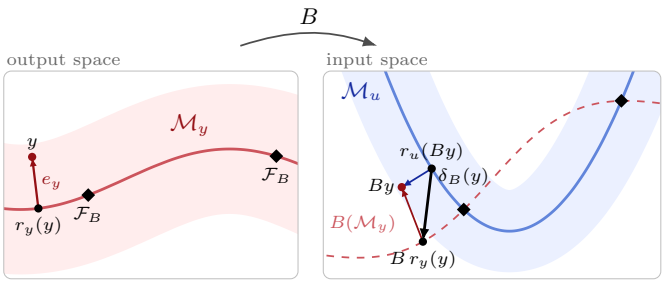


Figure 4: Compatibility defect for a linear channel B; the shaded bands are the tubes $\mathcal{N}_y$ and $\mathcal{N}_u$. The dashed curve is $B(\mathcal{M}_y)$, and the diamonds mark $\mathcal{F}_B$, discrete here, and its image. The defect $\delta_B(y) = B\, r_y(y) - r_u(By)$ (bold) is the mismatch between projecting then mapping and mapping then projecting.

collected into the *joint channel domain* $\mathcal{D}_\Gamma$ and the *joint feasibility set* $\mathcal{F}_\Gamma$.

$$\mathcal{D}_\Gamma := \mathcal{D}_{fc} \times \mathcal{D}_{rc}, \tag{24}$$

$$\mathcal{F}_\Gamma := \mathcal{F}_{fc} \times \mathcal{F}_{rc}, \tag{25}$$

so that $\mathcal{F}_\Gamma \subseteq \mathcal{D}_\Gamma$.

For the closed-loop realization, since the output maps h_i may contain direct feedthrough, algebraic well-posedness is not automatic. We therefore impose the following standing assumption.

Assumption 1 (Algebraic well-posedness)**.** *There exists an open set $\mathcal{X}_\Gamma \subseteq \mathcal{X}_1 \times \mathcal{X}_2$ such that, for every $x = (x_1, x_2) \in \mathcal{X}_\Gamma$, the realized port maps*

$$y_1(x) = h_1(x_1, u_1(x)), \quad u_2(x) = \Gamma y_1(x),$$
$$y_2(x) = h_2(x_2, u_2(x)), \quad u_1(x) = -\Gamma^\top y_2(x)$$

have a unique solution $(y_1(x), u_2(x), y_2(x), u_1(x))$ in $\mathcal{Y}_1 \times \mathcal{U}_2 \times \mathcal{Y}_2 \times \mathcal{U}_1$.

We further assume that the realized port maps are C^1 on the algebraic well-posedness domain $\mathcal{X}_\Gamma$. Throughout the remainder of this work, we adopt Assumption 1 and the C^1 regularity of the realized port maps as standing hypotheses and do not repeat them. The associated closed-loop maps, state-space sets, and preimages are understood relative to $\mathcal{X}_\Gamma$.

Within $\mathcal{X}_\Gamma$, we can define the realized output map and the closed-loop vector field by

$$h_\Gamma(x) := \big(y_1(x), y_2(x)\big), \quad F_{\mathrm{cl}}(x) := \begin{bmatrix} f_1(x_1, u_1(x)) \\ f_2(x_2, u_2(x)) \end{bmatrix}, \tag{26}$$

so that the closed loop is $\dot{x} = F_{\mathrm{cl}}(x)$. Then F_{cl} is locally Lipschitz on $\mathcal{X}_\Gamma$.

Restricting the realized ports to the subsystem tubular domains $\mathcal{D}_i$ gives the closed-loop tubular domain

$$\mathcal{D}_{\mathrm{cl}} := \{x \in \mathcal{X}_\Gamma : (x_i, u_i(x)) \in \mathcal{D}_i, \quad i = 1, 2\}, \tag{27}$$

where the realized transverse port errors are defined and both transversal-passivity inequalities apply. We use the shorthand $e_{y_i}(x) := e_{y_i}(y_i(x))$ and $e_{u_i}(x) := e_{u_i}(u_i(x))$, for $i = 1, 2$. The realized defect maps are $\delta_{fc}(x) := \delta_\Gamma(y_1(x))$ and $\delta_{rc}(x) := \delta_{-\Gamma^\top}(y_2(x))$. From now on, channel-indexed symbols δ_B denote signal-space defect maps, whereas δ_{fc} and δ_{rc} denote their closed-loop realizations.

Note that, by the feedback relations, $\mathcal{D}_{\mathrm{cl}}$ equals $h_\Gamma^{-1}(\mathcal{D}_\Gamma)$ and the closed-loop feasible state set is

$$\begin{aligned}\mathcal{Z}_{\mathrm{cl}} &:= h_\Gamma^{-1}(\mathcal{F}_\Gamma) \\ &= \{x \in \mathcal{D}_{\mathrm{cl}} : e_{y_i}(x) = e_{u_i}(x) = 0,\ i = 1, 2\}.\end{aligned} \tag{28}$$

The main point of this subsection is that the part of the supply that survives the interconnection can be attributed to the channels. In standard passivity, adding the subsystem inequalities cancels the supply rates because the interconnection is power conserving. Here the supply rates are built from transverse errors, and by (22) the channels transport these errors only up to the defects, $e_{u_2} = \Gamma e_{y_1} + \delta_{fc}$ and $e_{u_1} = -\Gamma^\top e_{y_2} + \delta_{rc}$. The Γ-terms still cancel by adjointness, but the defect terms do not, and they remain as a residual supply.

To state this, define the composite storage $V_{\mathrm{cl}}(x) := V_1(x_1)+V_2(x_2)$ and the quadratic dissipation term $Q(x) := \sum_{i=1}^2 \big(\varepsilon_i \|e_{y_i}(x)\|^2 + \gamma_i \|e_{u_i}(x)\|^2\big)$.

Theorem 1 (Closed-loop dissipation and residual supply)**.** *Consider subsystems Σ_1 and Σ_2 with storage functions V_1 and V_2. Then, for every $x \in \mathcal{D}_{\mathrm{cl}}$,*

$$\mathrm{d}_x V_{\mathrm{cl}}[F_{\mathrm{cl}}(x)] \le s_\delta(x) - Q(x), \tag{29}$$

where the residual transverse supply *s_δ is the total transverse supply of the two subsystems and admits the channel-defect form*

$$\begin{aligned}s_\delta(x) &:= \textstyle\sum_{i=1}^2 \langle e_{u_i}(x), e_{y_i}(x)\rangle \\ &= \langle \delta_{rc}(x), e_{y_1}(x)\rangle + \langle \delta_{fc}(x), e_{y_2}(x)\rangle.\end{aligned} \tag{30}$$

Proof. Summing the transversal-passivity inequalities of Σ_1 and Σ_2 on $\mathcal{D}_{\mathrm{cl}}$ gives (29) with s_δ given by the first expression in (30). By (19) and (22), the realized transverse input errors decompose as $e_{u_2}(x) = \Gamma e_{y_1}(x) + \delta_{fc}(x)$ and $e_{u_1}(x) = -\Gamma^\top e_{y_2}(x) + \delta_{rc}(x)$. Substituting these relations into the first expression, the channel terms $\langle \Gamma e_{y_1}, e_{y_2}\rangle - \langle \Gamma^\top e_{y_2}, e_{y_1}\rangle$ cancel by adjointness, which leaves the second expression. □

The two expressions in (30) indicate how the rest of this section proceeds. The first is the total transverse supply, which in standard passivity would vanish by power conservation. The second attributes the part that survives to the channels, and closed-loop dissipation estimates therefore reduce to analyzing these channel terms. Since each defect depends only on a channel and its prescribed manifolds, not on the subsystem dynamics, conditions under which the defects vanish can be checked in signal space, without the closed-loop realization; Subsection 4.2 develops such conditions. When the defects do not vanish, both factors in each term of the second expression still vanish on the closed-loop feasible state set, so the residual is locally quadratic there; Subsection 4.3 examines this quadratic part.

4.2 Vanishing Residuals and Exact Compatibility

If $s_\delta(x) = 0$, transversal passivity composes directly and (29) reduces to $\mathrm{d}_x V_{\mathrm{cl}}[F_{\mathrm{cl}}(x)] \le -Q(x)$. By (30), this is achieved when both channel defects vanish, providing a geometric sufficient condition that can be checked directly in signal space without solving for the closed-loop realization. Exact compatibility is a property of a region of output signals, typically a neighborhood of part of $\mathcal{M}_y$, on which the defect is checked.

Definition 2 (Exact channel compatibility)**.** *The channel B is* exactly compatible *on a relatively open set $\Omega_B \subseteq \mathcal{D}_B$ if δ_B vanishes identically on Ω_B. The interconnection is exactly compatible on $\Omega := \Omega_{fc} \times \Omega_{rc} \subseteq \mathcal{D}_\Gamma$ if this holds for $B = \Gamma$ on Ω_{fc} and for $B = -\Gamma^\top$ on Ω_{rc}.*

By (22), $\delta_B(y) = 0$ is equivalent to $r_u(By) = B r_y(y)$, so exact compatibility means that the channel commutes with the prescribed nearest-point retractions on Ω_B. In closed loop, this signal-space condition eliminates $s_\delta(x)$ and relates the realized transverse output errors to the feasible target $\mathcal{F}_\Gamma$.

Proposition 4 (Closed-loop dissipation and feasible distance under exact compatibility)**.** *Suppose that the interconnection is exactly compatible on $\Omega \subseteq \mathcal{D}_\Gamma$. Then, for every $x \in h_\Gamma^{-1}(\Omega)$, the residual supply vanishes, $s_\delta(x) = 0$, and the following two properties hold:*

(i) Dissipation: *the closed loop dissipates at least at rate Q,*

$$\mathrm{d}_x V_{\mathrm{cl}}[F_{\mathrm{cl}}(x)] \le -Q(x) \le 0; \tag{31}$$

(ii) Feasible distance: *the transverse output errors measure the distance to the feasible target,*

$$\mathrm{dist}\big(h_\Gamma(x), \mathcal{F}_\Gamma\big)^2 = \|e_{y_1}(x)\|^2 + \|e_{y_2}(x)\|^2. \tag{32}$$

Proof. Since $\Omega \subseteq \mathcal{D}_\Gamma$, we have $h_\Gamma^{-1}(\Omega) \subseteq h_\Gamma^{-1}(\mathcal{D}_\Gamma) = \mathcal{D}_{\mathrm{cl}}$. For every $x \in h_\Gamma^{-1}(\Omega)$, $h_\Gamma(x) \in \Omega$ implies that both realized defects vanish, so $s_\delta(x) = 0$ and (31) follows from (29). Moreover, for an exactly compatible channel B, (22) gives $B r_y(y) = r_u(By) \in \mathcal{M}_u$ on its compatibility region, hence $r_y(y) \in \mathcal{F}_B$. Since $\mathcal{F}_B \subseteq \mathcal{M}_y$,

$$\mathrm{dist}(y, \mathcal{M}_y) \le \mathrm{dist}(y, \mathcal{F}_B) \le \|y - r_y(y)\| = \mathrm{dist}(y, \mathcal{M}_y).$$

Applying this identity to the two realized outputs and using $\mathrm{dist}((z_1, z_2), A_1 \times A_2)^2 = \mathrm{dist}(z_1, A_1)^2 + \mathrm{dist}(z_2, A_2)^2$ for the Euclidean product norm proves (32). □

In general, the realized defects δ_{fc} and δ_{rc} do not vanish, so the residual transverse supply $s_\delta(x)$ is nonzero and sign-indefinite, and (29) does not automatically reduce to (31). Obtaining closed-loop dissipation estimates in this case requires analyzing the local structure of $s_\delta(x)$ around the feasible state set, which we do in the sequel.

4.3 Local Structure of Nonvanishing Residuals

Suppose that the four realized maps $e_{y_1}, e_{y_2}, \delta_{fc}, \delta_{rc}$ are C^2 near a state $p \in \mathcal{D}_{\mathrm{cl}}$. Since $s_\delta(x)$ in (30) is a sum of inner products of these maps, a second-order Taylor expansion about p gives

$$\begin{aligned} s_\delta(p+\xi) =& s_\delta(p) + \mathrm{d}_p s_\delta[\xi] + \frac{1}{2}\Big(\langle \mathrm{d}_p^2 \delta_{rc}[\xi,\xi], e_{y_1}(p)\rangle \\ &+ \langle \delta_{rc}(p), \mathrm{d}_p^2 e_{y_1}[\xi,\xi]\rangle + \langle \mathrm{d}_p^2 \delta_{fc}[\xi,\xi], e_{y_2}(p)\rangle \\ &+ \langle \delta_{fc}(p), \mathrm{d}_p^2 e_{y_2}[\xi,\xi]\rangle \Big) + \xi^\top R_p \xi + o(\|\xi\|^2), \end{aligned}$$

where $R_p := \mathrm{sym}\big((\mathrm{d}_p\delta_{rc})^\top \mathrm{d}_p e_{y_1} + (\mathrm{d}_p\delta_{fc})^\top \mathrm{d}_p e_{y_2}\big)$, with $\mathrm{sym}(L) := (L+L^\top)/2$. By the first expression in (30), equivalently $\xi^\top R_p \xi = \sum_i \langle \mathrm{d}_p e_{u_i}[\xi], \mathrm{d}_p e_{y_i}[\xi]\rangle$ when $p \in \mathcal{Z}_{\mathrm{cl}}$, so R_p is the quadratic form of the linearized total transverse supply.

On the closed-loop feasible state set $\mathcal{Z}_{\mathrm{cl}} = h_\Gamma^{-1}(\mathcal{F}_\Gamma)$, all four realized maps $e_{y_1}, e_{y_2}, \delta_{fc}, \delta_{rc}$ vanish by (22) and (28). Consequently, when $p \in \mathcal{Z}_{\mathrm{cl}}$, all preceding terms in the expansion vanish and only

$$s_\delta(p+\xi) = \xi^\top R_p \xi + o(\|\xi\|^2)$$

remains. One way to obtain closed-loop dissipation estimates near $\mathcal{Z}_{\mathrm{cl}}$ is to cancel this quadratic term so that $s_\delta(x)$ is locally cubic.

Proposition 5 (Quadratic cancellation of the residual)**.** *Let $n := n_1 + n_2$. Suppose that the realized maps $e_{y_1}, e_{y_2}, \delta_{fc}, \delta_{rc}$ are C^2 on an open set $\mathcal{O} \subseteq \mathcal{D}_{\mathrm{cl}}$, and let $\mathcal{K} \subseteq \mathcal{Z}_{\mathrm{cl}} \cap \mathcal{O}$ be nonempty and compact.*

Then there exist constants $r > 0$ and $C > 0$ such that the open r-neighborhood of $\mathcal{K}$ is contained in $\mathcal{O}$ and the following two properties hold:

(i) Quadratic approximation: *for every $p \in \mathcal{K}$ and every $\xi \in \mathbb{R}^n$ with $\|\xi\| < r$, the residual supply is quadratic up to a cubic error,*

$$\big| s_\delta(p+\xi) - \xi^\top R_p \xi \big| \le C\|\xi\|^3; \tag{33}$$

(ii) Cubic residual under quadratic cancellation: *if $R_p = 0$ for every $p \in \mathcal{K}$, then for every x with $\mathrm{dist}(x,\mathcal{K}) < r$, we have $|s_\delta(x)| \le C\,\mathrm{dist}(x,\mathcal{K})^3$ and*

$$\mathrm{d}_x V_{\mathrm{cl}}[F_{\mathrm{cl}}(x)] \le -Q(x) + C\,\mathrm{dist}(x,\mathcal{K})^3. \tag{34}$$

Proof. All four realized maps vanish on $\mathcal{K} \subseteq \mathcal{Z}_{\mathrm{cl}}$ by (22) and (28). Since $\mathcal{K}$ is compact and $\mathcal{O}$ is open, we may choose $0 < r \le 1$ such that $\mathcal{N} := \{z \in \mathbb{R}^n : \mathrm{dist}(z,\mathcal{K}) \le r\}$ is a compact subset of $\mathcal{O}$. For every $p \in \mathcal{K}$ and $\|\xi\| < r$, the segment $p + t\xi$, $t \in [0,1]$, lies in $\mathcal{N}$.

Since $\mathcal{N}$ is compact and the four maps are C^2, the constants

$$L := \max_g \sup_{z\in\mathcal{N}} \|\mathrm{d}_z g\|, \quad M := \tfrac{1}{2}\max_g \sup_{z\in\mathcal{N}} \|\mathrm{d}_z^2 g\|,$$

with g ranging over $\{e_{y_1}, e_{y_2}, \delta_{fc}, \delta_{rc}\}$, are finite; we may take $L, M > 0$. Since $g(p) = 0$, Taylor's formula with integral remainder along the segment $p+t\xi$ gives $g(p+\xi) = \mathrm{d}_p g[\xi] + \rho_{g,p}(\xi)$ with $\|\mathrm{d}_p g[\xi]\| \le L\|\xi\|$ and $\|\rho_{g,p}(\xi)\| \le M\|\xi\|^2$, for all $p \in \mathcal{K}$ and $\|\xi\| < r$. Substituting these expansions into (30), the products of the first differentials sum to $\xi^\top R_p \xi$, and the remaining terms satisfy

$$\begin{aligned} \big| s_\delta(p+\xi) - \xi^\top R_p \xi \big| &\le 4LM\|\xi\|^3 + 2M^2\|\xi\|^4 \\ &\le (4LM + 2M^2)\|\xi\|^3, \end{aligned}$$

because $\|\xi\| < r \le 1$. Thus (33) holds with $C := 4LM + 2M^2$.

For each $p \in \mathcal{K}$, symmetry of R_p implies that $R_p = 0$ is equivalent to the directional cancellation condition

$$\langle \mathrm{d}_p\delta_{rc}[v], \mathrm{d}_p e_{y_1}[v]\rangle + \langle \mathrm{d}_p\delta_{fc}[v], \mathrm{d}_p e_{y_2}[v]\rangle = 0, \tag{35}$$

for all $v \in \mathbb{R}^n$. When $R_p = 0$ on $\mathcal{K}$, for any x with $\mathrm{dist}(x,\mathcal{K}) < r$, compactness of $\mathcal{K}$ provides $p \in \mathcal{K}$ such that $\|x - p\| = \mathrm{dist}(x,\mathcal{K})$. Taking $\xi = x - p$ in (33) gives $|s_\delta(x)| \le C\,\mathrm{dist}(x,\mathcal{K})^3$, and (29) then yields (34). □

Checking $R_p = 0$ directly requires the differentials of the realized maps, and hence the closed-loop realization. The structure of R_p suggests a simpler sufficient condition: if both realized defects vanish to first order at p, that is, $\mathrm{d}_p\delta_{fc} = \mathrm{d}_p\delta_{rc} = 0$, then $R_p = 0$ regardless of the output errors. Since the realized defects are the signal-space defects evaluated along the realized outputs, this holds whenever the signal-space defects have vanishing differentials at the feasible targets.

Definition 3 (First-order channel compatibility)**.** *Suppose δ_B is of class C^1 on $\mathcal{D}_B$. The channel B is* first-order compatible along $\mathcal{S}_B \subseteq \mathcal{F}_B$ *if*

$$\mathrm{d}_y \delta_B = 0 \quad \text{for every } y \in \mathcal{S}_B. \tag{36}$$

The interconnection is first-order compatible along $\mathcal{S} := \mathcal{S}_{fc} \times \mathcal{S}_{rc} \subseteq \mathcal{F}_\Gamma$ if this holds for $B = \Gamma$ along $\mathcal{S}_{fc}$ and for $B = -\Gamma^\top$ along $\mathcal{S}_{rc}$.

Differentiating (22) shows that (36) is equivalent to

$$\mathrm{d}_{By} r_u \circ B = B \circ \mathrm{d}_y r_y, \quad y \in \mathcal{S}_B, \tag{37}$$

which depends only on the channel and the target retractions. If the interconnection is first-order compatible along

Table 1: Compatibility conditions and residual supply. The last column lists the port-convergence result of Section 5 for each case and serves as a roadmap for that section.

Condition	Residual near $\mathcal{Z}_{\mathrm{cl}}$	Result
Exact on Ω ($\delta \equiv 0$)	$s_\delta \equiv 0$, $\dot{V}_{\mathrm{cl}} \le -Q$ (Prop. 4)	Cor. 2
$R_p = 0$ on $\mathcal{K}$ (e.g., first-order along $\mathcal{S}$)	$\lvert s_\delta \rvert \le C\,\mathrm{dist}(x,\mathcal{K})^3$ (Prop. 5(ii))	Cor. 3
$R_p \neq 0$	$\xi^\top R_p \xi + O(\lVert\xi\rVert^3)$, Eq. (38)	Thm. 3

$\mathcal{S}$, then, for every $p \in h_\Gamma^{-1}(\mathcal{S})$, the chain rule gives $\mathrm{d}_p\delta_{fc} = \mathrm{d}_{y_1(p)}\delta_\Gamma \circ \mathrm{d}_p y_1 = 0$ and $\mathrm{d}_p\delta_{rc} = \mathrm{d}_{y_2(p)}\delta_{-\Gamma^\top} \circ \mathrm{d}_p y_2 = 0$, so $R_p = 0$. Under the hypotheses of Proposition 5, first-order compatibility along $\mathcal{S}$ therefore yields (34) for any compact $\mathcal{K} \subseteq h_\Gamma^{-1}(\mathcal{S}) \cap \mathcal{O}$.

Exact compatibility implies first-order compatibility. If the interconnection is exactly compatible on a relatively open Ω, then δ_Γ and $\delta_{-\Gamma^\top}$ vanish identically on open sets, so their differentials vanish there as well, and the interconnection is first-order compatible along $\mathcal{S} = \Omega \cap (\mathcal{M}_{y_1} \times \mathcal{M}_{y_2})$. This set lies in $\mathcal{F}_\Gamma$ by the proof of Proposition 4. The converse fails when the channel matches the tangent spaces of the targets but not their curvature. In closed loop, Ω pulls back to an open set of states on which s_δ vanishes identically, whereas $\mathcal{S}$ pulls back to a subset $h_\Gamma^{-1}(\mathcal{S})$ of the feasible state set $\mathcal{Z}_{\mathrm{cl}}$, near which s_δ is only cubically small.

When $R_p \neq 0$, the quadratic term $\xi^\top R_p \xi$ does not vanish identically, and we retain its signed contribution. Under the regularity and compactness hypotheses of Proposition 5, (29) and (33) give

$$\mathrm{d}_{p+\xi} V_{\mathrm{cl}}[F_{\mathrm{cl}}(p+\xi)] \le -Q(p+\xi) + \xi^\top R_p \xi + C\lVert\xi\rVert^3, \quad (38)$$

for every $p \in \mathcal{K}$ and $\lVert\xi\rVert < r$. Obtaining a closed-loop dissipation estimate then requires comparing the available dissipation (such as Q in (29)) directly with the signed quadratic residual $\xi^\top R_p \xi$, which we do in Subsection 5.4.

In summary, the conditions of this section form a hierarchy. Exact compatibility on Ω eliminates the residual supply (Proposition 4). First-order compatibility along $\mathcal{S}$, or more generally $R_p = 0$ on $\mathcal{K}$, makes it cubic (Proposition 5(ii)). Otherwise, the quadratic residual $\xi^\top R_p \xi$ remains and must be compared with the available dissipation. Section 5 shows how each case yields port convergence. Table 1 summarizes the hierarchy and, for later reference, lists the corresponding convergence result of Section 5 for each case.

5 Port Convergence under Transversal Passivity

This section establishes convergence of the realized port variables to their prescribed manifolds for the feedback interconnection of Section 4. The closed-loop dissipation inequalities of Section 4 hold only on the domain where the transverse errors are defined and both transversal-passivity inequalities apply. A nonincreasing storage therefore keeps a trajectory in a storage sublevel set only while the trajectory remains in this domain, and the sublevel sets may reach its boundary or be unbounded. We first use additional dynamical information, such as conserved quantities, to confine trajectories to a compact subset of the domain, which ensures that they exist for all time. We then show that the ports converge to their prescribed manifolds under a dissipation margin, for each of the residual cases of Table 1.

5.1 Compact Confinement

In classical Lyapunov arguments, a nonincreasing storage with compact sublevel sets keeps every trajectory in a compact set, so solutions exist for all time and remain bounded. Two features of the present setting prevent a direct use of this argument. First, the dissipation inequality holds only on a domain, so it constrains a trajectory only while the trajectory stays there. Second, the storage sublevel sets need not be compact. For example, they may be tubes about the target manifolds, which are often unbounded. We therefore complement the storage with a set that trajectories cannot leave, and require only that its intersection with a storage sublevel set be compact.

Throughout this section, let $\mathcal{O} \subseteq \mathcal{D}_{\mathrm{cl}}$, with $\mathcal{D}_{\mathrm{cl}}$ the closed-loop tubular domain (27), be a nonempty open analysis domain for the closed-loop dynamics $\dot{x} = F_{\mathrm{cl}}(x)$ in (26). Since the dissipation inequalities apply only while a trajectory remains in $\mathcal{O}$, a subset of $\mathcal{O}$ is called *forward invariant relative to* $\mathcal{O}$ if every solution initialized in it remains in it for as long as it remains in $\mathcal{O}$. The composite storage $V_{\mathrm{cl}}(x) = V_1(x_1) + V_2(x_2)$ of Theorem 1 is continuously differentiable and nonnegative on $\mathcal{O}$, and for $c \ge 0$ we write

$$\mathcal{L}_c := \{x \in \mathcal{O} : V_{\mathrm{cl}}(x) \le c\}$$

for its sublevel sets in $\mathcal{O}$.

A *restriction* is a set $\mathcal{H} \subseteq \mathcal{O}$ that is forward invariant relative to $\mathcal{O}$. The trivial choice $\mathcal{H} = \mathcal{O}$ imposes no restriction. A conserved quantity provides a nontrivial one. Suppose that a continuously differentiable map $\chi : \mathcal{O} \to \mathbb{R}^k$ satisfies

$$\mathrm{d}_x \chi[F_{\mathrm{cl}}(x)] = 0 \quad (39)$$

for every $x \in \mathcal{O}$. Such a χ is called a *first integral*, or *conserved quantity*, of the closed loop. Each initial condition $x_0 \in \mathcal{O}$ selects the conserved level set

$$\mathcal{H} := \{x \in \mathcal{O} : \chi(x) = \chi(x_0)\}. \quad (40)$$

Whenever a solution remains in $\mathcal{O}$, the chain rule gives $\chi(x(t)) = \chi(x_0)$, so $\mathcal{H}$ is a restriction. A conserved quantity constrains directions left unconstrained by V_{cl}, without a further dissipation inequality in those directions.

This use of conserved quantities to complement a storage function is in the same spirit as Casimir-based methods for interconnected Hamiltonian systems [29].

Proposition 6 (Compact confinement)**.** *Let $\mathcal{H} \subseteq \mathcal{O}$ be a restriction, let $x_0 \in \mathcal{H}$ and $c \geq V_{\mathrm{cl}}(x_0)$, and suppose that*

(i) $\mathrm{d}_x V_{\mathrm{cl}}[F_{\mathrm{cl}}(x)] \leq 0$ *for every $x \in \mathcal{H}$, and*

(ii) the closure of $\mathcal{L}_c \cap \mathcal{H}$ in $\mathbb{R}^n$, $n = n_1 + n_2$, is a compact subset of $\mathcal{O}$.

Then the solution initialized at x_0 is forward complete and remains in $\mathcal{L}_c \cap \mathcal{H}$ for all $t \geq 0$. Moreover, for every continuous function $\psi : \mathcal{O} \to \mathbb{R}$, the map $t \mapsto \psi(x(t))$ is uniformly continuous on $[0, \infty)$.

Proof. Let $x : [0, t_{\max}) \to \mathcal{X}_\Gamma$ be the maximal solution with $x(0) = x_0$, and write $\mathcal{A} := \mathcal{L}_c \cap \mathcal{H}$. As long as x remains in $\mathcal{O}$, relative forward invariance keeps it in $\mathcal{H}$, and (i) then gives $V_{\mathrm{cl}}(x(t)) \leq V_{\mathrm{cl}}(x_0) \leq c$, so $x(t) \in \mathcal{A}$. If $\tau := \sup\{t < t_{\max} : x([0,t]) \subseteq \mathcal{O}\} < t_{\max}$, then $x([0,\tau)) \subseteq \mathcal{A}$ and continuity give $x(\tau) \in \overline{\mathcal{A}} \subseteq \mathcal{O}$; since $\mathcal{O}$ is open, x remains in $\mathcal{O}$ beyond τ, a contradiction. Therefore $x(t) \in \mathcal{A}$ for all $t < t_{\max}$.

Next, suppose $t_{\max} < \infty$. Then $x([0, t_{\max}))$ lies in the compact set $\overline{\mathcal{A}} \subseteq \mathcal{X}_\Gamma$. Since F_{cl} in (26) is locally Lipschitz on $\mathcal{X}_\Gamma$, [6, Theorem 3.3] extends the solution beyond $t_{\max}$, which is a contradiction. Thus, $t_{\max} = \infty$ and $x(t) \in \mathcal{A}$ for all $t \geq 0$.

It remains to show that $t \mapsto \psi(x(t))$ is uniformly continuous, which is what allows Barbalat's lemma to be applied in Theorem 2. Continuity of F_{cl} on $\overline{\mathcal{A}}$ gives a constant $M_F \geq 0$ such that $\|F_{\mathrm{cl}}(x)\| \leq M_F$ there. Thus, $\|x(t) - x(s)\| \leq M_F |t - s|$ for all $s, t \geq 0$. Moreover, every continuous $\psi : \mathcal{O} \to \mathbb{R}$ is uniformly continuous on the compact set $\overline{\mathcal{A}}$, and therefore $t \mapsto \psi(x(t))$ is uniformly continuous on $[0, \infty)$. □

With $\mathcal{H} = \mathcal{O}$, Proposition 6 reduces to the classical requirement that the closure of $\mathcal{L}_c$ be a compact subset of $\mathcal{O}$. This requirement does not follow from the closed-loop dissipation inequalities, which penalize only the realized transverse port errors: a bound on these errors need not bound the state, so $\mathcal{L}_c$ may reach the boundary of $\mathcal{O}$ or be unbounded. A restriction removes the need for compact sublevel sets, since only the intersection $\mathcal{L}_c \cap \mathcal{H}$ has to be compact. The restriction may depend on the initial condition, as the level set (40) does, since only the solution from x_0 must be confined. Fig. 5 illustrates the sets involved.

Remark 4 (Conserved quantities from rank-deficient channels)**.** *Linear conserved quantities arise when a rank-deficient channel feeds integrating dynamics. Suppose that $\dot{x}_1 = G_1 u_1$ for a constant matrix G_1, and let v satisfy $G_1^\top v \in \ker \Gamma$. Then $\frac{\mathrm{d}}{\mathrm{d}t} v^\top x_1 = -(\Gamma G_1^\top v)^\top y_2 = 0$, so $\chi(x) = v^\top x_1$ satisfies (39) for any output map h_1, and the level set (40) is an affine slice. Symmetrically, if $\dot{x}_2 = G_2 u_2$, then $G_2^\top v \in \ker \Gamma^\top$ makes $v^\top x_2$ conserved.*

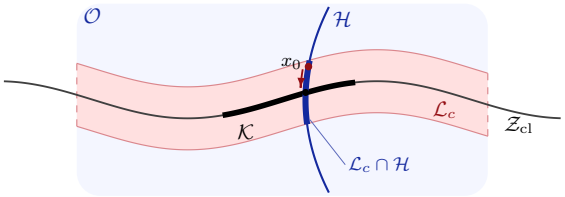


Figure 5: Sets of Section 5. The sublevel set $\mathcal{L}_c$ (red) is a tube about $\mathcal{Z}_{\mathrm{cl}}$ that reaches the boundary of $\mathcal{O}$ (blue); a restriction $\mathcal{H}$ meets it in a compact set (thick) that confines the solution from x_0 to a neighborhood of $\mathcal{K}$ (bold), by Proposition 6.

Diffusively coupled networks posses this structure [14, 16]. For integrator agents $\dot{x}_1 = u_1$ and $\Gamma = E^\top$, with E the incidence matrix of a connected graph, $\ker \Gamma = \mathrm{span}\{\mathbb{1}\}$ and the conserved quantity is the network average, which selects the agreement point in average consensus.

Remark 5 (Other confinement mechanisms)**.** *Proposition 6 uses only that the restriction $\mathcal{H}$ contains x_0 and is forward invariant relative to $\mathcal{O}$. Besides conserved level sets, such a restriction may be obtained as a sublevel set of a continuous function that is nonincreasing along the closed loop. For example, along a gradient flow $\dot{x} = -\nabla\phi(x)$ of a convex function ϕ, the distance to every minimizer is nonincreasing (Fejér monotonicity [30]), so small closed discs centered at a minimizer near x_0 are compact restrictions even when the storage sublevel sets are unbounded. In each case, conditions (i) and (ii) of Proposition 6 must still be verified.*

Port convergence thus rests on two ingredients that are separate from the dissipation inequality itself: compact confinement of the closed-loop state, provided by Proposition 6, and a dissipation margin that relates the dissipated power to the distance from the feasible target. This separation mirrors classical Lyapunov arguments, where boundedness and a decrease condition are established independently, and it allows each ingredient to be verified with tools suited to the system at hand. With forward completeness and uniform continuity ensured by compact confinement, we next establish port convergence under a dissipation margin.

5.2 Port convergence under a Dissipation Margin

Port convergence means that the realized inputs and outputs of Σ_1 and Σ_2 approach their prescribed manifolds:

$$\sum_{i=1}^{2} \Big(\mathrm{dist}\big(u_i(x(t)), \mathcal{M}_{u_i}\big)^2 + \mathrm{dist}\big(y_i(x(t)), \mathcal{M}_{y_i}\big)^2 \Big) \to 0. \tag{41}$$

By the feedback law (19), these inputs are determined by the realized outputs $h_\Gamma(x) = \big(y_1(x), y_2(x)\big)$, and all

four ports lie on their prescribed manifolds precisely when $h_\Gamma(x)$ belongs to the joint feasibility set $\mathcal{F}_\Gamma = \mathcal{F}_{fc} \times \mathcal{F}_{rc}$ of (25). Since $\| - \Gamma^\top \| = \|\Gamma\|$, comparing $h_\Gamma(x)$ with any point of $\mathcal{F}_\Gamma$ gives

$$\sum_{i=1}^{2} \big(\operatorname{dist}(u_i(x), \mathcal{M}_{u_i})^2 + \operatorname{dist}(y_i(x), \mathcal{M}_{y_i})^2\big) \leq (1 + \|\Gamma\|^2) \operatorname{dist}(h_\Gamma(x), \mathcal{F}_\Gamma)^2.$$

Consequently, a sufficient condition for (41) is that the distance $\operatorname{dist}\big(h_\Gamma(x(t)), \mathcal{F}_\Gamma\big)$ approaches zero as $t \to \infty$.

The next theorem combines compact confinement with a dissipation margin that controls the distance to $\mathcal{F}_\Gamma$.

Theorem 2 (Port convergence)**.** *Let $\mathcal{H}$, x_0, and c satisfy the hypotheses of Proposition 6. Suppose that there exist a continuous function $W : \mathcal{O} \to \mathbb{R}$ and a constant $\alpha_p > 0$ such that the following two conditions hold for every $x \in \mathcal{L}_c \cap \mathcal{H}$:*

(i) Dissipation: *the storage decreases at least at rate W,*

$$\mathrm{d}_x V_{\mathrm{cl}}[F_{\mathrm{cl}}(x)] \leq -W(x); \tag{42}$$

(ii) Margin: *W controls the distance to the feasible target,*

$$W(x) \geq \alpha_p \operatorname{dist}\big(h_\Gamma(x), \mathcal{F}_\Gamma\big)^2. \tag{43}$$

Then the solution initialized at x_0 is forward complete and satisfies

$$\lim_{t\to\infty} \operatorname{dist}\big(h_\Gamma(x(t)), \mathcal{F}_\Gamma\big) = 0. \tag{44}$$

Proof. By Proposition 6, the solution is forward complete, remains in $\mathcal{L}_c \cap \mathcal{H}$, and $t \mapsto W(x(t))$ is uniformly continuous. Along this solution, (43) gives $W(x(t)) \geq 0$. Integrating (42) over $[0, T]$ yields

$$\int_0^T W(x(t))\,\mathrm{d}t \leq V_{\mathrm{cl}}(x_0) - V_{\mathrm{cl}}(x(T)) \leq V_{\mathrm{cl}}(x_0).$$

Thus, $W(x(\cdot)) \in L^1([0,\infty))$, so Barbalat's lemma [6, Lemma 8.2] gives $\lim_{t\to\infty} W(x(t)) = 0$. The limit (44) then follows from (43). □

Theorem 2 applies whenever a dissipation margin (43) is available. When the residual supply vanishes, $W = Q$ is a natural choice, and output strictness of both subsystems supplies the margin.

Corollary 2 (Port convergence under exact compatibility)**.** *Suppose that the interconnection is exactly compatible on $\Omega \subseteq \mathcal{D}_\Gamma$, that $\varepsilon_1 > 0$ and $\varepsilon_2 > 0$, and that $\mathcal{O} \subseteq h_\Gamma^{-1}(\Omega)$. Then the conclusions of Theorem 2 hold with $W = Q$ and $\alpha_p = \min\{\varepsilon_1, \varepsilon_2\}$ for every restriction $\mathcal{H}$, initial condition $x_0 \in \mathcal{H}$, and $c \geq V_{\mathrm{cl}}(x_0)$ satisfying condition (ii) of Proposition 6.*

Proof. Proposition 4(i) gives (42) with $W = Q$ on $\mathcal{O}$, which also yields condition (i) of Proposition 6. Dropping the nonnegative input-strictness terms in Q and using (32) gives $Q(x) \geq \min\{\varepsilon_1, \varepsilon_2\}\big(\|e_{y_1}(x)\|^2 + \|e_{y_2}(x)\|^2\big) = \min\{\varepsilon_1, \varepsilon_2\} \operatorname{dist}\big(h_\Gamma(x), \mathcal{F}_\Gamma\big)^2$, which is (43). □

When the residual transverse supply $s_\delta(x)$ in (29) does not vanish, applying Theorem 2 requires controlling this residual. By Subsection 4.3, for perturbations $\xi = x - p$ about a feasible state $p \in \mathcal{Z}_{\mathrm{cl}}$, the residual reduces locally to $s_\delta(p+\xi) = \xi^\top R_p \xi + o(\|\xi\|^2)$. The next two subsections address the cases where the quadratic residual vanishes ($R_p = 0$) and where it remains nonzero ($R_p \neq 0$), respectively.

5.3 Port Convergence under Quadratic Cancellation

When $R_p = 0$ on a compact feasible set, Proposition 5(ii) bounds the residual supply by a cubic function of the distance to that set, and a quadratic dissipation margin absorbs it sufficiently close to the set. The argument uses the sets of Fig. 5: a nonempty compact set $\mathcal{K} \subseteq \mathcal{Z}_{\mathrm{cl}}$ on which $R_p = 0$, the open domain $\mathcal{O}$ of Proposition 5, which contains $\mathcal{K}$, and a restriction $\mathcal{H} \subseteq \mathcal{O}$, on which alone the margin is required. Since the estimates are local, $\mathcal{O}$ may have to be shrunk to a smaller open neighborhood of $\mathcal{K}$. We then replace $\mathcal{H}$ by its intersection with the smaller set, which is again a restriction: a solution in this intersection stays in $\mathcal{H}$ as long as it stays in the smaller set.

Corollary 3 (Cubic residual absorption)**.** *Suppose that the hypotheses of Proposition 5 hold with $R_p = 0$ for every $p \in \mathcal{K}$. Let $\mathcal{H} \subseteq \mathcal{O}$ be a restriction, and suppose that Q satisfies the quadratic margin*

$$Q(x) \geq \alpha \operatorname{dist}(x, \mathcal{K})^2, \quad \forall x \in \mathcal{H}, \tag{45}$$

for some $\alpha > 0$. Then, after shrinking $\mathcal{O}$ to a sufficiently small open neighborhood of $\mathcal{K}$, the following statements hold.

(i) Net margin. *There exists a constant $L > 0$ such that, for every $x \in \mathcal{H}$,*

$$\begin{aligned} Q(x) - s_\delta(x) &\geq \tfrac{\alpha}{2} \operatorname{dist}(x, \mathcal{K})^2 \\ &\geq \tfrac{\alpha}{2L^2} \operatorname{dist}\big(h_\Gamma(x), \mathcal{F}_\Gamma\big)^2. \end{aligned} \tag{46}$$

(ii) Port convergence. *Let $x_0 \in \mathcal{H}$ and $c \geq V_{\mathrm{cl}}(x_0)$ be such that the closure of $\mathcal{L}_c \cap \mathcal{H}$ is a compact subset of $\mathcal{O}$. Then the solution initialized at x_0 is forward complete and satisfies (44).*

Proof. (i) Let r and C be the constants of Proposition 5, so that $|s_\delta(x)| \leq C \operatorname{dist}(x, \mathcal{K})^3$ whenever $\operatorname{dist}(x, \mathcal{K}) < r$, and the closed r-neighborhood $\mathcal{N}$ of $\mathcal{K}$ is a compact subset of $\mathcal{O}$. Let $L > 0$ be an upper bound for $\|\mathrm{d}_x h_\Gamma\|$ on $\mathcal{N}$. For $\operatorname{dist}(x, \mathcal{K}) < r$, let $p \in \mathcal{K}$ be a nearest point to

x; the segment from p to x lies in $\mathcal{N}$ and $h_\Gamma(p) \in \mathcal{F}_\Gamma$, so $\operatorname{dist}(h_\Gamma(x), \mathcal{F}_\Gamma) \le \|h_\Gamma(x) - h_\Gamma(p)\| \le L \operatorname{dist}(x, \mathcal{K})$. Shrinking $\mathcal{O}$ to $\{x : \operatorname{dist}(x, \mathcal{K}) < \min\{r, \alpha/(2C)\}\}$ and combining these bounds with (45) gives (46).

(ii) By (29) and (46), the continuous function $W := Q - s_\delta$ satisfies (42)–(43) on $\mathcal{H}$ with $\alpha_p = \alpha/(2L^2)$. In particular, $\mathrm{d}_x V_{\mathrm{cl}}[F_{\mathrm{cl}}(x)] \le 0$ on $\mathcal{H}$, so the hypotheses of Proposition 6 hold, and Theorem 2 yields the claim. □

The margin (45) is the essential hypothesis of Corollary 3. Since Q penalizes only the realized transverse port errors, it may fail to detect certain state deviations from $\mathcal{K}$ at quadratic order. The following proposition identifies a geometric obstruction to obtaining such a margin from Q on a full neighborhood.

Proposition 7 (An obstruction to quadratic dissipation margins)**.** *Let $\mathcal{K} \subseteq \mathcal{Z}_{\mathrm{cl}}$ be a C^2 embedded submanifold in a neighborhood of $p \in \mathcal{K}$, and write $y_i^\star := y_i(p)$ and $u_i^\star := u_i(p)$ for $i = 1, 2$.*

If there exists a state direction $v \in N_p\mathcal{K} \setminus \{0\}$ such that $\mathrm{d}_p y_i[v] \in T_{y_i^\star}\mathcal{M}_{y_i}$ and $\mathrm{d}_p u_i[v] \in T_{u_i^\star}\mathcal{M}_{u_i}$ for $i = 1, 2$, then there is no constant $\alpha > 0$ such that $Q(x) \ge \alpha \operatorname{dist}(x, \mathcal{K})^2$ for every x in a neighborhood of p.

Proof. For sufficiently small τ, the state $x_\tau := p + \tau v$ belongs to $\mathcal{D}_{\mathrm{cl}}$. By Lemma 1(i), we have $\mathrm{d}_p e_{y_i}[v] = \mathrm{P}_{N_{y_i^\star}\mathcal{M}_{y_i}} \mathrm{d}_p y_i[v] = 0$ and, likewise, $\mathrm{d}_p e_{u_i}[v] = 0$ for $i = 1, 2$, where the differentials are taken along the realized port maps.

Since $p \in \mathcal{K} \subseteq \mathcal{Z}_{\mathrm{cl}}$, all four realized transverse port errors vanish at p. Their first-order expansions therefore imply $e_{y_i}(x_\tau) = o(|\tau|)$ and $e_{u_i}(x_\tau) = o(|\tau|)$ for $i = 1, 2$, so $Q(x_\tau) = o(\tau^2)$.

On the other hand, since $v \in N_p\mathcal{K} \setminus \{0\}$, the local nearest-point projection onto $\mathcal{K}$ yields

$$\operatorname{dist}(x_\tau, \mathcal{K})^2 = \tau^2 \|v\|^2 + o(\tau^2),$$

with $\|v\| > 0$. Thus, $\lim_{\tau \to 0,\ \tau \neq 0} \dfrac{Q(x_\tau)}{\operatorname{dist}(x_\tau, \mathcal{K})^2} = 0$, which excludes every positive quadratic lower bound on a neighborhood of p. □

If the realized transverse error maps are C^2, the same argument gives the sharper estimate $Q(x_\tau) = O(\tau^4)$. Proposition 7 concerns a full neighborhood of p. This is why Corollary 3 requires the margin only on the restriction $\mathcal{H}$: a restriction that excludes the obstructing directions v can recover the margin, as in Subsection 6.3.

5.4 Net Dissipation with Nonzero Quadratic Residuals

When $R_p \neq 0$, the quadratic residual $\xi^\top R_p \xi$ in (38) has the same order as $Q(p + \xi) = \frac{1}{2}\xi^\top \nabla^2 Q(p) \xi + o(\|\xi\|^2)$. Unlike in Corollary 3, shrinking the neighborhood absorbs only higher-order terms and cannot eliminate $\xi^\top R_p \xi$, so $\frac{1}{2}\nabla^2 Q(p)$ must be compared directly with R_p. The sets are those of Subsection 5.3: a compact feasible set $\mathcal{K}$, a domain $\mathcal{O}$ that may be shrunk about $\mathcal{K}$, and a restriction $\mathcal{H}$. When $\mathcal{H}$ is a C^2 submanifold containing $\mathcal{K}$, the comparison can be carried out on its tangent space, in directions normal to $\mathcal{K}$.

Theorem 3 (A restricted net-dissipation criterion)**.** *Suppose that the realized maps $e_{y_1}, e_{y_2}, \delta_{fc}, \delta_{rc}$ are C^2 on $\mathcal{O}$. Let $\mathcal{H} \subseteq \mathcal{O}$ be a restriction that is a C^2 embedded submanifold, and let $\mathcal{K} \subseteq \mathcal{H} \cap \mathcal{Z}_{\mathrm{cl}}$ be a nonempty compact C^2 embedded submanifold without boundary. For $p \in \mathcal{K}$, define the net quadratic form*

$$H_p := \tfrac{1}{2}\nabla^2 Q(p) - R_p, \tag{47}$$

and suppose that there exists $\lambda > 0$ such that

$$v^\top H_p v \ge \lambda \|v\|^2, \quad \forall p \in \mathcal{K},\ v \in T_p\mathcal{H} \cap N_p\mathcal{K}. \tag{48}$$

Then, after shrinking $\mathcal{O}$ to a sufficiently small open neighborhood of $\mathcal{K}$, the following statements hold.

(i) Net margin. *There exist constants $\kappa, L > 0$ such that, for every $x \in \mathcal{H}$,*

$$\begin{aligned} Q(x) - s_\delta(x) &\ge \kappa \operatorname{dist}(x, \mathcal{K})^2 \\ &\ge \tfrac{\kappa}{L^2} \operatorname{dist}\big(h_\Gamma(x), \mathcal{F}_\Gamma\big)^2. \end{aligned} \tag{49}$$

(ii) Port convergence. *Let $x_0 \in \mathcal{H}$ and $c \ge V_{\mathrm{cl}}(x_0)$ be such that the closure of $\mathcal{L}_c \cap \mathcal{H}$ is a compact subset of $\mathcal{O}$. Then the solution initialized at x_0 is forward complete and satisfies (44).*

(iii) Local exponential convergence. *Suppose that $\mathcal{K} = \{p\}$ with $F_{\mathrm{cl}}(p) = 0$, that $\mathcal{H}$ is relatively closed in $\mathcal{O}$, and that there exist $c_1, c_2 > 0$ such that*

$$c_1 \|x - p\|^2 \le V_{\mathrm{cl}}(x) - V_{\mathrm{cl}}(p) \le c_2 \|x - p\|^2 \tag{50}$$

for all $x \in \mathcal{H}$ near p. Then, for every initial state $x(0) \in \mathcal{H}$ sufficiently close to p, the solution remains in $\mathcal{O}$, achieves port convergence, and satisfies

$$\|x(t) - p\| \le \sqrt{c_2/c_1}\, e^{-\kappa t/(2c_2)} \|x(0) - p\|, \quad t \ge 0.$$

Proof. (i) By the feedback law, all e_{y_i} and e_{u_i} are C^2 on $\mathcal{O}$ and vanish on $\mathcal{K} \subseteq \mathcal{Z}_{\mathrm{cl}}$, so Q is C^2 on $\mathcal{O}$ with $Q = 0$ and $\mathrm{d}_p Q = 0$ on $\mathcal{K}$. By Taylor's formula, Proposition 5(i), and compactness of $\mathcal{K}$,

$$Q(p + \xi) - s_\delta(p + \xi) = \xi^\top H_p \xi + o(\|\xi\|^2)$$

uniformly for $p \in \mathcal{K}$. For $x \in \mathcal{H}$ near $\mathcal{K}$, let $p = r_{\mathcal{K}}(x)$ be its nearest point on $\mathcal{K}$ and $\xi = x - p$, so that $\xi \in N_p\mathcal{K}$. Since $\mathcal{H}$ is a C^2 embedded submanifold and $\mathcal{K} \subseteq \mathcal{H}$ is compact, finitely many slice charts of $\mathcal{H}$ cover $\mathcal{K}$, so near $\mathcal{K}$ the set $\mathcal{H}$ is the graph of a C^2 map over $T_p\mathcal{H}$ with constants uniform in $p \in \mathcal{K}$; embeddedness excludes other sheets of $\mathcal{H}$ accumulating on $\mathcal{K}$. Hence there is $C_{\mathcal{H}} > 0$ such that $\eta := \mathrm{P}_{T_p\mathcal{H}}\xi$ satisfies $\|\xi - \eta\| \le C_{\mathcal{H}} \|\xi\|^2$. Since $T_p\mathcal{K} \subseteq T_p\mathcal{H}$,

the projection $\mathrm{P}_{T_p\mathcal{H}}$ maps $N_p\mathcal{K}$ into itself, so $\eta \in T_p\mathcal{H} \cap N_p\mathcal{K}$ and (48) gives $\eta^\top H_p \eta \geq \lambda\|\eta\|^2$. As H_p is bounded uniformly on $\mathcal{K}$, $\xi^\top H_p \xi = \eta^\top H_p \eta + O(\|\xi\|^3)$, and $\|\eta\|^2 = \|\xi\|^2 - \|\xi - \eta\|^2 = \|\xi\|^2 + O(\|\xi\|^4)$, so $Q(x) - s_\delta(x) \geq (\lambda + o(1))\|\xi\|^2$. Shrinking $\mathcal{O}$ gives the first inequality in (49) with $\kappa = \lambda/2$, and the Lipschitz bound for h_Γ in the proof of Corollary 3(i) gives the second.

(ii) As in the proof of Corollary 3(ii), (29) and (49) show that $W := Q - s_\delta$ satisfies (42)–(43) on $\mathcal{H}$ and that $\mathrm{d}_x V_{\mathrm{cl}}[F_{\mathrm{cl}}(x)] \leq 0$ there, so Proposition 6 and Theorem 2 yield the claim.

(iii) Let $B \subseteq \mathcal{O}$ be a closed ball of radius $\rho > 0$ about p on which (50) holds for $x \in \mathcal{H}$, and let B° be its interior. By the left inequality in (50), p is a strict minimizer of V_{cl} on $\mathcal{H} \cap B$. Fix $\varepsilon \in (0, c_1\rho^2)$ and let $\mathcal{S} := \{x \in \mathcal{H} \cap B : V_{\mathrm{cl}}(x) \leq V_{\mathrm{cl}}(p) + \varepsilon\}$. Since $\mathcal{H}$ is relatively closed in $\mathcal{O}$ and $B \subseteq \mathcal{O}$ is compact, $\mathcal{H} \cap B$ and hence $\mathcal{S}$ are compact, and every $x \in \mathcal{S}$ satisfies $\|x - p\| \leq (\varepsilon/c_1)^{1/2} < \rho$, so $\mathcal{S} \subseteq B^\circ$. We apply Proposition 6 with $\mathcal{O}$, $\mathcal{H}$, and c replaced by B°, $\mathcal{H} \cap B^\circ$, and $V_{\mathrm{cl}}(p) + \varepsilon$. The set $\mathcal{H} \cap B^\circ$ is a restriction relative to B°, because $\mathcal{H}$ is one relative to $\mathcal{O} \supseteq B^\circ$; condition (i) follows from (29) and (49); and condition (ii) holds because the relevant sublevel set is $\mathcal{S}$, which is compact and contained in B°. If $x(0) \in \mathcal{H}$ and $\|x(0) - p\| \leq (\varepsilon/c_2)^{1/2}$, then $x(0) \in B^\circ$, since $c_1 \leq c_2$, and $V_{\mathrm{cl}}(x(0)) \leq V_{\mathrm{cl}}(p) + \varepsilon$ by the right inequality in (50). Hence the solution is forward complete and remains in $\mathcal{S} \subseteq B^\circ$. There, (29), (49), and (50) give $\frac{\mathrm{d}}{\mathrm{d}t}\big(V_{\mathrm{cl}}(x) - V_{\mathrm{cl}}(p)\big) \leq -\kappa\|x - p\|^2 \leq -(\kappa/c_2)\big(V_{\mathrm{cl}}(x) - V_{\mathrm{cl}}(p)\big)$; integrating and using (50) gives the exponential bound, and (49) gives port convergence. □

Since Q and s_δ vanish on $\mathcal{Z}_{\mathrm{cl}}$, the net margin (49) implies $\mathcal{H} \cap \mathcal{Z}_{\mathrm{cl}} \cap \mathcal{O} = \mathcal{K}$ for the shrunken $\mathcal{O}$. Thus condition (48) itself guarantees that $\mathcal{K}$ is isolated in $\mathcal{H} \cap \mathcal{Z}_{\mathrm{cl}}$, and it fails whenever $\mathcal{H}$ contains feasible states outside $\mathcal{K}$ arbitrarily close to it. As in Proposition 7, the restriction must therefore exclude the feasible directions normal to $\mathcal{K}$.

Theorem 3 also applies when $R_p = 0$, in which case $H_p = \frac{1}{2}\nabla^2 Q(p)$ and (48) provides a second-order test for the restricted margin (45) on $\mathcal{H}$. Since the transverse port errors vanish on $\mathcal{K}$, $\frac{1}{2}v^\top \nabla^2 Q(p) v = \sum_{i=1}^{2} \big(\varepsilon_i \|\mathrm{d}_p e_{y_i}[v]\|^2 + \gamma_i \|\mathrm{d}_p e_{u_i}[v]\|^2\big)$, so the test requires that no direction in $T_p\mathcal{H} \cap N_p\mathcal{K} \setminus \{0\}$ be annihilated by the weighted port-error differentials. With $\mathcal{H} = \mathcal{O}$ and all strictness indices positive, this is exactly the negation of the hypothesis of Proposition 7 at every $p \in \mathcal{K}$, since $\mathrm{d}_p e_{y_i}[v] = 0$ if and only if $\mathrm{d}_p y_i[v] \in T_{y_i(p)}\mathcal{M}_{y_i}$, and likewise for the inputs. Under the regularity hypotheses of Theorem 3, the obstruction is then sharp: a quadratic margin on a full neighborhood of $\mathcal{K}$ exists if and only if no obstructing direction exists. Moreover, Corollary 3 and Theorem 3 remain valid if Q is replaced by any continuous (respectively, C^2) function $W_0 : \mathcal{O} \to \mathbb{R}_{\geq 0}$ satisfying

$$\mathrm{d}_x V_{\mathrm{cl}}[F_{\mathrm{cl}}(x)] \leq s_\delta(x) - W_0(x), \tag{51}$$

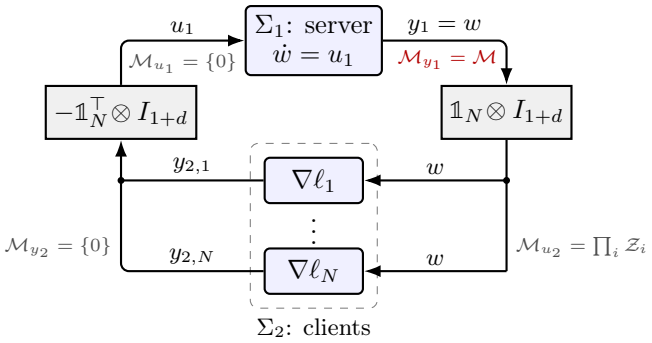


Figure 6: Federated learning as the interconnection of Fig. 3: the forward channel broadcasts the server model, and the reverse channel aggregates the client gradients. Port targets are in gray; only the server target $\mathcal{M}$ (red) is varied.

with $W_0 = 0$ on $\mathcal{K}$ in Theorem 3, for instance when an unrelaxed subsystem supply term is retained in place of a conservative quadratic bound Q.

In summary, applying the results of this section involves three steps. First, choose a restriction $\mathcal{H}$, for instance a conserved level set (40), and verify compact confinement (Proposition 6). Second, identify the residual case of Table 1. Third, verify the corresponding margin: output strictness under exact compatibility (Corollary 2), the quadratic margin (45) on $\mathcal{H}$ when $R_p = 0$ (Corollary 3), or the net positivity condition (48) when $R_p \neq 0$ (Theorem 3). Section 6 follows these steps.

6 Case Study: Federated Learning With Heterogeneous Data

In federated learning, clients exchange model parameters or gradients with a server rather than raw data [31, 32]. For overparameterized models the interpolating models can form a manifold [5], and the central difficulty, data heterogeneity [32], has a geometric meaning: each client has its own zero-loss manifold, and the clients agree only on their intersection. Throughout this section, the closed loop is fixed and only the server output target is varied. We show that heterogeneity precludes channel compatibility for every server target, yet port convergence can still be certified: via quadratic cancellation when the server target is tangent to the shared manifold, and via net dissipation on an invariant set when a nontangential target leaves a nonzero quadratic residual.

6.1 Model, Transversal Passivity, and Compatibility

Learning problem. Consider N clients that jointly fit a linear predictor $\xi \mapsto \langle \theta, \xi \rangle$, $\theta \in \mathbb{R}^d$, through the factorization

$\theta = ab$ with parameters $w = (a, b) \in \mathbb{R} \times \mathbb{R}^d$. This scalar–vector factorization is the simplest low-rank parametrization, and it makes the problem overparameterized, since each nonzero θ is represented by a curve of parameters w. Client i holds one sample (ξ_i, τ_i) generated by a ground truth $\theta^\star$, with $\tau_i = \langle \theta^\star, \xi_i \rangle \neq 0$, and has the squared loss

$$\ell_i(w) = \tfrac{1}{2} r_i(w)^2, \quad r_i(w) := a\langle b, \xi_i \rangle - \tau_i,$$

with gradient $\nabla \ell_i = r_i \nabla r_i$ and $\nabla r_i = (\langle b, \xi_i \rangle, a\xi_i)$. We assume that the samples span $\mathbb{R}^d$, so that $X := \sum_i \xi_i \xi_i^\top \succ 0$, and write $m := \|\theta^\star\|$ and $\hat{\theta} := \theta^\star / m$.

Zero-loss manifolds and heterogeneity. Client i attains zero loss on $\mathcal{Z}_i := \{r_i = 0\}$. In the coordinates $s_i := \langle b, \xi_i \rangle / \|\xi_i\|$ and $\mu_i := \tau_i / \|\xi_i\|$, this is the hyperbolic cylinder $\{as_i = \mu_i\}$. The clients agree only on the common zero-loss set

$$\mathcal{Z} := \textstyle\bigcap_i \mathcal{Z}_i = \{ab = \theta^\star\},$$

where the equality uses that the ξ_i span $\mathbb{R}^d$. Since $ab = \theta^\star$ forces b to be parallel to $\hat{\theta}$, this set lies in the plane

$$\mathcal{P} := \{(a, \beta\hat{\theta}) : a, \beta \in \mathbb{R}\},$$

where it is the hyperbola $a\beta = m$ with reach $\sqrt{2m}$. We use (a, β), with $\beta = \langle b, \hat{\theta} \rangle$, as coordinates on $\mathcal{P}$. For $d = 1$, all $\mathcal{Z}_i$ coincide with $\mathcal{Z}$. For $d \geq 2$, they do not, since their normals ∇r_i point in different directions; this is the geometric form of data heterogeneity.

Training as a feedback interconnection. In federated SGD with full participation [31], taken in continuous time, the server broadcasts its model w, each client returns its gradient at w, and the server integrates the negative sum of these gradients. This is the interconnection of Fig. 3, detailed in Fig. 6, with

$$\Sigma_1 : \ \dot{w} = u_1, \ y_1 = w, \quad \Sigma_2 : \ y_{2,i} = \nabla \ell_i(u_{2,i}),$$

and the broadcast channel $\Gamma = \mathbb{1}_N \otimes I_{1+d}$. The server has no feedthrough and the clients are static, so the realized ports are explicit and Assumption 1 holds. The closed loop is the gradient flow of the global loss $L := \sum_i \ell_i$,

$$\dot{a} = -\langle b, Xg \rangle, \quad \dot{b} = -aXg, \quad g := ab - \theta^\star, \tag{52}$$

and it conserves the imbalance $\chi := a^2 - \|b\|^2$ [33], since $\dot{\chi} = -2a\langle b, Xg \rangle + 2a\langle b, Xg \rangle = 0$. The hyperbola $\mathcal{Z}$ has two branches, $a > 0$ and $a < 0$, related by the sign change $w \mapsto -w$. Along $\mathcal{Z}$ we have $b = \theta^\star / a$, so $\chi = a^2 - m^2/a^2$ increases strictly with $|a|$ on each branch. Hence each imbalance value χ_0 is attained at exactly one interpolant on each branch. We denote the one with $a > 0$ by p_{χ_0}, so that the other is $-p_{\chi_0}$. Since χ is conserved, p_{χ_0} is the only interpolant with $a > 0$ that a trajectory starting at imbalance χ_0 can reach. The balanced case $\chi_0 = 0$ gives $p_0 = (\sqrt{m}, \theta^\star / \sqrt{m})$, the interpolant of smallest norm, $\|p_0\| = \sqrt{2m}$.

Port targets. The learning objective fixes $\mathcal{M}_{u_1} = \mathcal{M}_{y_2} = \{0\}$, so that the aggregated gradient vanishes, and $\mathcal{M}_{u_2} = \prod_i \mathcal{Z}_i$, so that each client receives a model it interpolates. The server target $\mathcal{M}_{y_1} =: \mathcal{M}$ is the design choice. Denoting the nearest-point projection and transverse error of a manifold $\mathcal{M}$ by $r_{\mathcal{M}}$ and $e_{\mathcal{M}}$, Lemma 1(ii) with $V_1 = \frac{1}{2}\operatorname{dist}(w, \mathcal{M})^2$ gives $\dot{V}_1 = \langle u_1, e_{\mathcal{M}}(w) \rangle$, so the server is $(\{0\}, \mathcal{M})$-passive for every $\mathcal{M}$. Each client is a static map, so its transversal passivity reduces to a pointwise inequality between its transverse input error and its gradient. The reach of $\mathcal{Z}_i$ is $\sqrt{2|\mu_i|}$, the radius of curvature of the hyperbola $as_i = \mu_i$ at its vertex. We work on the tube of half this radius, $\eta_i := \frac{1}{2}\sqrt{2|\mu_i|}$, truncated to a ball of radius $R > 0$,

$$\mathcal{N}_i := \big\{u : \ \operatorname{dist}(u, \mathcal{Z}_i) < \eta_i, \ \|u\| < R\big\}. \tag{53}$$

The truncation is needed because the strictness index below decreases with R.

Lemma 2 (Client transversal passivity)**.** *Let $R > 0$. Then the following statements hold.*

(i) Aiming: *for every client i and every $u \in \mathcal{N}_i$,*

$$\langle e_{\mathcal{Z}_i}(u), \nabla \ell_i(u) \rangle \geq \frac{2}{3\|\xi_i\|^2 R^2} \|\nabla \ell_i(u)\|^2.$$

(ii) Transversal passivity: Σ_2 *is output-strictly $(\prod_i \mathcal{Z}_i, \{0\})$-passive on $\prod_i \mathcal{N}_i$, with storage $V_2 = 0$ and index $\varepsilon_2 = 2/(3 \max_i \|\xi_i\|^2 R^2)$.*

Proof. Both sides of (i) vanish on $\mathcal{Z}_i$, so let $u \in \mathcal{N}_i \setminus \mathcal{Z}_i$. Since $\ell_i = \frac{\|\xi_i\|^2}{2}(as_i - \mu_i)^2$ depends on u only through $z = (a, s_i)$, it suffices to treat the planar hyperbola $\{as = \mu_i\}$ with $\mu_i > 0$. Write the transverse error of z as νn, with base point q, unit normal $n = (q_2, q_1)/\|q\|$, and $|\nu| < \eta_i$. Expanding the quadratic $as - \mu_i$ about q gives $\langle e_{\mathcal{Z}_i}, \nabla \ell_i \rangle / \|\nabla \ell_i\|^2 = (1+2\zeta)/\big(\|\xi_i\|^2(1+\zeta)\|z\|^2\big)$ with $\zeta := \nu\mu_i / \|q\|^3$, and $|\zeta| < 1/4$ because $\|q\| \geq \sqrt{2\mu_i}$. With $\|z\| < R$ this gives (i), and summing (i) over the clients gives (ii). □

The inequality of Lemma 2 is the aiming condition of overparameterized learning [34], although the clients are neither passive nor monotone, since ℓ_i is nonconvex. A client aims at its own $\mathcal{Z}_i$, not at $\mathcal{Z}$: the supply $\langle e_{\mathcal{Z}}, \nabla \ell_i \rangle$ can be negative, at fourth order, so $\mathcal{M}_{u_2} = \mathcal{Z}^N$ would not work.

Since $\delta_{rc} \equiv 0$ and $e_{u_2}(\Gamma w) = (e_{\mathcal{Z}_i}(w))_i$, the forward defect, the residual supply, and the feasible sets are

$$\begin{aligned} &\delta_{fc}(w) = \big(r_{\mathcal{M}}(w) - r_{\mathcal{Z}_i}(w)\big)_i, \\ &s_\delta(w) = \textstyle\sum_i \langle \delta_{fc,i}(w), \nabla \ell_i(w) \rangle, \\ &\mathcal{F}_\Gamma = (\mathcal{M} \cap \mathcal{Z}) \times \{0\}, \quad \mathcal{Z}_{\mathrm{cl}} = \mathcal{M} \cap \mathcal{Z}, \end{aligned} \tag{54}$$

and Theorem 1 gives $\dot{V}_{\mathrm{cl}} \leq s_\delta - Q$ with $V_{\mathrm{cl}} = V_1$ and $Q = \varepsilon_2 \sum_i \|\nabla \ell_i\|^2$.

Proposition 8 (Heterogeneity and compatibility)**.** *Let $\mathcal{M}$ be a server target and let $p = (a_p, b_p) \in \mathcal{M} \cap \mathcal{Z}$ be a feasible point. Then the following statements hold.*

(i) First-order compatibility: *the interconnection is first-order compatible along* $\{p\}\times\{0\}$ *if and only if* $T_p\mathcal{M} = T_p\mathcal{Z}_i$ *for every client* i*. For* $d \geq 2$*, this condition fails for every server target.*

(ii) Quadratic cancellation: $R_p = 0$ *if and only if* $T_p\mathcal{M} \subseteq T_p\mathcal{Z}$.

Proof. (i) Lemma 1(i) gives $\mathrm{d}_p\delta_{fc,i} = \mathrm{P}_{T_p\mathcal{M}} - \mathrm{P}_{T_p\mathcal{Z}_i}$, which vanishes if and only if $T_p\mathcal{M} = T_p\mathcal{Z}_i$. Since each $\mathcal{Z}_i$ is a hypersurface, this equality for every i forces the normals $\nabla r_i(p) = (\langle b_p, \xi_i\rangle, a_p\xi_i)$ to be pairwise parallel. As $a_p \neq 0$ on $\mathcal{Z}$, the ξ_i are then pairwise parallel and span at most a line, which contradicts $X \succ 0$ when $d \geq 2$. (ii) Since $r_i(p) = 0$, $\mathrm{d}_p\nabla\ell_i[v] = \langle\nabla r_i(p), v\rangle\nabla r_i(p)$, which is normal to $\mathcal{Z}_i$. In (35), $v^\top R_p v$ therefore reduces to $v^\top \mathrm{P}_{T_p\mathcal{M}} G v$ with $G := \sum_i \nabla r_i(p)\nabla r_i(p)^\top \succeq 0$ and $\ker G = \ker \mathrm{d}_p(ab) = T_p\mathcal{Z}$. It vanishes for all v iff $T_p\mathcal{M} \subseteq \ker G$. □

In words, first-order compatibility would require the server target to be tangent to every client's zero-loss manifold at once, which heterogeneity rules out, whereas quadratic cancellation only requires it to be tangent to their intersection $\mathcal{Z}$. We compare three server targets, each encoding a different specification for the trained server model. The target $\mathcal{M} = \mathcal{Z}$ asks only that the server model interpolate the data of all clients, which is the federated learning objective itself. The circles $\mathcal{C}_\varrho := \{w \in \mathcal{P} : \|w\| = \varrho\}$ in addition prescribe the norm of the trained parameters, a simple way to express a preference among the many interpolants of an overparameterized model. The radius $\varrho = \sqrt{2m}$ selects the minimum-norm interpolant p_0, while a radius $\varrho > \sqrt{2m}$ prescribes a larger norm, attained by two interpolants on each branch of $\mathcal{Z}$. Geometrically, the three targets meet $\mathcal{Z}$ along all of $\mathcal{Z}$, tangentially at a single point, and transversally, respectively. By Proposition 8(ii), this geometry determines the residual: $R_p = 0$ for the first two targets and $R_p \neq 0$ for the third; see Fig. 7. The first two instantiate the row $R_p = 0$ of Table 1 (Corollary 3) and the third the row $R_p \neq 0$ (Theorem 3). The row of exact compatibility is absent, since exact compatibility implies first-order compatibility, which heterogeneity rules out for $d \geq 2$.

6.2 Shared Target: Quadratic Cancellation and Implicit Bias

Let $\mathcal{M} = \mathcal{Z}$. The defect grows linearly (Fig. 7(d)), yet by Proposition 8(ii) we have $R_p = 0$ along all of $\mathcal{Z}$: the linear part of each client's defect is tangent to $\mathcal{Z}_i$, while its output varies normally to $\mathcal{Z}_i$. Work on the tube $\mathcal{O} := \{w : \mathrm{dist}(w, \mathcal{Z}) < \eta,\ \|w\| < R\}$ with $\eta < \min\{\sqrt{2m}, \min_i \eta_i\}$, where all subsystem inequalities hold because $\mathrm{dist}(w, \mathcal{Z}_i) \leq \mathrm{dist}(w, \mathcal{Z})$.

Because $\mathcal{Z}$ is unbounded, the storage sublevel sets $\mathcal{L}_c$ reach the boundary of $\mathcal{O}$; the conserved imbalance confines them through the restriction $\mathcal{H} = \{\chi = \chi_0\}$ (Proposition 6). With $p = r_{\mathcal{Z}}(w)$ and $e = w - p$, $g = \mathrm{d}_p(ab)[e] + e_a e_b$ and $\|\mathrm{d}_p(ab)\| = \|p\|$ give

$$\|w\|^4 = \chi_0^2 + 4\|ab\|^2 \leq \chi_0^2 + 4\big(m + \eta\|w\| + \tfrac{3}{2}\eta^2\big)^2.$$

so for $R > \eta + \sqrt{|\chi_0| + 2m + 4\eta^2}$ and $c < \eta^2/2$, the closure of $\mathcal{L}_c \cap \{\chi = \chi_0\}$ is a compact subset of $\mathcal{O}$, on which $|a|$ and the first coordinate of p exceed some $a_- > 0$ in magnitude.Since $\mathrm{d}_p(ab)$ has smallest singular value at least a_- on $N_p\mathcal{Z}$ and $\|\nabla r_i\| \geq |a|\|\xi_i\|$, we obtain

$$\textstyle\sum_i \|\nabla\ell_i(w)\|^2 \geq c_{\mathcal{Z}}\|e\|^2 \tag{55}$$

for $\|e\| \leq a_-$, with $c_{\mathcal{Z}} := \frac{1}{4}\lambda_{\min}(X)\min_i\|\xi_i\|^2 a_-^4$. Corollary 3 now applies on the positive branch with $\mathcal{K} = \{p_{\chi_0}\} = \mathcal{Z}_{\mathrm{cl}} \cap \mathcal{H}$, where $R_{p_{\chi_0}} = 0$. Since χ is strictly monotone along each branch of $\mathcal{Z}$, the level set $\mathcal{H}$ is transverse to $\mathcal{Z}$ at p_{χ_0}, so $\|w - p_{\chi_0}\| \leq C_T\|e\|$ for $w \in \mathcal{H}$ near p_{χ_0} and some $C_T > 0$. Together with (55), this gives the margin (45), $Q(w) \geq \varepsilon_2 c_{\mathcal{Z}}\|e\|^2 \geq (\varepsilon_2 c_{\mathcal{Z}}/C_T^2)\|w - p_{\chi_0}\|^2$, on $\mathcal{H}$ near p_{χ_0}. Hence, for c small enough that the closure of $\mathcal{L}_c \cap \mathcal{H}$ lies in the neighborhood of Corollary 3, the ports converge. Since the solution remains in $\mathcal{H}$ near p_{χ_0} and $e(t) \to 0$, its ω-limit set lies in $\mathcal{Z} \cap \mathcal{H} = \{p_{\chi_0}\}$ there, so $w(t) \to p_{\chi_0}$, and symmetrically $w(t) \to -p_{\chi_0}$ on the negative branch. Heterogeneity leaves the local implicit bias of factorized models unchanged, since balanced initializations ($\chi_0 = 0$) converge to p_0. What changes is the certificate: quadratic cancellation replaces exact compatibility, which holds for homogeneous data.

6.3 Minimum-Norm Target: The Margin Obstruction

Let $\mathcal{M} = \mathcal{C}_{\sqrt{2m}}$, so $\mathcal{Z}_{\mathrm{cl}} = \mathcal{M} \cap \mathcal{Z} = \{p_0\}$ on the positive branch, with $T_{p_0}\mathcal{M} = T_{p_0}\mathcal{Z}$. Taking $\mathcal{K} = \{p_0\}$, Proposition 8(ii) gives $R_{p_0} = 0$, and Proposition 5(ii) yields $|s_\delta(w)| \leq C\|w - p_0\|^3$ near p_0 (Fig. 7(d)). For any nonzero $v \in T_{p_0}\mathcal{Z}$, however, $v \in N_{p_0}\mathcal{K} \setminus \{0\}$, while $\mathrm{d}_{p_0}y_1[v] \in T_{p_0}\mathcal{M}$, $\mathrm{d}_{p_0}u_{2,i}[v] \in T_{p_0}\mathcal{Z}_i$, and $\mathrm{d}_{p_0}y_2[v] = \mathrm{d}_{p_0}u_1[v] = 0$. Proposition 7 therefore excludes a quadratic margin $Q(w) \geq \alpha\|w - p_0\|^2$ on a full neighborhood of p_0, and the failure is genuine: since $\chi(p_0) = 0$, initializations with $\chi_0 \neq 0$ converge to $p_{\chi_0} \neq p_0$, so $\mathrm{dist}(y_1(t), \mathcal{M}) \to \|p_{\chi_0}\| - \sqrt{2m} > 0$ (Fig. 7(e)). On the balanced restriction $\mathcal{H} = \{\chi = 0\}$, whose normal $\nabla\chi(p_0)$ spans $T_{p_0}\mathcal{Z}$, transversality gives $\|w - p_0\| \leq 2\|e\|$ near p_0. Then (55) yields (45) with $\alpha = \frac{1}{4}\varepsilon_2 c_{\mathcal{Z}}$ on $\{\chi = 0\}$ near p_0, where $T_{p_0}\{\chi = 0\} = N_{p_0}\mathcal{M}$ ensures compact local storage sublevel components, and Corollary 3 gives $w(t) \to p_0$.

6.4 Larger Circle: Nonzero Quadratic Residuals

For $\mathcal{M} = \mathcal{C}_\varrho$ with $\varrho > \sqrt{2m}$, the circle crosses $\mathcal{Z}$ transversally within $\mathcal{P}$, so $T_p\mathcal{M} \not\subseteq T_p\mathcal{Z}$ at each feasible point

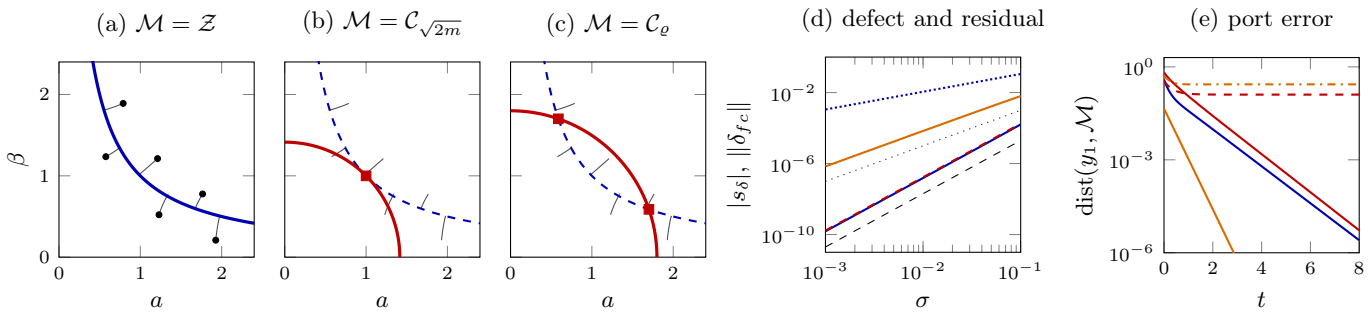


Figure 7: Heterogeneous data with $d = 2$, $\theta^\star = (0.8, 0.6)$, $\xi_1 = (1, 0)$, $\xi_2 = (0.5, 1)$, $\xi_3 = (1, -0.5)$, $\varrho = 1.8$. (a)–(c) The same trajectories (gray) projected onto $\mathcal{P}$, with $\mathcal{Z}$ (blue), the circle targets (red), and their feasible points (squares). (d) Defect $\|\delta_{fc}\|$ for $\mathcal{M} = \mathcal{Z}$ (dotted) and residual $|s_\delta|$ for the three targets (blue, red dashed, orange) along perturbations of size σ from a feasible point; reference slopes σ^2 and σ^3. (e) Server port error for $\mathcal{Z}$ (blue), the minimum circle with $\chi_0 = 0$ (red) and $\chi_0 \neq 0$ (red dashed), and the larger circle on matched (orange) and unmatched (orange dash-dotted) imbalance levels. The matched trajectory starts at $b(0) = (a_p^{-1} + 0.03)\hat{\theta} + 0.04(-0.6, 0.8)$ on $\{\chi = \chi(p)\}$; integration by fourth-order Runge–Kutta with step 0.002.

$p \in \mathcal{Z}_{\mathrm{cl}} = \mathcal{M} \cap \mathcal{Z}$. Proposition 8(ii) therefore gives $R_p \neq 0$ (Fig. 7(d)), and the quadratic residual must be compared with the available dissipation. To apply Theorem 3 via (51) without the conservatism of Q, we retain the unrelaxed client supply $W_0(w) := \sum_i \langle e_{\mathcal{Z}_i}(w), \nabla \ell_i(w) \rangle \geq 0$, for which (54) gives $W_0(w) - s_\delta(w) = \langle e_{\mathcal{M}}(w), \nabla L(w) \rangle = -\dot{V}_{\mathrm{cl}}(w)$. Since $e_{\mathcal{M}}(p) = \nabla L(p) = 0$ and $\mathrm{d}_p e_{\mathcal{M}} = \mathrm{P}_{N_p \mathcal{M}}$, the net quadratic form (47) at p is

$$H_p = \tfrac{1}{2} \nabla^2 W_0(p) - R_p = \mathrm{sym}\big(\mathrm{P}_{N_p \mathcal{M}} \nabla^2 L(p)\big). \tag{56}$$

For the data in Fig. 7, $X = \mathrm{diag}(9/4, 5/4)$, $m = 1$, and $\varrho = 9/5$. At the feasible equilibrium $p = (a_p, \theta^\star / a_p)$ with $a_p^2 = \frac{81 + \sqrt{4061}}{50}$, $a_p > 0$, and $\chi(p) = \frac{\sqrt{4061}}{25}$, an orthonormal basis of $T_p \mathcal{H}$ for the matched restriction $\mathcal{H} = \{\chi = \chi(p)\}$ is

$$E = \begin{bmatrix} 1/(a_p \varrho) & 0 \\ 4a_p/(5\varrho) & -3/5 \\ 3a_p/(5\varrho) & 4/5 \end{bmatrix}.$$

Direct calculation from (56) yields

$$E^\top \mathrm{P}_{N_p \mathcal{M}} E = \mathrm{diag}(2500/6561, 1) \succ 0,$$
$$E^\top H_p E = \begin{bmatrix} \frac{7}{3} & -\frac{18122 a_p}{30375} \\ -\frac{18122 a_p}{30375} & \frac{161 a_p^2}{100} \end{bmatrix} \succ 0,$$

while $(E^\top R_p E)_{11} = 28427/7500 > 0$, so the quadratic residual remains nonzero on $T_p \mathcal{H}$. Since $\nabla^2 V_{\mathrm{cl}}(p) = \mathrm{P}_{N_p \mathcal{M}}$, the two positive-definite matrices verify (50) and (48) on $\mathcal{H}$ with $\mathcal{K} = \{p\}$. Theorem 3(iii) therefore yields local exponential convergence to p on $\mathcal{H}$ (Fig. 7(e), orange solid). If $\chi_0 \notin \{\pm\sqrt{\varrho^4 - 4m^2}\}$, then $\{\chi = \chi_0\} \cap \mathcal{Z}_{\mathrm{cl}} = \emptyset$: under the local positive-branch hypotheses of Subsection 6.2, $w(t) \to p_{\chi_0} \notin \mathcal{C}_\varrho$ and $\mathrm{dist}(y_1(t), \mathcal{C}_\varrho) \to \big|\|p_{\chi_0}\| - \varrho\big| > 0$ (Fig. 7(e), orange dash-dotted).

6.5 Discussion

Since the clients are neither passive nor monotone, passivity- and monotonicity-based analyses of networked optimization [14, 16] do not apply, and convergence analyses of overparameterized learning [5, 34] certify convergence to $\mathcal{Z}$ only. Transversal passivity, the hierarchy of Table 1, and conserved-quantity confinement certify the same closed loop relative to other targets, identify through the imbalance which targets are reachable, and explain through the margin obstruction when certification fails. Heterogeneity enters only through the tangency condition of Proposition 8(ii), which is checked from the targets alone.

7 Conclusion

We introduced transversal passivity, a storage inequality in the transverse port errors with respect to prescribed manifolds, and characterized it by a necessary condition in linear matrix inequality form, a constructive sufficient condition, and its relation to standard, shifted, and incremental passivity. For feedback through negative-adjoint linear channels, the non-cancelling supply is governed by the compatibility defects of the channels, which yields a hierarchy from exact compatibility through quadratic cancellation to nonzero quadratic residuals. Together with compact confinement and a dissipation margin, each level of this hierarchy gives port convergence. A federated learning example with heterogeneous data showed that convergence can be certified although no server target is compatible, and that a conserved quantity of the learning dynamics determines which targets are reachable.

Future work includes nonlinear interconnections, the synthesis of transversally passive systems, the role of dynamical and interconnection symmetries, and networked

systems on directed graphs, whose channels need not form a negative-adjoint pair [18,19], with decentralized learning as a particular application.

Acknowledgment

This work was supported by the Israel Science Foundation grant no. 453/24. The authors used Claude (Anthropic) and ChatGPT (OpenAI) to assist with copy-editing, with generating the code for the figures, and with the numerical studies in Section 6. All AI-assisted content was reviewed and verified by the authors, who take full responsibility for this article.

References

[1] W. Yao, H. G. de Marina, B. Lin, and M. Cao, "Singularity-free guiding vector field for robot navigation," *IEEE Transactions on Robotics*, vol. 37, no. 4, pp. 1206–1221, 2021.

[2] T. A. Johansen and T. I. Fossen, "Control allocation—a survey," *Automatica*, vol. 49, no. 5, pp. 1087–1103, 2013.

[3] F. Dörfler and B. Francis, "Geometric analysis of the formation problem for autonomous robots," *IEEE Transactions on Automatic Control*, vol. 55, no. 10, pp. 2379–2384, 2010.

[4] A. Sarlette and R. Sepulchre, "Consensus optimization on manifolds," *SIAM Journal on Control and Optimization*, vol. 48, no. 1, pp. 56–76, 2009.

[5] C. Liu, L. Zhu, and M. Belkin, "Loss landscapes and optimization in over-parameterized non-linear systems and neural networks," *Applied and Computational Harmonic Analysis*, vol. 59, pp. 85–116, 2022.

[6] H. K. Khalil, *Nonlinear Systems, 3rd ed.* Prentice Hall, 2002.

[7] J. C. Willems, "Dissipative dynamical systems Part I: General theory," *Archive for Rational Mechanics and Analysis*, vol. 45, no. 5, pp. 321–351, 1972.

[8] I. R. Manchester and J.-J. E. Slotine, "Transverse contraction criteria for existence, stability, and robustness of a limit cycle," *Systems & Control Letters*, vol. 63, pp. 32–38, 2014.

[9] V. Andrieu, B. Jayawardhana, and L. Praly, "Transverse exponential stability and applications," *IEEE Transactions on Automatic Control*, vol. 61, no. 11, pp. 3396–3411, 2016.

[10] C. Nielsen and M. Maggiore, "On local transverse feedback linearization," *SIAM Journal on Control and Optimization*, vol. 47, no. 5, pp. 2227–2250, 2008.

[11] A. Astolfi and R. Ortega, "Immersion and invariance: A new tool for stabilization and adaptive control of nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 48, no. 4, pp. 590–606, 2003.

[12] A. Isidori and C. I. Byrnes, "Output regulation of nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 35, no. 2, pp. 131–140, 1990.

[13] J. M. Montenbruck, M. Arcak, and F. Allgöwer, "An input–output framework for submanifold stabilization," *IEEE Transactions on Automatic Control*, vol. 62, no. 10, pp. 5170–5184, 2017.

[14] M. Arcak, "Passivity as a design tool for group coordination," *IEEE Transactions on Automatic Control*, vol. 52, no. 8, pp. 1380–1390, 2007.

[15] A. J. van der Schaft and B. M. Maschke, "Port-Hamiltonian systems on graphs," *SIAM Journal on Control and Optimization*, vol. 51, no. 2, pp. 906–937, 2013.

[16] M. Bürger, D. Zelazo, and F. Allgöwer, "Duality and network theory in passivity-based cooperative control," *Automatica*, vol. 50, no. 8, pp. 2051–2061, 2014.

[17] N. Chopra and M. W. Spong, "Passivity-based control of multi-agent systems," in *Advances in Robot Control: From Everyday Physics to Human-Like Movements*, S. Kawamura and M. Svinin, Eds. Berlin, Germany: Springer, 2006, pp. 107–134.

[18] F.-Y. Yue and D. Zelazo, "A passivity analysis for nonlinear consensus on digraphs," in *2025 IEEE 64th Conference on Decision and Control (CDC)*, 2025, pp. 6216–6221.

[19] ——, "A passivity analysis for nonlinear consensus on balanced digraphs," *European Journal of Control*, vol. 86, no. Part A, p. 101368, 2025.

[20] M. I. El-Hawwary and M. Maggiore, "Reduction principles and the stabilization of closed sets for passive systems," *IEEE Transactions on Automatic Control*, vol. 55, no. 4, pp. 982–987, 2010.

[21] V. Guillemin and A. Pollack, *Differential topology*. American Mathematical Society, 2010, vol. 370.

[22] J. M. Lee, *Introduction to Smooth Manifolds*. Springer, 2003.

[23] G. Leobacher and A. Steinicke, "Existence, uniqueness and regularity of the projection onto differentiable manifolds," *Annals of Global Analysis and Geometry*, vol. 60, no. 3, pp. 559–587, 2021.

[24] J. C. Willems, "Dissipative dynamical systems Part II: Linear systems with quadratic supply rates," *Archive for Rational Mechanics and Analysis*, vol. 45, no. 5, pp. 352–393, 1972.

[25] B. Brogliato, R. Lozano, B. Maschke, and O. Egeland, *Dissipative Systems Analysis and Control: Theory and Applications*, 3rd ed. Cham, Switzerland: Springer, 2020.

[26] N. Monshizadeh, P. Monshizadeh, R. Ortega, and A. van der Schaft, "Conditions on shifted passivity of port-Hamiltonian systems," *Systems & Control Letters*, vol. 123, pp. 55–61, 2019.

[27] G. H. Hines, M. Arcak, and A. K. Packard, "Equilibrium-independent passivity: A new definition and numerical certification," *Automatica*, vol. 47, no. 9, pp. 1949–1956, 2011.

[28] A. J. van der Schaft, "On differential passivity," *IFAC Proceedings Volumes*, vol. 46, no. 23, pp. 21–25, 2013.

[29] B. Maschke, R. Ortega, and A. J. van der Schaft, "Energy-based Lyapunov functions for forced Hamiltonian systems with dissipation," *IEEE Transactions on Automatic Control*, vol. 45, no. 8, pp. 1498–1502, 2000.

[30] H. H. Bauschke and P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, 2nd ed. Cham, Switzerland: Springer, 2017.

[31] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. Agüera y Arcas, "Communication-efficient learning of deep networks from decentralized data," in *Proc. 20th Int. Conf. Artificial Intelligence and Statistics (AISTATS)*, ser. PMLR, vol. 54, 2017, pp. 1273–1282.

[32] P. Kairouz, H. B. McMahan *et al.*, "Advances and open problems in federated learning," *Foundations and Trends in Machine Learning*, vol. 14, no. 1–2, pp. 1–210, 2021.

[33] S. S. Du, W. Hu, and J. D. Lee, "Algorithmic regularization in learning deep homogeneous models: Layers are automatically balanced," in *Advances in Neural Information Processing Systems*, vol. 31, 2018.

[34] C. Liu, D. Drusvyatskiy, M. Belkin, D. Davis, and Y.-A. Ma, "Aiming towards the minimizers: Fast convergence of SGD for overparametrized problems," in *Advances in Neural Information Processing Systems*, vol. 36, 2023.