P\textsuperscript{TH} MOMENT NOISE-TO-STATE STABILITY OF STOCHASTIC DIFFERENTIAL EQUATIONS WITH PERSISTENT NOISE\textsuperscript{*}

DAVID MATEOS-N\textsuperscript{Ú}NEZ AND JORGE CORT\textsuperscript{É}S\textsuperscript{†}

Abstract. This paper studies the stability properties of stochastic differential equations subject to persistent noise (including the case of additive noise), which is noise that is present even at the equilibria of the underlying differential equation and does not decay with time. The class of systems we consider exhibit disturbance attenuation outside a closed, not necessarily bounded, set. We identify conditions, based on the existence of Lyapunov functions, to establish the noise-to-state stability in probability and in \( p \)-th moment of the system with respect to a closed set. As part of our analysis, we study the concept of two functions being proper with respect to each other formalized via pair of inequalities with comparison functions. We show that such inequalities define several equivalence relations for increasingly strong refinements on the comparison functions. We also provide a complete characterization of the properties that a pair of functions must satisfy to belong to the same equivalence class. This characterization allows us to provide checkable conditions to determine whether a function satisfies the requirements to be a strong NSS-Lyapunov function in probability or a \( p \)-th moment NSS-Lyapunov function. Several examples illustrate our results.

1. Introduction. Stochastic differential equations (SDEs) go beyond ordinary differential equations (ODEs) to deal with systems subject to stochastic perturbations of a particular type, known as white noise. Applications are numerous and include option pricing in the stock market, networked systems with noisy communication channels, and, in general, scenarios whose complexity cannot be captured by deterministic models. In this paper, we study SDEs subject to persistent noise (including the case of additive noise), i.e., systems for which the noise is present even at the equilibria of the underlying ODE and does not decay with time. Such scenarios arise, for instance, in control-affine systems when the input is corrupted by persistent noise. For these systems, the presence of persistent noise makes it impossible to establish in general a stochastic notion of asymptotic stability for the (possibly unbounded) set of equilibria of the underlying ODE. Our aim here is to develop notions and tools to study the stability properties of these systems and provide probabilistic guarantees on the size of the state of the system.

Literature review: In general, it is not possible to obtain explicit descriptions of the solutions of SDEs. Fortunately, the Lyapunov techniques used to study the qualitative behavior of ODEs [6, 10] can be adapted to study the stability properties of SDEs as well [7, 13, 27]. Depending on the notion of stochastic convergence used, there are several types of stability results in SDEs. These include stochastic stability (or stability in probability), stochastic asymptotic stability, almost sure exponential stability, and \( p \)-th moment asymptotic stability, see e.g., [13, 14, 26, 27]. However, these notions are not appropriate in the presence of persistent noise because they require the effect of the noise on the set of equilibria to either vanish or decay with time. To deal with persistent noise, as well as other system properties like delays, a concept of ultimate boundedness is required that generalizes the notion of convergence. As an example, for stochastic delay differential equations, [28] considers a notion of ultimate...
bound in $p$th moment [21] and employs Lyapunov techniques to establish it. More generally, for mean-square random dynamical systems, the concept of forward attractor [9] describes a notion of convergence to a dynamic neighborhood and employs contraction analysis to establish it. Similar notions of ultimate boundedness for the state of a system, now in terms of the size of the disturbance, are also used for differential equations, and many of these notions are inspired by dissipativity properties of the system that are captured via dissipation inequalities of a suitable Lyapunov function: such inequalities encode the fact that the Lyapunov function decreases along the trajectories of the system as long as the state is “big enough” with regards to the disturbance. As an example, the concept of input-to-state stability (ISS) goes hand in hand with the concept of ISS-Lyapunov function, since the existence of the second implies the former (and, in many cases, a converse result is also true [24]). Along these lines, the notion of practical stochastic input-to-state stability (SISS) in [12, 29] generalizes the concept of ISS to SDEs where the disturbance or input affects both the deterministic term of the dynamics and the diffusion term modeling the role of the noise. Under this notion, the state bound is guaranteed in probability, and also depends, as in the case of ISS, on a decaying effect of the initial condition plus an increasing function of the sum of the size of the input and a positive constant related to the persistent noise. For systems where the input modulates the covariance of the noise, SISS corresponds to noise-to-state-stability (NSS) [3], which guarantees, in probability, an ultimate bound for the state that depends on the magnitude of the noise covariance. That is, the noise in this case plays the main role, since the covariance can be modulated arbitrarily and can be unknown. This is the appropriate notion of stability for the class of SDEs with persistent noise considered in this paper, which are nonlinear systems affine in the input, where the input corresponds to white noise with locally bounded covariance. Such systems cannot be studied under the ISS umbrella, because the stochastic integral against Brownian motion has infinite variation, whereas the integral of a legitimate input for ISS must have finite variation.

Statement of contributions: The contributions of this paper are twofold. Our first contribution concerns the noise-to-state stability of systems described by SDEs with persistent noise. We generalize the notion of noise-dissipative Lyapunov function, which is a positive semidefinite function that satisfies a dissipation inequality that can be nonexponential (by this we mean that the inequality admits a convex $\mathcal{K}_\infty$ gain instead of the linear gain characteristic of exponential dissipativity). We also introduce the concept of $p$thNSS-Lyapunov function with respect to a closed set, which is a noise-dissipative Lyapunov function that in addition is proper with respect to the set with a convex lower-bound gain function. Using this framework, we show that noise-dissipative Lyapunov functions have NSS dynamics and we characterize the overshoot gain. More importantly, we show that the existence of a $p$thNSS-Lyapunov function with respect to a closed set implies that the system is NSS in $p$th moment with respect to the set. Our second contribution is driven by the aim of providing alternative, structured ways to check the hypotheses of the above results. We introduce the notion of two functions being proper with respect to each other as a generalization of the notion of properness with respect to a set. We then develop a methodology to verify whether two functions are proper with respect to each other by analyzing the associated pair of inequalities with increasingly strong refinements that involve the classes $\mathcal{K}, \mathcal{K}_\infty$, and $\mathcal{K}_\infty$ plus a convexity property. We show that these refinements define equivalence relations between pairs of functions, thereby producing nested partitions on the space of functions. This provides a useful way to
deal with these inequalities because the construction of the gains is explicit when the transitivity property is exploited. This formalism motivates our characterization of positive semidefinite functions that are proper, in the various refinements, with respect to the Euclidean distance to their nullset. This characterization is technically challenging because we allow the set to be noncompact, and thus the pre-comparison functions can be discontinuous. We devote special attention to the case when the set is a subspace and examine the connection with seminorms. Finally, we show how this framework allows us to develop an alternative formulation of our stability results.

Organization: The paper is organized as follows. Section 2 introduces preliminaries on seminorms, comparison functions, and SDEs. Section 3 presents the NSS stability results and Section 4 develops the methodology to help verify their hypotheses. Finally, Section 5 discusses our conclusions and ideas for future work.

2. Preliminary notions. This section reviews some notions on comparison functions and stochastic differential equations that are used throughout the paper.

2.1. Notational conventions. Let \( \mathbb{R} \) and \( \mathbb{R}_{\geq 0} \) be the sets of real and nonnegative real numbers, respectively. We denote by \( \mathbb{R}^n \) the \( n \)-dimensional Euclidean space. A subspace \( \mathcal{U} \subseteq \mathbb{R}^n \) is a subset which is also a vector space. Given a matrix \( A \in \mathbb{R}^{m \times n} \), its nullspace \( \mathcal{N}(A) \equiv \{ x \in \mathbb{R}^n : Ax = 0 \} \) is a subspace. Given \( \mathcal{D} \subseteq \mathbb{R}^n \), we denote by \( C(\mathcal{D}; \mathbb{R}_{\geq 0}) \) and \( C^2(\mathcal{D}; \mathbb{R}_{\geq 0}) \) the set of positive semidefinite functions defined on \( \mathcal{D} \) that are continuous and continuously twice differentiable (if \( \mathcal{D} \) is open), respectively. Given \( V \in C^2(\mathbb{R}^n; \mathbb{R}_{\geq 0}) \), we denote its gradient by \( \nabla V \) and its Hessian by \( \nabla^2 V \). A seminorm is a function \( S : \mathbb{R}^n \to \mathbb{R} \) that is positively homogeneous, i.e., \( S(\lambda x) = |\lambda|S(x) \) for any \( \lambda \in \mathbb{R} \), and satisfies the triangular inequality, i.e., \( S(x + y) \leq S(x) + S(y) \) for any \( x, y \in \mathbb{R}^n \). From these properties it can be deduced that \( S \in C(\mathbb{R}^n; \mathbb{R}_{\geq 0}) \) and its nullset is always a subspace. If, moreover, the function \( S \) is positive definite, i.e., \( S(x) = 0 \) implies \( x = 0 \), then \( S \) is a norm. The Euclidean norm of \( x \in \mathbb{R}^n \) is denoted by \( \|x\|_2 \), and the Frobenius norm of the matrix \( A \in \mathbb{R}^{m \times n} \) is \( \|A\|_F \equiv \sqrt{\text{trace}(A^T A)} = \sqrt{\text{trace}(AA^T)} \). For any matrix \( A \in \mathbb{R}^{m \times n} \), the function \( \|x\|_A \equiv \|Ax\|_2 \) is a seminorm and can be viewed as a distance to \( \mathcal{N}(A) \).

For a symmetric positive semidefinite real matrix \( A \in \mathbb{R}^{n \times n} \), we order its eigenvalues as \( \lambda_1(A) \geq \cdots \geq \lambda_n(A) \geq 0 \), so if the dimension of \( \mathcal{N}(A) \) verifies \( \dim(\mathcal{N}(A)) = k \leq n \), then \( \lambda_{n-k}(A) \) is the minimum nonzero eigenvalue of \( A \). The Euclidean distance from \( x \) to a set \( \mathcal{U} \subseteq \mathbb{R}^n \) is defined by \( \|x\|_\mathcal{U} \equiv \inf \{ \|x - u\|_2 : u \in \mathcal{U} \} \).

The function \( \|\cdot\|_\mathcal{U} \) is continuous when \( \mathcal{U} \) is closed. Given \( f, g : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0} \), we say that \( f(s) \) is in \( \mathcal{O}(g(s)) \) as \( s \to \infty \) if there exist constants \( \kappa, s_0 > 0 \) such that \( f(s) < \kappa g(s) \) for all \( s > s_0 \). Finally, we denote by ess sup the essential supremum operator.

2.2. Comparison, convex, and concave functions. Here we introduce some classes of comparison functions following [6] that are useful in our technical treatment.

A continuous function \( \alpha : [0, b) \to \mathbb{R}_{\geq 0} \), for \( b > 0 \) or \( b = \infty \), is class \( \mathcal{K} \) if \( \alpha(0) = 0 \) and is strictly increasing. A function \( \alpha : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0} \) is class \( \mathcal{K}_\infty \) if \( \alpha \in \mathcal{K} \) and is unbounded. A continuous function \( \mu : \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0} \) is class \( \mathcal{KL} \) if, for each fixed \( s \geq 0 \), the function \( r \mapsto \mu(r, s) \) is class \( \mathcal{K} \), and, for each fixed \( r \geq 0 \), the function \( s \mapsto \mu(r, s) \) is decreasing and \( \lim_{s \to \infty} \mu(r, s) = 0 \). If \( \alpha_1, \alpha_2 \) are class \( \mathcal{K} \) and the domain of \( \alpha_1 \) contains the range of \( \alpha_2 \), then their composition \( \alpha_1 \circ \alpha_2 \) is class \( \mathcal{K} \) too. If \( \alpha_3, \alpha_4 \) are class \( \mathcal{K}_\infty \), then both the inverse function \( \alpha_3^{-1} \) and their composition \( \alpha_3 \circ \alpha_4 \) are class \( \mathcal{K}_\infty \). In our technical treatment, it is sometimes convenient to require comparison functions to be class \( \mathcal{K}_\infty \).
functions to satisfy additional convexity properties. A real-valued function $f$ defined in a convex set $X$ in a vector space is convex if $f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y)$ for each $x, y \in X$ and any $\lambda \in [0, 1]$, and is concave if $-f$ is convex. By [2, Ex. 3.3], if $f : [a, b] \rightarrow [f(a), f(b)]$ is a strictly increasing convex (respectively, concave) function, then the inverse function $f^{-1} : [f(a), f(b)] \rightarrow [a, b]$ is strictly increasing and concave (respectively, convex). Also, following [2, Section 3], if $f, g : \mathbb{R} \rightarrow \mathbb{R}$ are convex (respectively, concave) and $f$ is nondecreasing, then the composition $f \circ g$ is also convex (respectively, concave).

2.3. Brownian motion. We review some basic notions on probability and introduce the notion of Brownian motion following [14]. Throughout the paper, we assume that $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P})$ is a complete probability space, where $\mathbb{P}$ is a probability measure defined on the $\sigma$-algebra $\mathcal{F}$, which contains all the subsets of $\Omega$ of probability 0. The filtration $\{\mathcal{F}_t\}_{t \geq 0}$ is a family of sub-$\sigma$-algebras of $\mathcal{F}$ satisfying $\mathcal{F}_t \subseteq \mathcal{F}_s \subseteq \mathcal{F}$ for any $0 \leq t < s < \infty$; we assume it is right continuous, i.e., $\mathcal{F}_t = \bigcap_{s > t} \mathcal{F}_s$ for any $t \geq 0$, and $\mathcal{F}_0$ contains all the subsets of $\Omega$ of probability 0. The Borel $\sigma$-algebra in $\mathbb{R}^n$, denoted by $\mathcal{B}^n$, or in $[t_0, \infty)$, denoted by $\mathcal{B}([t_0, \infty))$, are the smallest $\sigma$-algebras that contain all the open sets in $\mathbb{R}^n$ or $[t_0, \infty)$, respectively. A function $X : \Omega \rightarrow \mathbb{R}^n$ is $\mathcal{F}$-measurable if the set $\{\omega \in \Omega : X(\omega) \in A\}$ belongs to $\mathcal{F}$ for any $A \in \mathcal{B}^n$. We call such function a ($\mathcal{F}$-measurable) $\mathbb{R}^n$-valued random variable. If $X$ is a real-valued random variable that is integrable with respect to $\mathbb{P}$, its expectation is $\mathbb{E}[X] = \int_{\Omega} X(\omega) d\mathbb{P}(\omega)$.

A function $f : \Omega \times [t_0, \infty) \rightarrow \mathbb{R}^n$ is $\mathcal{F} \times \mathcal{B}$-measurable (or just measurable) if the set $\{(\omega, t) \in \Omega \times [t_0, \infty) : f(\omega, t) \in A\}$ belongs to $\mathcal{F} \times \mathcal{B}([t_0, \infty))$ for any $A \in \mathcal{B}^n$. We call such function an $\{\mathcal{F}_t\}$-adapted process if $f(\cdot, t) : \Omega \rightarrow \mathbb{R}^n$ is $\mathcal{F}_t$-measurable for every $t \geq t_0$. At times, we omit the dependence on “$\omega$”, in the sense that we refer to the indexed family of random variables, and refer to the random process $f = \{f(t)\}_{t \geq t_0}$.

We define $\mathcal{L}^1([t_0, \infty); \mathbb{R}^n)$ as the set of all $\mathbb{R}^n$-valued measurable $\{\mathcal{F}_t\}$-adapted processes $f$ such that $\mathbb{P}(\{\omega \in \Omega : \int_{t_0}^T \|f(\omega, s)\|^2 ds < \infty\}) = 1$ for every $T > t_0$.

Similarly, $\mathcal{L}^2([t_0, \infty); \mathbb{R}^{n \times m})$ denotes the set of all $\mathbb{R}^{n \times m}$-matrix-valued measurable $\{\mathcal{F}_t\}$-adapted processes $G$ such that $\mathbb{P}(\{\omega \in \Omega : \int_{t_0}^T \|G(\omega, s)\|^2 ds < \infty\}) = 1$ for every $T > t_0$.

A one-dimensional Brownian motion $B : \Omega \times [t_0, \infty) \rightarrow \mathbb{R}$ defined in the probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P})$ is an $\{\mathcal{F}_t\}$-adapted process such that

- $\mathbb{P}(\{\omega \in \Omega : B(\omega, t_0) = 0\}) = 1$;
- the mapping $B(\cdot, \cdot) : [t_0, \infty) \rightarrow \mathbb{R}$, called sample path, is continuous also with probability 1;
- the increment $B(\cdot, t) - B(\cdot, s) : \Omega \rightarrow \mathbb{R}$ is independent of $\mathcal{F}_s$ for $t_0 \leq s < t < \infty$ (i.e., if $S_b \triangleq \{\omega \in \Omega : B(\omega, t) - B(\omega, s) \in (-\infty, b)\}$, for $b \in \mathbb{R}$, then $\mathbb{P}(A \cap S_b) = \mathbb{P}(A)\mathbb{P}(S_b)$ for all $A \in \mathcal{F}_s$ and all $b \in \mathbb{R}$). In addition, this increment is normally distributed with zero mean and variance $t-s$.

An $m$-dimensional Brownian motion $B : \Omega \times [t_0, \infty) \rightarrow \mathbb{R}^m$ is given by $B(\omega, t) = [B_1(\omega, t), \ldots, B_m(\omega, t)]^T$, where each $B_i$ is a one-dimensional Brownian motion and, for each $t \geq t_0$, the random variables $B_1(t), \ldots, B_m(t)$ are independent.

2.4. Stochastic differential equations. Here we review some basic notions on stochastic differential equations (SDEs) following [14]; other useful references are [7,
existence and uniqueness of solutions of (2.1) for each initial condition. According to [14, Th. 3.6, p. 58], Assumption 2.1 is sufficient to guarantee global existence and uniqueness of solutions of (2.1). We conclude this section by presenting a useful operator in the stability analysis of SDEs.

Consider the n-dimensional SDE

\[ dx(\omega, t) = f(x(\omega, t), t)dt + G(x(\omega, t), t)\Sigma(t)dB(\omega, t), \]

where \( x(\omega, t) \in \mathbb{R}^n \) is a realization at time \( t \) of the random variable \( x(., t) : \Omega \to \mathbb{R}^n \), for \( t \in [t_0, \infty) \). The initial condition is given by \( x(\omega, t_0) = x_0 \) with probability 1 for some \( x_0 \in \mathbb{R}^n \). The functions \( f : \mathbb{R}^n \times [t_0, \infty) \to \mathbb{R}^n \), \( G : \mathbb{R}^n \times [t_0, \infty) \to \mathbb{R}^{n \times q} \), and \( \Sigma : [t_0, \infty) \to \mathbb{R}^{q \times m} \) are measurable. The functions \( f \) and \( G \) are regarded as a model for the architecture of the system and, instead, \( \Sigma \) is part of the model for the stochastic disturbance; at any given time \( \Sigma \) determines a linear transformation of the \( m \)-dimensional Brownian motion \( \{B(t)\}_{t \geq t_0} \), so that at time \( t \geq t_0 \) the input to the system is the process \( \{\Sigma(t)B(t)\}_{t \geq t_0} \), with covariance \( \int_{t_0}^{t} \Sigma(t)\Sigma(t)^Tds \). The distinction between the roles of \( G \) and \( \Sigma \) is irrelevant for the SDE; both together determine the effect of the Brownian motion. The integral form of (2.1) is given by

\[
\begin{align*}
  x(\omega, t) &= x_0 + \int_{t_0}^{t} f(x(\omega, s), s)ds + \int_{t_0}^{t} G(x(\omega, s), s)\Sigma(s)dB(\omega, s),
\end{align*}
\]

where the second integral is a stochastic integral [14, p. 18]. A \( \mathbb{R}^n \)-valued random process \( \{x(t)\}_{t \geq t_0} \) is a solution of (2.1) with initial value \( x_0 \) if

(i) is continuous with probability 1, \( \{\mathcal{F}_t\} \)-adapted, and satisfies \( x(\omega, t_0) = x_0 \) with probability 1,

(ii) the processes \( \{f(x(t), t)\}_{t \geq t_0} \) and \( \{G(x(t), t)\}_{t \geq t_0} \) belong to \( \mathcal{L}^1([t_0, \infty); \mathbb{R}^n) \) and \( \mathcal{L}^2([t_0, \infty); \mathbb{R}^{n \times m}) \) respectively, and

(iii) equation (2.1) holds for every \( t \geq t_0 \) with probability 1.

A solution \( \{x(t)\}_{t \geq t_0} \) of (2.1) is unique if any other solution \( \{\bar{x}(t)\}_{t \geq t_0} \) with \( \bar{x}(t_0) = x_0 \) differs from it only in a set of probability 0, that is, \( P(\{x(t) = \bar{x}(t) : \forall t \geq t_0\}) = 1 \).

We make the following assumptions on the objects defining (2.1) to guarantee existence and uniqueness of solutions.

**Assumption 2.1.** We assume \( \Sigma \) is essentially locally bounded. Furthermore, for any \( T > t_0 \) and \( n \geq 1 \), we assume there exists \( K_{T,n} > 0 \) such that, for almost every \( t \in [t_0, T] \) and all \( x, y \in \mathbb{R}^n \) with \( \max \{\|x\|_2, \|y\|_2\} \leq n \),

\[
\max\{\|f(x, t) - f(y, t)\|_2^2, \|G(x, t) - G(y, t)\|_2^2\} \leq K_{T,n}\|x - y\|_2^2.
\]

Finally, we assume that for any \( T > t_0 \), there exists \( K_T > 0 \) such that, for almost every \( t \in [t_0, T] \) and all \( x \in \mathbb{R}^n \),

\[
\|\int_{t_0}^{t} f(x(s), s)ds\|_2^2 \leq K_T(1 + \|x\|_2^2).
\]

According to [14, Th. 3.6, p. 58], Assumption 2.1 is sufficient to guarantee global existence and uniqueness of solutions of (2.1) for each initial condition \( x_0 \in \mathbb{R}^n \).

We conclude this section by presenting a useful operator in the stability analysis of SDEs. Given a function \( V \in \mathcal{C}^2(\mathbb{R}^n; \mathbb{R}_+^n) \), we define the generator of (2.1) acting on the function \( V \) as the mapping \( \mathcal{L}[V] : \mathbb{R}^n \times [t_0, \infty) \to \mathbb{R} \) given by

\[
\mathcal{L}[V](x, t) = \nabla V(x)^T f(x, t) + \frac{1}{2} \text{tr} \left( \Sigma(t)^T G(x, t)^T \nabla^2 V(x) G(x, t) \Sigma(t) \right).
\]

It can be shown that \( \mathcal{L}[V](x, t) \) gives the expected rate of change of \( V \) along a solution of (2.1) that passes through the point \( x \) at time \( t \), so it is a generalization of the Lie bracket.
derivative. According to [14, Th. 6.4, p. 36], if we evaluate \( V \) along the solution \( \{x(t)\}_{t \geq t_0} \) of (2.1), then the process \( \{V(x(t))\}_{t \geq t_0} \) satisfies the new SDE

\[
(2.3) \quad V(x(t)) = V(x_0) + \int_{t_0}^{t} \mathcal{L}[V](x(s), s)ds + \int_{t_0}^{t} \nabla V(x(s))^T G(x(s), s)\Sigma(s)dB(s).
\]

Equation (2.3) is known as Itô’s formula and corresponds to the stochastic version of the chain rule.

3. Noise-to-state stability via noise-dissipative Lyapunov functions. In this section, we study the stability of stochastic differential equations subject to persistent noise. Our first step is the introduction of a novel notion of stability. This captures the behavior of the \( p \)th moment of the distance (of the state) to a given closed set, as a function of two objects: the initial condition and the maximum size of the covariance. After this, our next step is to derive several Lyapunov-type stability results that help determine whether a stochastic differential equation enjoys these stability properties. The following definition generalizes the concept of noise-to-state stability given in [3].

**Definition 3.1.** (Noise-to-state stability with respect to a set). The system (2.1) is noise-to-state stable (NSS) in probability with respect to the set \( U \subseteq \mathbb{R}^n \) if for any \( \epsilon > 0 \) there exist \( \mu \in \mathcal{KL} \) and \( \theta \in \mathcal{K} \) (that might depend on \( \epsilon \)), such that

\[
(3.1) \quad \mathbb{P} \left\{ |x(t)|^p > \mu(|x_0|_u, t - t_0) + \theta \left( \text{ess sup}_{t_0 \leq s \leq t} |\Sigma(s)|_F \right) \right\} \leq \epsilon,
\]

for all \( t \geq t_0 \) and any \( x_0 \in \mathbb{R}^n \). And the system (2.1) is \( p \)th moment noise-to-state stable (\( p \)thNSS) with respect to \( U \) if there exist \( \mu \in \mathcal{KL} \) and \( \theta \in \mathcal{K} \), such that

\[
(3.2) \quad \mathbb{E}[|x(t)|^p] \leq \mu(|x_0|_u, t - t_0) + \theta \left( \text{ess sup}_{t_0 \leq s \leq t} |\Sigma(s)|_F \right),
\]

for all \( t \geq t_0 \) and any \( x_0 \in \mathbb{R}^n \). The gain functions \( \mu \) and \( \theta \) are the overshoot gain and the noise gain, respectively.

The quantity \( |\Sigma(t)|_F = \sqrt{\text{trace}(\Sigma(t)\Sigma(t)^T)} \) is a measure of the size of the noise because it is related to the infinitesimal covariance \( \Sigma(t)\Sigma(t)^T \). The choice of the \( p \)th power is irrelevant in the statement in probability since one could take any \( \mathcal{K}_\infty \) function evaluated at \( |x(t)|_u \). However, this would make a difference in the statement in expectation. (Also, we use the same power for convenience.) When the set \( U \) is a subspace, we can substitute \( |\cdot|_u \) by \( \|\cdot\|_A \), for some matrix \( A \in \mathbb{R}^{m \times n} \) with \( N(A) = U \). In such a case, the definition above does not depend on the choice of the matrix \( A \).

**Remark 3.2.** (NSS is not a particular case of ISS). The concept of NSS is not a particular case of input-to-state stability (ISS) [23] for systems that are affine in the input, namely,

\[
\dot{y} = f(y, t) + G(y, t)u(t) \iff y(t) = y(t_0) + \int_{t_0}^{t} f(y(s), s)ds + \int_{t_0}^{t} G(y(s), s)u(s)ds,
\]

where \( u : [t_0, \infty) \to \mathbb{R}^q \) is measurable and essentially locally bounded [22, Sec. C.2]. The reason is the following: the components of the vector-valued function
\begin{align*}
\int_{t_0}^{t} G(y(s), s) u(s) \, ds \text{ are differentiable almost everywhere by the Lebesgue fundamental theorem of calculus [16, p. 289], and thus absolutely continuous [16, p. 292] and with bounded variation [16, Prop. 8.5]. On the other hand, at any time previous to } t_k(t) \triangleq \min\{t, \inf\{s \geq t_0 : \|x(s)\|_2 \geq k\}\}, \text{ the driving disturbance of (2.1) is the vector-valued function } \int_{t_0}^{t_k(t)} G(x(s), s) \Sigma(s) \, dB(s), \text{ whose } i\text{th component has quadratic variation [14, Th. 5.14, p. 25] equal to } \int_{t_0}^{t_k(t)} \sum_{j=1}^{m} \sum_{l=1}^{n} G(x(s), s)_{ij} \Sigma(s)_{lj}^2 \, ds > 0.
\end{align*}

Since a continuous process that has positive quadratic variation must have infinite variation [8, Th. 1.10], we conclude that the driving disturbance in this case is not allowed in the ISS framework.

Our first goal now is to provide tools to establish whether a stochastic differential equation enjoys the noise-to-state stability properties given in Definition 3.1. To achieve this, we look at the dissipativity properties of a special kind of energy functions along the solutions of (2.1).

**Definition 3.3. (Noise-dissipative Lyapunov function).** A function \( V \in C^2(\mathbb{R}^n; \mathbb{R}_\geq 0) \) is a noise-dissipative Lyapunov function for (2.1) if there exist \( W \in C(\mathbb{R}^n; \mathbb{R}_\geq 0), \) \( \sigma \in K, \) and concave \( \eta \in K_\infty \) such that

\begin{align*}
V(x) \leq \eta(W(x)), & \quad \text{for all } x \in \mathbb{R}^n, \text{ and the following dissipation inequality holds:} \\
\mathcal{L}[V](x, t) \leq -W(x) + \sigma(|\Sigma(t)|_F), & \quad \text{for all } (x, t) \in \mathbb{R}^n \times [t_0, \infty).
\end{align*}

**Remark 3.4. (Itô formula and exponential dissipativity).** Interestingly, the conditions (3.3) and (3.4) are equivalent to

\begin{align*}
\mathcal{L}[V](x, t) \leq -\eta^{-1}(V(x)) + \sigma(|\Sigma(t)|_F),
\end{align*}

for all \( x \in \mathbb{R}^n, \) where \( \eta^{-1} \in K_\infty \) is convex. Note that, since \( \mathcal{L}[V] \) is not the Lie derivative of \( V \) (as it contains the Hessian of \( V \)), one cannot directly deduce from (3.5) the existence of a continuously twice differentiable function \( \tilde{V} \) such that

\begin{align*}
\mathcal{L}[\tilde{V}](x, t) \leq -c\tilde{V}(x) + \tilde{\sigma}(|\Sigma(t)|_F),
\end{align*}

as instead can be done in the context of ISS, see e.g. [19].

**Example 3.5. (A noise-dissipative Lyapunov function).** Assume that \( h : \mathbb{R}^n \to \mathbb{R} \) is continuously differentiable and verifies

\begin{align*}
\gamma(||x - x'||_2^2) \leq (x - x')^T(\nabla h(x) - \nabla h(x'))
\end{align*}

for some convex function \( \gamma \in K_\infty \) for all \( x, x' \in \mathbb{R}^n. \) In particular, this implies that \( h \) is strictly convex. (Incidentally, any strongly convex function verifies (3.7) for some choice of \( \gamma \) linear and strictly increasing.) Consider now the dynamics

\begin{align*}
\text{d}x(\omega, t) = -\left(\delta Lx(\omega, t) + \nabla h(x(\omega, t))\right) \, dt + \Sigma(t) \, dB(\omega, t),
\end{align*}
for all \( t \in [t_0, \infty) \), where \( x(\omega, t_0) = x_0 \) with probability 1 for some \( x_0 \in \mathbb{R}^n \), and \( \delta > 0 \). Here, the matrix \( L \in \mathbb{R}^{n \times n} \) is symmetric and positive semidefinite, and the matrix-valued function \( \Sigma : [t_0, \infty) \to \mathbb{R}^{n \times m} \) is measurable and essentially locally bounded. This dynamics corresponds to the SDE (2.1) with \( f(x,t) \triangleq -\delta Lx - \nabla h(x) \) and \( G(x,t) \triangleq I_n \) for all \( (x,t) \in \mathbb{R}^n \times [t_0, \infty) \).

Let \( x^* \in \mathbb{R}^n \) be the unique solution of the Karush-Kuhn-Tucker [2] condition \( \delta L x^* = -\nabla h(x^*) \), corresponding to the unconstrained minimization of \( F(x) \triangleq \frac{1}{2} x^T Lx + h(x) \). Consider then the candidate Lyapunov function \( V \in \mathcal{C}^2(\mathbb{R}^n; \mathbb{R}_+ \cup \{0\}) \) given by \( V(x) \triangleq \frac{1}{2} (x-x^*)^T(x-x^*) \). Using (2.2), we obtain that, for all \( x \in \mathbb{R}^n \),

\[
\mathcal{L}[V](x,t) = -(x-x^*)^T(\delta L x + \nabla h(x)) + \frac{1}{2} \text{tr} \left( \Sigma(t)^T \Sigma(t) \right) \\
= -\delta(x-x^*)^T L (x-x^*) -(x-x^*)^T \left( \nabla h(x) - \nabla h(x^*) \right) + \frac{1}{2} |\Sigma(t)|_F^2 \\
\leq -\gamma (\|x-x^*\|_2^2) + \frac{1}{2} |\Sigma(t)|_F^2.
\]

We note that \( W \in \mathcal{C}(\mathbb{R}^n; \mathbb{R}_+ \cup \{0\}) \) defined by \( W(x) \triangleq \gamma (\|x-x^*\|_2^2) \) verifies

\[
V(x) = \frac{1}{2} \gamma^{-1}(W(x)) \quad \forall x \in \mathbb{R}^n,
\]

where \( \gamma^{-1} \) is concave and belongs to the class \( \mathcal{K}_\infty \) as explained in Section 2.2. Therefore, \( V \) is a noise-dissipative Lyapunov function for (3.8), with concave \( \eta \in \mathcal{K}_\infty \) given

\[
\eta(r) = \frac{1}{2} \gamma^{-1}(r) \quad \forall r \in \mathbb{R}_+.
\]

The next result generalizes [4, Th. 4.1] to positive semidefinite Lyapunov functions that satisfy weaker dissipativity properties (cf. (3.5)) than the typical exponential-like inequality (3.6), and characterizes the overshoot gain.

**Theorem 3.6.** (Noise-dissipative Lyapunov functions have an NSS dynamics). Under Assumption 2.1, suppose that \( V \) is a noise-dissipative Lyapunov function for (2.1). Then,

\[
\mathbb{E}[V(x(t))] \leq \mu(V(x_0), t-t_0) + \eta \left( 2 \sigma \left( \text{ess sup}_{t_0 \leq s \leq t} |\Sigma(s)|_F \right) \right),
\]

for all \( t \geq t_0 \), where the class \( \mathcal{KL} \) function \( (r,s) \mapsto \mu(r,s) \) is well defined as the solution \( y(s) \) to the initial value problem

\[
y'(s) = -\frac{1}{2} \eta^{-1}(y(s)), \quad y(0) = r.
\]

**Proof.** Recall that Assumption 2.1 guarantees the global existence and uniqueness of solutions of (2.1). Given the process \( \{V(x(t))\}_{t \geq t_0} \), the proof strategy is to obtain an ordinary integral inequality for \( \mathbb{E}[V(x(t))] \) using Itô formula (2.3), and then use a comparison principle to translate the problem into one of standard input-to-state stability for an appropriate choice of the input. To carry out this strategy, we first need to ensure that the expectation of the integral against Brownian motion in (2.3) is 0. Let \( S_k = \{ x \in \mathbb{R}^n : \|x\|_2 \leq k \} \) be the ball of radius \( k \) centered at the origin. Fix \( x_0 \in \mathbb{R}^n \) and denote \( \tau_k \) as the first exit time of \( x(t) \) from \( S_k \) for integer values of \( k \) greater than \( \|x_0\|_2 \), namely, \( \tau_k \triangleq \inf \{ s \geq t_0 : \|x(s)\|_2 \geq k \} \), for \( k > \|x_0\|_2 \). Since the event \( \{ \omega \in \Omega : \tau_k \leq t \} \) belongs to \( \mathcal{F}_t \), \( \tau_k \) is an \( \{\mathcal{F}_t\} \)-stopping time. Now, for each \( k \) fixed, if
we consider the random variable \( t_k(t) \triangleq \min\{t, \tau_k\} \) and define \( I(t) \) as the stochastic integral in (2.3), then the process \( \hat{I}_k(t) \triangleq I(t_k(t)) \) has zero expectation as we show next. The function \( X : S_k \times [t_0, t] \to \mathbb{R} \) given by \( X(x, s) \triangleq \nabla V(x)^T G(x, s) \Sigma(s) \) is essentially bounded, and thus \( \mathbb{E}\left[ \int_{t_0}^{t} 1_{[t_0, t_k(t)]}(s) X(x(s), t)^2 \, ds \right] < \infty \), where \( 1_{[t_0, t_k(t)]}(s) \) is the indicator function of the set \([t_0, t_k(t)]\). Therefore, \( \mathbb{E}\left[ \hat{I}_k(t) \right] = 0 \) by [14, Th. 5.16, p. 26].

Define now \( \tilde{V}(t) \triangleq \mathbb{E}[V(x(t))] \) and \( W(t) \triangleq \mathbb{E}[W(x(t))] \) in \( \Gamma \triangleq \{ t \geq t_0 : \tilde{V}(t) < \infty \} \). By the above, taking expectations in (2.3) and using (3.4), we obtain that

\[
\tilde{V}(t) = \mathbb{E}\left[ \int_{t_0}^{t} \mathcal{L}[V](x(s), s) \, ds \right] \leq \tilde{V}(t_0) - \mathbb{E}\left[ \int_{t_0}^{t} W(x(s)) \, ds \right] + \mathbb{E}\left[ \int_{t_0}^{t} \sigma(|\Sigma(s)|_x) \, ds \right],
\]

for all \( t \in \Gamma \). Next we use the facts that \( V \) is continuous and \( \{x(t)\}_{t \geq t_0} \) is also continuous with probability 1. In addition, according to Fatou’s lemma [16, p. 123] for convergence in the probability measure, we get that, for all \( t \in \Gamma \),

\[
\tilde{V}(t) = \mathbb{E}\left[ V\left(\liminf_{k \to \infty} t_k(t)\right)\right] = \mathbb{E}\left[ \liminf_{k \to \infty} V(x(t_k(t))) \right] \leq \liminf_{k \to \infty} \mathbb{E}\left[ V(x(t_k(t))) \right] = \liminf_{k \to \infty} \tilde{V}(t_k(t)).
\]

Moreover, using the monotone convergence [16, p. 176] when \( k \to \infty \) in both Lebesgue integrals in (3.11) (because both integrands are nonnegative and \( 1_{[t_0, t_k(t)]} \) converges monotonically to \( 1_{[t_0, t]} \) as \( k \to \infty \)), we obtain from (3.12) that, for all \( t \in \Gamma \),

\[
\tilde{V}(t) \leq \tilde{V}(t_0) - \mathbb{E}\left[ \int_{t_0}^{t} W(x(s)) \, ds \right] + \mathbb{E}\left[ \int_{t_0}^{t} \sigma(|\Sigma(s)|_x) \, ds \right].
\]

Before resuming the argument we make two observations. First, applying Tonelli’s theorem [16, p. 212] to the nonnegative process \( \{W(x(s))\}_{s \geq t_0} \), it follows that

\[
\mathbb{E}\left[ \int_{t_0}^{t} W(x(s)) \, ds \right] = \int_{t_0}^{t} \tilde{W}(x) \, ds.
\]

Second, using (3.3) and Jensen’s inequality [1, Ch. 3], we get that

\[
\tilde{V}(t) = \mathbb{E}[V(x(t))] \leq \mathbb{E}[\eta(W(x(t)))] \leq \eta\left(\mathbb{E}[W(x(t))]\right) = \eta(\tilde{W}(t)),
\]

because \( \eta \) is concave, so \( \tilde{W}(t) \geq \eta^{-1}(\tilde{V}(t)) \). Hence, (3.13) and (3.14) yield

\[
\tilde{V}(t) \leq \tilde{V}(t_0) - \int_{t_0}^{t} \tilde{W}(s) \, ds + \int_{t_0}^{t} \sigma(|\Sigma(s)|_x) \, ds \\
\leq \tilde{V}(t_0) + \int_{t_0}^{t} \left( - \eta^{-1}(\tilde{V}(s)) + \sigma(|\Sigma(s)|_x) \right) \, ds,
\]

for all \( t \in \Gamma \). Define now the nonnegative input \( d(t) \triangleq \sigma(|\Sigma(t)|_x) \), which is Borel measurable and essentially locally bounded. According to the comparison principle, the solutions [22, Sec. C.2] of the initial value problem

\[
\dot{U}(t) = -\eta^{-1}(U(t)) + d(t), \quad U_0 \triangleq U(t_0) = \tilde{V}(t_0),
\]
satisfy that \( U(t) \geq \bar{V}(t) (\geq 0) \) wherever they exist. Regarding existence and uniqueness of solutions \([25, \text{Sec. I. A}]\) of (3.16), consider the function \( h : \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \to \mathbb{R} \)
given by \( h(y,d) \triangleq -\eta^{-1}(y) + d \). Since \( \alpha \triangleq \eta^{-1} \) is convex and class \( K_\infty \) (see Section 2.2), it holds that

\[
\alpha(s') \leq \alpha(s) \leq \alpha(s') + \frac{\alpha(s''') - \alpha(s')}{(s''') - (s')}(s - s')
\]

for all \( s \in [s', s'''] \), for any \( s''' > s' \geq 0 \). Thus, \( |\alpha(s) - \alpha(s')| = \alpha(s) - \alpha(s') \leq L(s-s') \), for any \( s'' > s \geq s' \geq 0 \), where \( L \triangleq (\alpha(s''') - \alpha(s'))/(s''' - s') \), so \( \eta^{-1} \) is locally Lipschitz. Hence, \( h \) is locally Lipschitz in \( \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \). Therefore, given the input \( d \geq 0 \) and any \( U_0 \geq 0 \), there is a unique maximal solution of (3.16), \( U(U_0, t_0; t) \), defined in a maximal interval \([t_0, t_{\text{max}}(U_0, t_0)) \). (As a by-product, the initial value problem (3.10), which can be written as \( \dot{y}(s) = \frac{1}{2}h(y(s), 0), \, y(0) = r \), has a unique and strictly decreasing solution in \([0, \infty) \), so \( \bar{\mu} \) in the statement is well defined and in class \( KL \).)

We show next that (3.16) is input-to-state stable (ISS) following a similar argument as in the proof of [23, Th. 5]; as a consequence, we obtain that \( t_{\text{max}}(U_0, t_0) = \infty \).

Firstly, if \( \eta^{-1}(U) \geq 2d \), then \( U(t) \leq \frac{1}{2} \eta^{-1}(U(t)) \), which implies that the trajectory \( U \) is nonincreasing outside the set \( S \triangleq \{ t \geq t_0 : U(t) \leq \eta(2d(t)) \} \). Thus, if \( t^* \) belongs to \( S \), then so does every \( t \in [t^*, t_{\text{max}}(U_0, t_0)) \), so the trajectory \( U \) is essentially locally bounded because \( d \geq 0 \) is too (and, additionally, \( U(t) \geq 0 \) by the comparison principle used above); hence, \( t_{\text{max}}(U_0, t_0) = \infty \) (and thus \( \Gamma = [t_0, \infty) \)). Therefore, for all \( t \geq t_0 \), and for \( \bar{\mu} \) as in the statement (which we have shown is well defined), we have that

\[
\bar{V}(t) \leq U(t) \leq \max \left\{ \bar{\mu}(\bar{V}(t_0), t - t_0) , \eta\left(2 \text{ess sup} \, d(s)\right) \right\}
\]

Since the maximum of two quantities is upper bounded by the sum, and using the definition of \( d \) together with the monotonicity of \( \sigma \), it follows that

\[
\bar{V}(t) \leq \bar{\mu}(V(x_0), t - t_0) + \eta\left(2\sigma\left(\text{ess sup} \, |\Sigma(s)|_x\right)\right),
\]

where we also used that \( \bar{V}(t_0) = V(x_0) \). This completes the proof. \( \square \)

Of particular interest to us is the case when the function \( V \) is lower and upper bounded by class \( K_\infty \) functions of the distance to a closed, not necessarily bounded, set.

DEFINITION 3.7. (NSS-Lyapunov functions). A function \( V \in C^2(\mathbb{R}^n; \mathbb{R}_{\geq 0}) \) is a strong

NSS-Lyapunov function in probability with respect to \( U \subseteq \mathbb{R}^n \) for (2.1) if \( V \) is a

noise-dissipative Lyapunov function and, in addition, there exist \( p > 0 \) and class \( K_\infty \)

functions \( \alpha_1 \) and \( \alpha_2 \) such that

\[
(3.17) \quad \alpha_1(||x||_p^p) \leq V(x) \leq \alpha_2(||x||_p^p), \quad \forall x \in \mathbb{R}^n.
\]

If, moreover, \( \alpha_1 \) is convex, then \( V \) is a \( p \)-th moment NSS-Lyapunov function with
respect to \( U \).

Note that a strong NSS-Lyapunov function in probability with respect to a set satisfies

an inequality of the type (3.17) for any \( p > 0 \), whereas the choice of \( p \) is relevant

when \( \alpha_1 \) is required to be convex. The reason for the ‘strong’ terminology is that

we require (3.5) to be satisfied with convex \( \eta^{-1} \in K_\infty \). Instead, a standard

NSS-Lyapunov function in probability satisfies the same inequality with a class \( K_\infty \) function
which is not necessarily convex. We also note that (3.17) implies that \( \mathcal{U} = \{ x \in \mathbb{R}^n : V(x) = 0 \} \), which is closed because \( V \) is continuous.

**Example 3.8.** (Example 3.5–revisited: an NSS-Lyapunov function). Consider the function \( V \) introduced in Example 3.5. For each \( p \in (0, 2] \), note that

\[
\alpha_1 p (\| x - x^\ast \|^p_2) \leq V(x) \leq \alpha_2 p (\| x - x^\ast \|^p_2) \quad \forall x \in \mathbb{R}^n,
\]

for the convex functions \( \alpha_1 p (r) = \alpha_2 p (r) \equiv r^{2/p} \), which are in the class \( \mathcal{K}_{\infty} \). (Recall that \( \alpha_2 \) in Definition 3.7 is only required to be \( r^{1/2} \) in the first equation, Chebyshev’s inequality [1, Ch. 3], and the upper bound obtained in Theorem 3.6, cf. (3.9), in the last inequality (leveraging the monotonicity (3.20)).

Substituting now \( \hat{\alpha} \), that

\[
\mu(p, s) \triangleq \alpha_1^{-1} \left( \frac{\bar{\mu}(\alpha_2(p^p), s)}{2} \right), \quad \theta(r) \triangleq \alpha_1^{-1} \left( \frac{\eta(2\sigma(r))}{2} \right);
\]

where we have used the strict monotonicity of \( \alpha_1 \) in the first equation, Chebyshev’s inequality [1, Ch. 3] in the second inequality, and the upper bound for \( \mathbb{E}[V(x(t))] \) obtained in Theorem 3.6, cf. (3.9), in the last inequality (leveraging the monotonicity of \( \bar{\mu} \) in the first argument and the fact that \( V(x) \leq \alpha_2 (|x|^p) \) for all \( x \in \mathbb{R}^n \)). Also, for any function \( \alpha \in \mathcal{K} \), we have that \( \alpha(2r) + \alpha(2s) \geq \alpha(r + s) \) for all \( r, s \geq 0 \). Thus,

\[
\rho(\epsilon, x_0, t) \triangleq \mu(|x_0|_2, t - t_0) + \theta\left( \text{ess sup}_{t_0 \leq s \leq t} |\Sigma(s)|_2 \right) \geq \alpha_1^{-1} \left( \frac{1}{\epsilon} \bar{\mu}(\alpha_2(|x_0|^p), t - t_0) + \frac{1}{\epsilon} \eta\left( \text{ess sup}_{t_0 \leq s \leq t} |\Sigma(s)|_2 \right) \right) \triangleq \hat{\rho}(\epsilon).
\]

Substituting now \( \hat{\rho} \) in (3.19), and using that \( \rho(\epsilon, x_0, t) \geq \hat{\rho}(\epsilon) \), we get that

\[
\mathbb{P}\{ |x(t)|^p_2 > \rho(\epsilon, x_0, t) \} \leq \mathbb{P}\{ |x(t)|^p_2 > \hat{\rho}(\epsilon) \} \leq \epsilon.
\]

To show (ii), since \( \alpha_1^{-1} \) is concave, applying Jensen’s inequality [1, Ch. 3], we get

\[
\mathbb{E}[|x(t)|^p_2] \leq \mathbb{E}\left[ \alpha_1^{-1}\left( V(x(t)) \right) \right] \leq \alpha_1^{-1}\left( \mathbb{E}[V(x(t))] \right) \leq \hat{\rho}(1) \leq \rho(1, x_0, t),
\]
where in the last two inequalities we have used the bound for $\mathbb{E}[V(x(t))]$ in (3.19) and the definition of $\beta(\varepsilon)$ in (3.20).

**Example 3.10.** (Example 3.5–revisited: illustration of Corollary 3.9). Consider again Example 3.5. Since $V$ is a $p$th moment NSS-Lyapunov function for (3.8) with respect to the point $x^*$ for $p \in (0, 2]$, as shown in Example 3.8, Corollary 3.9 implies that

$$
E[\|x - x^*\|_2^p] \leq \mu(\|x_0 - x^*\|_2, t - t_0) + \theta\left(\text{ess sup}_{t_0 \leq s \leq t} |\Sigma(s)|_F\right),
$$

for all $t \geq t_0$, $x_0 \in \mathbb{R}^n$, and $p \in (0, 2]$, where

$$
\mu(r, s) = (2\hat{\mu}(r^2, s))^{p/2}, \quad \theta(r) = (\gamma^{-1}(r^2))^{p/2},
$$

and the class $KL$ function $\hat{\mu}$ is defined as the solution to the initial value problem (3.10) with $\eta(r) = \frac{1}{2}\gamma^{-1}(r)$. Figure 3.1 illustrates this noise-to-state stability property. We note that if the function $h$ is strongly convex, i.e., if $\gamma(r) = c_\gamma r$ for some constant $c_\gamma > 0$, then $\hat{\mu} : \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ becomes $\hat{\mu}(r, s) = re^{-c_\gamma s}$, and $\mu(r, s) = 2^{p/2} r^p e^{-c_\gamma s} p/2^s$, so the bound for $E[\|x - x^*\|_2^p]$ in (3.21) decays exponentially with time to $\theta(\text{ess sup}_{t_0 \leq s \leq t} |\Sigma(s)|_F)$.

**4. Refinements of the notion of proper functions.** In this section, we analyze in detail the inequalities between functions that appear in the definition of noise-dissipative Lyapunov function, strong NSS-Lyapunov function in probability, and $p$th moment NSS-Lyapunov function. In Section 4.1, we establish that these inequalities can be regarded as equivalence relations. In Section 4.2, we make a complete
characterization of the properties of two functions related by these equivalence relations. Finally, in Section 4.3, these results lead us to obtain an alternative formulation of Corollary 3.9.

4.1. Proper functions and equivalence relations. Here, we provide a refinement of the notion of proper functions with respect to each other. Proper functions play an important role in stability analysis, see e.g., [6, 23].

Definition 4.1. (Refinements of the notion of proper functions with respect to each other). Let $D \subseteq \mathbb{R}^n$ and the functions $V, W : D \to \mathbb{R}_{\geq 0}$ be such that

$$
\alpha_1(W(x)) \leq V(x) \leq \alpha_2(W(x)), \quad \forall x \in D,
$$

for some functions $\alpha_1, \alpha_2 : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$. Then,

(i) if $\alpha_1, \alpha_2 \in K$, we say that $V$ is $K$-dominated by $W$ in $D$, and write $V \preceq^K W$ in $D$;

(ii) if $\alpha_1, \alpha_2 \in K_\infty$, we say that $V$ and $W$ are $K_\infty$-proper with respect to each other in $D$, and write $V \sim_{K_\infty} W$ in $D$;

(iii) if $\alpha_1, \alpha_2 \in K_\infty$ are convex and concave, respectively, we say that $V$ and $W$ are $K_{\infty}^{cc}$-(convex-concave) proper with respect to each other in $D$, and write $V \sim_{K_{\infty}^{cc}} W$ in $D$;

(iv) if $\alpha_1(r) \triangleq c_{\alpha_1} r$ and $\alpha_2(r) \triangleq c_{\alpha_2} r$, for some constants $c_{\alpha_1}, c_{\alpha_2} > 0$, we say that $V$ and $W$ are equivalent in $D$, and write $V \sim W$ in $D$.

Note that the relations in Definition 4.1 are nested, i.e., given $V, W : D \to \mathbb{R}_{\geq 0}$, the following chain of implications hold in $D$:

$$
(4.1) \quad V \sim W \Rightarrow V \sim_{K_\infty} W \Rightarrow V \sim_{K_{\infty}^{cc}} W \Rightarrow V \preceq^K W.
$$

Also, note that if $W(x) = \|x\|_2$, $D$ is a neighborhood of 0, and $\alpha_1, \alpha_2$ are class $K$, then we recover the notion of $V$ being a proper function [6]. If $D = \mathbb{R}^n$, and $V$ and $W$ are seminorms, then the relation $\sim$ corresponds to the concept of equivalent seminorms.

The relation $\sim_{K_\infty}$ is relevant for ISS and NSS in probability, whereas the relation $\sim_{K_{\infty}^{cc}}$ is important for $p$th moment NSS. The latter is because the inequalities in $\sim_{K_{\infty}^{cc}}$ are still valid, thanks to Jensen inequality, if we substitute $V$ and $W$ by their expectations along a stochastic process. Another fact about the relation $\sim_{K_{\infty}^{cc}}$ is that $\alpha_1, \alpha_2 \in K_\infty$, convex and concave, respectively, must be asymptotically linear if $V(D) \supseteq [s_0, \infty)$, for some $s_0 \geq 0$, so that $\alpha_1(s) \leq \alpha_2(s)$ for all $s \geq s_0$. This follows from Lemma A.1.

Remark 4.2. (Quadratic forms in a constrained domain). It is sometimes convenient to view the functions $V, W : D \to \mathbb{R}_{\geq 0}$ as defined in a domain where their functional expression becomes simpler. To make this idea precise, assume there exist $i : D \subset \mathbb{R}^n \to \mathbb{R}^m$, with $m \geq n$, and $V, W : \tilde{D} \to \mathbb{R}_{\geq 0}$, where $\tilde{D} = i(D)$, such that $V = \hat{V} \circ i$ and $W = \hat{W} \circ i$. If this is the case, then the existence of $\alpha_1, \alpha_2 : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ such that $\alpha_1(\hat{W}(\hat{x})) \leq \hat{V}(\hat{x}) \leq \alpha_2(\hat{W}(\hat{x}))$, for all $\hat{x} \in \tilde{D}$, implies that $\alpha_1(W(x)) \leq V(x) \leq \alpha_2(W(x))$, for all $x \in D$. The reason is that for any $x \in D$ there exists $\hat{x} \in \tilde{D}$, given by $\hat{x} = i(x)$, such that $V(x) = \hat{V}(\hat{x})$ and $W(x) = \hat{W}(\hat{x})$, so

$$
\alpha_1(W(x)) = \alpha_1(\hat{W}(\hat{x})) \leq V(x) = \hat{V}(\hat{x}) \leq \alpha_2(\hat{W}(\hat{x})) = \alpha_2(W(x)).
$$

Consequently, if any of the relations given in Definition 4.1 is satisfied by $\hat{V}, \hat{W}$ in $\tilde{D}$, then the corresponding relation is satisfied by $V, W$ in $D$. For instance, in
some scenarios this procedure can allow us to rewrite the original functions $V, W$ as quadratic forms $\hat{V}, \hat{W}$ in a constrained set of an extended Euclidean space, where it is easier to establish the appropriate relation between the functions. We make use of this observation in Section 4.3 below.

**Lemma 4.3.** (Powers of seminorms with the same nullspace) Let $A$ and $B$ in $\mathbb{R}^{m \times n}$ be nonzero matrices with the same nullspace, $N(A) = N(B)$. Then, for any $p, q > 0$, the inequalities $\alpha_1(\|x\|_A^p) \leq \|x\|_B^p \leq \alpha_2(\|x\|_A^p)$ are verified with

$$
\alpha_1(r) = \alpha_1 \left( \left( \frac{\lambda_{n-k}(B^T B)}{\lambda_{n-k}(A^T A)} \right)^{\frac{q}{p}} \right)^{\frac{q}{p}}, \quad \alpha_2(r) = \alpha_2 \left( \sqrt{\lambda_{\max}(B^T B)/\lambda_{\max}(A^T A)} \right) \left( \frac{\|x\|_A^p}{\|x\|_B^p} \right),
$$

where $k \triangleq \dim(N(A))$. In particular, $\|\cdot\|_A^p \sim \|\cdot\|_B^p$ and $\|\cdot\|_A^p \sim \|\cdot\|_B^p$ in $\mathbb{R}^n$ for any real numbers $p, q > 0$.

**Proof.** For $U \triangleq N(A)$, write any $x \in \mathbb{R}^n$ as $x = x_\ell + x_{\ell^*}$, where $x_\ell \in U$ and $x_{\ell^*} \in \{ x \in \mathbb{R}^n : x^T u = 0, \forall u \in U \}$, so that $Ax = A(x_\ell + x_{\ell^*}) = Ax_\ell + Bx_{\ell^*}$ and $Bx = Bx_{\ell^*}$ because $N(A) = N(B) = U$. Using the formulas for the eigenvalues in [5, p. 178], we see that the next chain of inequalities hold:

$$
\alpha_1(\|x\|_A^p) = \alpha_1 \left( \left( \frac{\lambda_{n-k}(B^T B)}{\lambda_{n-k}(A^T A)} \right)^{\frac{q}{p}} \right) \leq \alpha_1 \left( \left( \lambda_{\max}(A^T A) x^T_{\ell^*} x_{\ell^*} \right) \right) \leq \alpha_2 \left( \left( \frac{\lambda_{n-k}(B^T B)}{\lambda_{n-k}(A^T A)} \right)^{\frac{q}{p}} \right) \leq \alpha_2(\|x\|_A^p),
$$

where $\|x\|_B^p = (x^T_{\ell^*} B^T B x_{\ell^*})^{\frac{q}{p}}$. From this we conclude that $\|\cdot\|_A^p \sim \|\cdot\|_B^p$ in $\mathbb{R}^n$. Finally, when $p = q$, the class $\kappa_\sim$ functions $\alpha_1, \alpha_2$ in the statement are linear, so we obtain that $\|\cdot\|_A^p \sim \|\cdot\|_B^p$ in $\mathbb{R}^n$. \( \square \)

Next we show that $\sim_{\kappa_\sim}$ and $\sim_{\kappa_{\sim}}$ are reflexive, symmetric, and transitive, and hence define equivalence relations.

**Lemma 4.4.** (The $\kappa_\sim$- and $\kappa_{\sim}$-proper relations are equivalence relations). The relations $\sim_{\kappa_\sim}$ and $\sim_{\kappa_{\sim}}$ in any set $D \subseteq \mathbb{R}^n$ are both equivalence relations.

**Proof.** For convenience, we represent both relations by $\sim^\ast$. Both are reflexive, i.e., $V \sim^\ast V$, because one can take $\alpha_1(r) = \alpha_2(r) = r$ noting that a linear function is both convex and concave. Both are symmetric, i.e., $V \sim^\ast W$ if and only if $W \sim^\ast V$, because if $\alpha_1 \circ W \leq V \leq \alpha_2 \circ W$ in $D$, then $\alpha_2^{-1} \circ V \leq W \leq \alpha_1^{-1} \circ V$ in $D$. In the case of $\sim_{\kappa_\sim}$, the inverse of a class $\kappa_\sim$ functions is class $\kappa_\sim$. Additionally, in the case of $\sim_{\kappa_{\sim}}$, if $\alpha \in \kappa_{\sim}$ is convex (respectively, concave), then $\alpha^{-1} \in \kappa_{\sim}$ is concave (respectively, convex). Finally, both are transitive, i.e., $U \sim^\ast V$ and $V \sim^\ast W$ imply $U \sim^\ast W$, because if $\alpha_1 \circ V \leq U \leq \alpha_2 \circ V$ and $\alpha_1 \circ W \leq \alpha_2 \circ W$ in $D$, then $\alpha_1 \circ \alpha_2 \circ W \leq U \leq \alpha_2 \circ \alpha_1 \circ W$ in $D$. In the case of $\sim_{\kappa_{\sim}}$, the composition of two class $\kappa_{\sim}$ functions is class $\kappa_{\sim}$. Additionally, in the case of $\sim_{\kappa_{\sim}}$, if $\alpha_1, \alpha_2 \in \kappa_{\sim}$ are both convex (respectively, concave), then the compositions $\alpha_1 \circ \alpha_2$ and $\alpha_2 \circ \alpha_1$ belong to $\kappa_{\sim}$ and are convex (respectively, concave), as explained in Section 2.2. \( \square \)

**Remark 4.5.** (The relation $\triangleleft^\ast$ is not an equivalence relation). The proof above also shows that the relation $\triangleleft^\ast$ is reflexive and transitive. However, it is not symmetric: consider $V, W \in \mathcal{C}(\mathbb{R}^n; \mathbb{R}_+)$ given by $V(x) = 1 - e^{-\|x\|_2}$ and $W(x) = \|x\|_2$. Clearly, $V \triangleleft^\ast W$ in $\mathbb{R}^n$ by taking $\alpha_1 = \alpha_2 = \alpha \in \kappa$, with $\alpha(s) = 1 - e^{-s}$. On the other
hand, if there exist $\tilde{\alpha}_1, \tilde{\alpha}_2 \in \mathcal{K}$ such that $\tilde{\alpha}_1(V(x)) \leq W(x) \leq \tilde{\alpha}_2(V(x))$ for all $x \in \mathbb{R}^n$, then we reach the contradiction, by continuity of $\tilde{\alpha}_2$, that $\lim_{\|x\|_2 \to \infty} \|x\|_2 \leq \tilde{\alpha}_2\left(\lim_{\|x\|_2 \to \infty} (1 - e^{-\|x\|_2^2})\right) = \tilde{\alpha}_2(1) < \infty$.

4.2. Characterization of proper functions with respect to each other.
In this section, we provide a complete characterization of the properties that two functions must satisfy to be related by the equivalence relations defined in Section 4.1. For $\mathcal{D} \subseteq \mathbb{R}^n$, consider $V_1, V_2 : \mathcal{D} \to \mathbb{R}_{\geq 0}$. Given a real number $p > 0$, define

$$\phi_p(s) \triangleq \sup_{\{y \in \mathcal{D} : V_2(y) \leq s\}} V_1(y),$$

$$\psi_p(s) \triangleq \inf_{\{y \in \mathcal{D} : V_2(y) \geq s\}} V_1(y),$$

for $s \geq 0$. The value $\phi_p(s)$ gives the supremum of the function $V_1$ in the $\sqrt{s}$-sublevel set of $V_2$, and $\psi_p(s)$ is the infimum of $V_1$ in the $\sqrt{s}$-superlevel set of $V_2$. Thus, the functions $\phi_p$ and $\psi_p$ satisfy

$$(4.2) \psi_p(V_2(x)^p) = \inf_{\{y \in \mathcal{D} : V_2(y) \leq V_2(x)\}} V_1(y) \leq V_1(x) \leq \sup_{\{y \in \mathcal{D} : V_2(y) \leq V_2(x)\}} V_1(y) = \phi_p(V_2(x)^p),$$

for all $x \in \mathcal{D}$, which suggests $\phi_p$ and $\psi_p$ as pre-comparison functions to construct $\alpha_1$ and $\alpha_2$ in Definition 4.1. To this end, we find it useful to formulate the following properties of the function $V_1$ with respect to $V_2$:

P0: The set $\{x \in \mathcal{D} : V_2(x) = s\}$ is nonempty for all $s \geq 0$.

P1: The nullsets of $V_1$ and $V_2$ are the same, i.e., $\{x \in \mathcal{D} : V_1(x) = 0\} = \{x \in \mathcal{D} : V_2(x) = 0\}$.

P2: The function $\phi_1$ is locally bounded in $\mathbb{R}_{\geq 0}$ and right continuous at 0, and $\psi_1$ is positive definite.

P3: The next limit holds: $\lim_{s \to \infty} \psi_1(s) = \infty$.

P4 (as a function of $p > 0$): The asymptotic behavior of $\phi_p$ and $\psi_p$ is such that $\phi_p(s)$ and $s^2/\psi_p(s)$ are both in $O(s)$ as $s \to \infty$.

The next result shows that these properties completely characterize whether the functions $V_1$ and $V_2$ are related through the equivalence relations defined in Section 4.1. This result generalizes [6, Lemma 4.3] in several ways: the notions of proper functions considered here are more general and are not necessarily restricted to a relationship between an arbitrary function and the distance to a compact set.

**Theorem 4.6.** (Characterizations of proper functions with respect to each other).
Let $V_1, V_2 : \mathcal{D} \to \mathbb{R}_{\geq 0}$, and assume $V_2$ satisfies P0. Then

(i) $V_1$ satisfies $\{P_i\}_{i=1}^3$ with respect to $V_2$ $\iff$ $V_1 \preccurlyeq^\mathcal{E} V_2$ in $\mathcal{D}$;

(ii) $V_1$ satisfies $\{P_i\}_{i=1}^3$ with respect to $V_2$ $\iff$ $V_1 \sim^\mathcal{E} V_2$ in $\mathcal{D}$;

(iii) $V_1$ satisfies $\{P_i\}_{i=1}^3$ with respect to $V_2$ for $p > 0$ $\iff$ $V_1 \sim^\mathcal{E} V_2^p$ in $\mathcal{D}$.

**Proof.** We begin by establishing a few basic facts about the pre-comparison functions $\psi_p$ and $\phi_p$. By definition and by P0, it follows that $0 \leq \psi_1(s) \leq \phi_1(s)$ for all $s \geq 0$. Since $\phi_1$ is locally bounded by P2, then so is $\psi_1$. In particular, $\phi_1$ and $\psi_1$ are well defined in $\mathbb{R}_{\geq 0}$. Moreover, both $\phi_1$ and $\psi_1$ are nondecreasing because if $s_2 \geq s_1$, then
the supremum is taken in a larger set, \( \{x \in \mathcal{D} : V_2(x) \leq s_2\} \supseteq \{x \in \mathcal{D} : V_2(x) \leq s_1\} \), and the infimum is taken in a smaller set, \( \{x \in \mathcal{D} : V_2(x) \geq s_2\} \subseteq \{x \in \mathcal{D} : V_2(x) \geq s_1\} \). Furthermore, for any \( q > 0 \), the functions \( \phi_q \) and \( \psi_q \) are also monotonic and positive definite because \( \phi_q(s) = \phi_1(\sqrt{s}) \) and \( \psi_q(s) = \psi_1(\sqrt{s}) \) for all \( s \geq 0 \). We now use these properties of the pre-comparison functions to construct \( \alpha_1, \alpha_2 \) in Definition 4.1 required by the implications from left to right in each statement.

Proof of (i) \( \Rightarrow \). To show the existence of \( \alpha_2 \in \mathcal{K} \) such that \( \alpha_2(s) \geq \phi_1(s) \) for all \( s \in \mathbb{R}_{\geq 0} \), we proceed as follows. Since \( \phi_1 \) is locally bounded and nondecreasing, given a strictly increasing sequence \( \{b_k\}_{k \geq 1} \subseteq \mathbb{R}_{\geq 0} \) with \( \lim_{k \to \infty} b_k = \infty \), we choose the sequence \( \{M_k\}_{k \geq 1} \subseteq \mathbb{R}_{\geq 0} \), setting \( M_0 = 0 \), in the following way:

\[
M_k \triangleq \max \left\{ \sup_{s \in [0,b_k]} \phi_1(s), M_{k-1} + 1/k^2 \right\} = \max \{\phi_1(b_k), M_{k-1} + 1/k^2\}.
\]

This choice guarantees that \( \{M_k\}_{k \geq 1} \) is strictly increasing and, for each \( k \geq 1 \),

\[
0 \leq M_k - \phi_1(b_k) \leq \sum_{i=1}^{k} \frac{1}{i^2} \leq \pi^2/6.
\]

Also, since \( \phi_1 \) is right continuous at 0, we can choose \( b_1 > 0 \) such that there exists \( \alpha_2 : [0,b_1] \to \mathbb{R}_{\geq 0} \) continuous, positive definite and strictly increasing, satisfying that \( \alpha_2(s) \geq \phi_1(s) \) for all \( s \in [0,b_1] \) and with \( \alpha_2(b_1) = M_2 \). (This is possible because the only function that cannot be upper bounded by an arbitrary continuous function in some arbitrarily small interval \( [0,b_1] \) is the function that has a jump at 0.) The rest of the construction is explicit. We define \( \alpha_2 \) as a piecewise linear function in \( (b_1,\infty) \) in the following way: for each \( k \geq 2 \), we define

\[
\alpha_2(s) \triangleq \alpha_2(b_{k-1}) + \frac{M_{k+1} - \alpha_2(b_{k-1})}{b_k - b_{k-1}}(s - b_{k-1}), \quad \forall s \in (b_{k-1},b_k).
\]

The resulting \( \alpha_2 \) is continuous by construction. Also, \( \alpha_2(b_1) = M_2 \), so that, inductively, \( \alpha_2(b_{k-1}) = M_k \) for \( k \geq 2 \). Two facts now follow: first, \( M_{k+1} - \alpha_2(b_{k-1}) = M_k - M_k \geq 1/(k+1)^2 \) for \( k \geq 2 \), so \( \alpha_2 \) has positive slope in each interval \( (b_{k-1},b_k] \) and thus is strictly increasing in \( (b_1,\infty) \); second, \( \alpha_2(s) > \alpha_2(b_{k-1}) = M_k \geq \phi_1(b_k) \geq \phi_1(s) \) for all \( s \in (b_{k-1},b_k] \), for each \( k \geq 2 \), so \( \alpha_2(s) \geq \phi_1(s) \) for all \( s \in (b_1,\infty) \).

We have left to show the existence of \( \alpha_1 \in \mathcal{K} \) such that \( \alpha_1(s) \leq \psi_1(s) \) for all \( s \in \mathbb{R}_{\geq 0} \). First, since \( 0 \leq \psi_1(s) \leq \phi_1(s) \) for all \( s \geq 0 \) by definition and by P0, using the sandwich theorem \([11, \text{p. 107}]\), we derive that \( \psi_1 \) is right continuous at 0 the same as \( \phi_1 \). In addition, since \( \psi_1 \) is nondecreasing, it can only have a countable number of jump discontinuities (none of them at 0). Therefore, we can pick \( c_1 > 0 \) such that a continuous and nondecreasing function \( \hat{\psi}_1 \) can be constructed in \([0,c_1]\) by removing the jumps of \( \psi_1 \), so that \( \hat{\psi}_1(s) \leq \psi_1(s) \). Moreover, since \( \hat{\psi}_1 \) is positive definite and right continuous at 0, then \( \hat{\psi}_1 \) is also positive definite. Thus, there exists \( \alpha_1 \) in \([0,c_1]\) continuous, positive definite, and strictly increasing, such that, for some \( r < 1 \),

\[
\alpha_1(s) \leq r\hat{\psi}_1(s) \leq r\psi_1(s)
\]

for all \( s \in [0,c_1] \). To extend \( \alpha_1 \) to a function in class \( \mathcal{K} \) in \( \mathbb{R}_{\geq 0} \), we follow a similar strategy as for \( \alpha_2 \). Given a strictly increasing sequence \( \{c_k\}_{k \geq 2} \subseteq \mathbb{R}_{\geq 0} \) with
Second, the construction of $\alpha$ so there exists
\begin{equation}
\psi(s) = \psi_1(s) - \frac{\psi_1(c_1) - \alpha_1(c_1)}{1 + k^2} = \psi_1(c_k) - \frac{\psi_1(c_1) - \alpha_1(c_1)}{1 + k^2},
\end{equation}
Next we define $\alpha_1$ in $[c_1, \infty)$ as the piecewise linear function
\begin{equation}
\alpha_1(s) \triangleq \alpha_1(c_k) + \frac{m_k - \alpha_1(c_k)}{c_{k+1} - c_k} (s - c_k), \quad \forall s \in [c_k, c_{k+1}),
\end{equation}
for all $k \geq 1$, so $\alpha_1$ is continuous by construction. It is also strictly increasing because $\alpha_1(c_2) = m_1 = (\psi_1(c_1) + \alpha_1(c_1))/2 > \alpha_1(c_1)$ by (4.5), and also, for each $k \geq 2$, the slopes are positive because $m_k - \alpha_1(c_k) = m_k - m_{k-1} > 0$ (due to the fact that $\{m_k\}_{k \geq 1}$ in (4.6) is strictly increasing because $\psi_1$ is nondecreasing). Finally, $\alpha_1(s) < \alpha_1(c_{k+1}) = m_k < \psi_1(c_k) \leq \psi_1(s)$ for all $s \in [c_k, c_{k+1})$, for all $k \geq 1$ by (4.6).

Equipped with $\alpha_1$, $\alpha_2$ as defined above, and as a consequence of (4.2), we have that
\begin{equation}
\alpha_1(V_2(x)) \leq \psi_1(V_2(x)) \leq V_1(x) \leq \phi_1(V_2(x)) \leq \alpha_2(V_2(x)), \quad \forall x \in D.
\end{equation}
This concludes the proof of (i) $\Rightarrow$.

As a preparation for (ii)-(iii) $\Rightarrow$, and assuming $P3$, we derive two facts regarding the functions $\alpha_1$ and $\alpha_2$ constructed above. First, we establish that
\begin{equation}
\alpha_2(s) \in \mathcal{O}(\phi_1(s)) \text{ as } s \to \infty.
\end{equation}
To show this, we argue that
\begin{equation}
\lim_{k \to \infty} \sup_{s \in [b_{k-1}, b_k]} (\alpha_2(s) - \phi_1(s)) \leq \lim_{k \to \infty} (\phi_1(b_{k+1}) - \phi_1(b_{k-1})) + \frac{\pi^2}{6},
\end{equation}
so that there exist $C, s_1 > 0$ such that $\alpha_2(s) \leq 3\phi_1(s) + C$, for all $s \geq s_1$. Thus, noting that $\lim_{s \to \infty} \phi_1(s) = \infty$, as a consequence of $P3$, the expression (4.8) holds.

To establish (4.9), we use the monotonicity of $\alpha_2$ and $\phi_1$, (4.3) and (4.4). For $k \geq 2$,
\begin{align*}
\sup_{s \in [b_{k-1}, b_k]} (\alpha_2(s) - \phi_1(s)) & \leq \alpha_2(b_k) - \phi_1(b_{k-1}) = M_{k+1} - \phi_1(b_{k-1}) \\
& = \max \{ \phi_1(b_{k+1}) - \phi_1(b_{k-1}), M_k + 1/(k + 1)^2 - \phi_1(b_{k-1}) \} \\
& \leq \max \{ \phi_1(b_{k+1}) - \phi_1(b_{k-1}), \phi_1(b_k) + \pi^2/6 + 1/(k + 1)^2 - \phi_1(b_{k-1}) \}.
\end{align*}
Second, the construction of $\alpha_1$ guarantees that
\begin{equation}
\psi_1(s) \in \mathcal{O}(\alpha_1(s)) \text{ as } s \to \infty,
\end{equation}
because, as we show next,
\begin{equation}
\lim_{k \to \infty} \sup_{s \in [c_k, c_{k+1})} (\psi_1(s) - \alpha_1(s)) \leq \lim_{k \to \infty} (\alpha_1(c_{k+2}) - \alpha_1(c_k))
\end{equation}
so there exists $s_2 > 0$ such that $\psi_1(s) \leq 3\alpha_1(s)$ for all $s \geq s_2$. To obtain (4.11), we leverage the monotonicity of $\psi_1$ and $\alpha_1$, and (4.6); namely, for $k \geq 2$,
\begin{align*}
\sup_{s \in [c_k, c_{k+1})} (\psi_1(s) - \alpha_1(s)) & \leq \psi_1(c_{k+1}) - \alpha_1(c_k) \\
& = m_{k+1} + \frac{\psi_1(c_1) - \alpha_1(c_1)}{1 + (k+1)^2} - \alpha_1(c_k) = \alpha_1(c_{k+2}) + \frac{\psi_1(c_1) - \alpha_1(c_1)}{1 + (k+1)^2} - \alpha_1(c_k),
\end{align*}
Equipped with (4.8) and (4.10), we prove next (ii)-(iii) \( \Rightarrow \).

Proof of (ii) \( \Rightarrow \): If, in addition, P3 holds, then \( \lim_{s \to \infty} \phi_1(s) \geq \lim_{s \to \infty} \psi_1(s) = \infty \). This guarantees that \( \alpha_2 \in K_{\infty} \). Also, according to (4.10), P3 implies that \( \alpha_1 \) is unbounded, and thus in \( K_{\infty} \) as well. The result now follows by (4.7).

Proof of (iii) \( \Rightarrow \): Finally, assume that P4 also holds for some \( p > 0 \). We show next the existence of the required convex and concave functions involved in the relation \( \sim_{K_{\infty}} \). Let \( \alpha_{1,p}(s) \equiv \alpha_1(\sqrt[p]{s}) \) and \( \alpha_{2,p}(s) \equiv \alpha_2(\sqrt[p]{s}) \) for \( s \geq 0 \), so that

\[
\alpha_{1,p}(s) = \alpha_1(\sqrt[p]{s}) \leq \psi_1(\sqrt[p]{s}) = \psi_p(s) \quad \text{and} \quad \phi_p(s) = \phi_1(\sqrt[p]{s}) \leq \alpha_2(\sqrt[p]{s}) = \alpha_{2,p}(s).
\]

From (4.8) and P4, it follows that there exist \( s', c_1, c_2 > 0 \) such that \( \alpha_{2,s}(s) \leq c_1 \phi_1(s) \) and \( \phi_p(s) \leq c_2 \psi_p(s) \) for all \( s \geq s' \). Thus,

\[
\alpha_{2,p}(s) = \alpha_2(\sqrt[p]{s}) \leq c_1 \phi_1(\sqrt[p]{s}) = c_1 \phi_p(s) \leq c_1 c_2 s,
\]

for all \( s \geq s' \), so \( \alpha_{2,p}(s) \) is in \( O(s) \) as \( s \to \infty \). Similarly, according to (4.10) and P4, there are constants \( s'', c_3, c_4 > 0 \) such that \( \psi_1(s) \leq c_3 \alpha_1(s) \) and \( s^2 \leq c_4 \psi_p(s) \) for all \( s \geq s'' \). Thus,

\[
s \alpha_{1,p}(s) = s \alpha_1(\sqrt[p]{s}) \geq s \frac{1}{c_3} \psi_1(\sqrt[p]{s}) = s \frac{1}{c_3} \psi_p(s) \geq \frac{1}{c_3 c_4} s^2,
\]

for all \( s \geq s'' \), so \( s^2/\alpha_{1,p}(s) \) is in \( O(s) \) as \( s \to \infty \). Summarizing, the construction of \( \alpha_1, \alpha_2 \) guarantees, under P4, that \( \alpha_{1,p}, \alpha_{2,p} \) satisfy that \( s^2/\alpha_{1,p}(s) \) and \( \alpha_{2,p}(s) \) are in \( O(s) \) as \( s \to \infty \) (and, as a consequence, so are \( s^2/\alpha_{2,p}(s) \) and \( \alpha_{1,p}(s) \)). Therefore, according to Lemma A.1, we can leverage (4.7) by taking \( \tilde{\alpha}_1, \tilde{\alpha}_2 \in K_{\infty} \), convex and concave, respectively, such that, for all \( x \in D \),

\[
\tilde{\alpha}_1(V_2(x)^p) \leq \alpha_{1,p}(V_2(x)^p) = \alpha_1(V_2(x)) \leq \psi_1(V_2(x)) \leq V_1(x) \leq \phi_1(V_2(x)) \leq \alpha_{2,p}(V_2(x)^p) \leq \tilde{\alpha}_2(V_2(x)^p).
\]

Proof of (i) \( \Leftarrow \): If there exist class \( K \) functions \( \alpha_1, \alpha_2 \) such that \( \alpha_1(V_2(x)) \leq V_1(x) \leq \alpha_2(V_2(x)) \) for all \( x \in D \), then the nullsets of \( V_1 \) and \( V_2 \) are the same, which is the property P1. In addition, \( 0 \leq \phi_1(s) \leq \alpha_2(s) \) for all \( s \geq 0 \), so \( \phi_1 \) is locally bounded and, moreover, the sandwich theorem guarantees that \( \phi_1 \) is right continuous at 0. Also, since \( \alpha_1(s) \leq \psi_1(s) \), for all \( s \geq 0 \), and \( \psi_1(0) = 0 \), it follows that \( \psi_1 \) is positive definite. Therefore, P2 also holds.

Proof of (ii) \( \Leftarrow \): Since \( \psi_1(s) \geq \alpha_1(s) \) for all \( s \geq 0 \), the property P3 follows because

\[
\lim_{s \to \infty} \psi_1(s) \geq \lim_{s \to \infty} \alpha_1(s) = \infty.
\]

Proof of (iii) \( \Leftarrow \): If \( V_1 \sim_{K_{\infty}} V_2 \), then \( V_1 \sim_{K_{\infty}} V_2 \) by (4.1). Also, we have trivially that \( V_2 \sim_{K_{\infty}} V_2 \). Since \( \sim_{K_{\infty}} \) is an equivalence relation by Lemma 4.4, it follows that \( V_1 \sim_{K_{\infty}} V_2 \), so the properties \( \{P1\}_{i=1}^3 \) hold as in (ii) \( \Leftarrow \). We have left to derive P4. If \( V_1 \sim_{K_{\infty}} V_2 \), then there exist \( \alpha_1, \alpha_2 \in K_{\infty} \) convex and concave, respectively, such that \( \alpha_1(V_2(x)^p) \leq V_1(x) \leq \alpha_2(V_2(x)^p) \) for all \( x \in D \). Hence, by the definition of \( \psi_p \) and \( \phi_p \), and P0, and by the monotonicity of \( \alpha_1 \) and \( \alpha_2 \), we have that, for all \( s \geq 0 \),

\[
\alpha_1(s) \leq \inf_{\{x \in D : V_2(x)^p \geq s\}} \alpha_1(V_2(x)^p) \leq \inf_{\{x \in D : V_2(x)^p \geq s\}} V_1(x) = \psi_p(s) \leq \phi_p(s) = \sup_{\{x \in D : V_2(x)^p \leq s\}} \alpha_2(V_2(x)^p) \leq \alpha_2(s).
\]
Now, since $\alpha_1, \alpha_2 \in \mathcal{K}_\infty$ are convex and concave, respectively, it follows by Lemma A.1 that $s^2/\alpha_1(s)$ and $\alpha_2(s)$ are in $\mathcal{O}(s)$ as $s \to \infty$. Knowing from (4.12) that $\alpha_1(s) \leq \psi_p(s) \leq \alpha_2(s)$ for all $s \geq 0$, we conclude that the functions $s^2/\psi_p(s)$ and $\phi_p(s)$ are also in $\mathcal{O}(s)$ as $s \to \infty$, which is the property P4. □

The following example shows ways in which the conditions of Theorem 4.6 might fail.

**Example 4.7.** (Illustration of Theorem 4.6). Let $V_2 : \mathbb{R}^2 \to \mathbb{R}_{\geq 0}$ be the distance to the set $\{(x_1, x_2) \in \mathbb{R}^2 : x_1 = 0\}$, i.e., $V_2(x_1, x_2) = |x_1|$. Consider the following cases:

- **$P2$ fails ($\psi$ is not positive definite):** Let $V_1(x_1, x_2) = |x_1|e^{-|x_2|}$ for $(x_1, x_2) \in \mathbb{R}^2$. Note that $V_1$ is not $\mathcal{K}$-dominated by $V_2$ because, given any $\alpha_1 \in \mathcal{K}$, for every $x_1 \in \mathbb{R}$ with $|x_1| > 0$ there exists $x_2 \in \mathbb{R}$ such that the inequality $\alpha_1(|x_1|) \leq |x_1|e^{-|x_2|}$ does not hold (just choose $x_2$ satisfying $|x_2| > \log \left(\frac{|x_1|}{\alpha_1(|x_1|)}\right)$). Thus, there must be some of the hypotheses on Theorem 4.6 that fail to be true. In this case, we observe that

$$\psi_1(s) = \inf_{\{(x_1, x_2) \in \mathbb{R}^2 : |x_1| \geq s\}} |x_1|e^{-|x_2|}$$

is identically 0 for all $s \geq 0$, so it is not positive definite as required in P2.

- **$P2$ fails ($\phi$ is not locally bounded):** Let $V_1(x_1, x_2) = |x_1|e^{|x_2|}$ for $(x_1, x_2) \in \mathbb{R}^2$. As above, one can show that $\alpha_2$ does not exist in the required class; in this case, the hypothesis P2 is not satisfied because $\phi_1$ is not locally bounded in $(0, \infty)$:

$$\phi_1(s) = \sup_{\{(x_1, x_2) \in \mathbb{R}^2 : |x_1| \leq s\}} |x_1|e^{|x_2|} = \infty, \quad \forall s > 0.$$

- **$P2$ fails ($\phi$ is not right continuous):** Let $V_1(x_1, x_2) = |x_1|^4 + |\sin(x_1x_2)|$ for $(x_1, x_2) \in \mathbb{R}^2$. For every $p > 0$, we have that

$$\phi_p(s) = \sup_{\{(x_1, x_2) \in \mathbb{R}^2 : |x_1|^p \leq s\}} |x_1|^4 + |\sin(x_1x_2)| \leq s^{4/p} + 1,$$

so $\phi_p$ is locally bounded in $\mathbb{R}_{\geq 0}$, and, again for every $p > 0$,

$$\psi_p(s) = \inf_{\{(x_1, x_2) \in \mathbb{R}^2 : |x_1|^p \geq s\}} |x_1|^4 + |\sin(x_1x_2)| \geq s^{4/p},$$

so $\psi_p$ is positive definite. However, $\phi_p$ is not right continuous at 0 because $\sup_{\{(x_1, x_2) \in \mathbb{R}^2 : |x_1|^p \leq s_0\}} |\sin(x_1x_2)| = 1$ for any $s_0 > 0$, so by Theorem 4.6 (ii), it follows that $V_1$ is not $\mathcal{K}$-dominated by $V_2$.

- **$P4$ fails (non-compliant asymptotic behavior):** Let $V_1(x_1, x_2) = |x_1|^4$ for $(x_1, x_2) \in \mathbb{R}^2$. Then P2 is satisfied and P3 also holds because $\lim_{s \to \infty} \psi_1(s) = \lim_{s \to \infty} s^4 = \infty$, so Theorem 4.6 (ii) implies that $V_1$ and $V_2$ are $\mathcal{K}_\infty$-proper with respect to each other. However, in this case $\phi_p(s) = \psi_p(s) = s^{4/p}$, which implies that $\phi_p$ is not in $\mathcal{O}(s)$ as $s \to \infty$ when $p \in (0, 4)$, and $s^4/\psi_p(s)$ is not in $\mathcal{O}(s)$ as $s \to \infty$ when $p > 4$. Thus P4 is satisfied only for $p = 4$, so Theorem 4.6 (iii) implies that only in this case $V_1$ and $V_2$ are $\mathcal{K}_\infty$-proper with respect to each other. Namely, for $p > 4$, one cannot choose a convex $\alpha_1 \in \mathcal{K}_\infty$ such that $\alpha_1(|x_1|^p) \leq |x_1|^4$ for all $x_1 \in \mathbb{R}$ and, if $p < 4$, one cannot choose a concave $\alpha_2 \in \mathcal{K}_\infty$ such that $|x_1|^4 \leq \alpha_2(|x_1|^p)$ for all $x_1 \in \mathbb{R}$. □

4.3. **Application to noise-to-state stability.** In this section we use the results of Sections 4.1 and 4.2 to study the noise-to-state stability properties of stochastic
differential equations of the form (2.1). Our first result provides a way to check whether a candidate function that satisfies a dissipation inequality of the type (3.3) is in fact a noise-dissipative Lyapunov function, a strong NSS-Lyapunov function in probability, or a pth moment NSS-Lyapunov function.

**Corollary 4.8.** (Establishing proper relations between pairs of functions through seminorms). Consider $V_1, V_2 : D \to \mathbb{R}_{\geq 0}$ such that their nullset is a subspace $\mathcal{U}$. Let $A, \tilde{A} \in \mathbb{R}^{m \times n}$ be such that $\mathcal{N}(A) = \mathcal{U} = \mathcal{N}(\tilde{A})$. Assume that $V_1$ and $V_2$ satisfy \{P\}_i^{3} with respect to $\| \cdot \|_A$ and $\| \cdot \|_{\tilde{A}}$, respectively. Then, for any $q > 0$,

$$V_1 \sim^{\mathcal{K}-} V_2, \quad V_1 \sim^{\mathcal{K}^*} \| \cdot \|_A, \quad V_2 \sim^{\mathcal{K}^*} \| \cdot \|_{\tilde{A}} \text{ in } D.$$

If, in addition, $V_1$ and $V_2$ satisfy P4 with respect to $\| \cdot \|_A$ and $\| \cdot \|_{\tilde{A}}$, respectively, for some $p > 0$, then

$$V_1 \sim^{\mathcal{K}^*} V_2, \quad V_1 \sim^{\mathcal{K}^*} \| \cdot \|_A, \quad V_2 \sim^{\mathcal{K}^*} \| \cdot \|_{\tilde{A}} \text{ in } D.$$

**Proof.** The statements follow from the characterizations in Theorem 4.6 (ii) and (iii), and from the fact that the relations $\sim^{\mathcal{K}^*}$ and $\sim^{\mathcal{K}^*}$ are equivalence relations as shown in Lemma 4.4. That is, under the hypothesis P0,

$V_1$ satisfies \{P\}_i^{3} w/ respect to $\| \cdot \|_A \Leftrightarrow V_1 \sim^{\mathcal{K}^*} \| \cdot \|_A \text{ in } D \} \Rightarrow V_1 \sim^{\mathcal{K}^*} V_2 \text{ in } D,$

$V_2$ satisfies \{P\}_i^{3} w/ respect to $\| \cdot \|_{\tilde{A}} \Leftrightarrow V_2 \sim^{\mathcal{K}^*} \| \cdot \|_{\tilde{A}} \text{ in } D \} \Rightarrow V_1 \sim^{\mathcal{K}^*} V_2 \text{ in } D.$

Note that, by Lemma 4.3 and (4.1), the equivalences

$$\| \cdot \|_A \sim^{\mathcal{K}^*} \| \cdot \|_A^q \text{ in } D, \quad \| \cdot \|_A \sim^{\mathcal{K}^*} \| \cdot \|_A^p \text{ in } D,$$

hold for any $p, q > 0$ and any matrices $A, \tilde{A} \in \mathbb{R}^{m \times n}$ with $\mathcal{N}(A) = \mathcal{N}(\tilde{A})$. \qed

We next build on this result to provide an alternative formulation of Corollary 3.9. To do so, we employ the observation made in Remark 4.2 about the possibility of interpreting the candidate functions as defined on a constrained domain of an extended Euclidean space.

**Corollary 4.9.** (The existence of a pthNSS-Lyapunov function implies pth moment NSS –revisited). Under Assumption 2.1, let $V \in \mathcal{C}^2(\mathbb{R}^{n}; \mathbb{R}_{\geq 0})$, $W \in \mathcal{C}(\mathbb{R}^{m}; \mathbb{R}_{\geq 0})$ and $\sigma \in \mathcal{K}$ be such that the dissipation inequality (3.4) holds. Let $R : \mathbb{R}^{n} \to \mathbb{R}^{(m-n)}$, with $m \geq n$, $D \subset \mathbb{R}^{m}$, $\tilde{V} \in \mathcal{C}^2(D; \mathbb{R}_{\geq 0})$ and $\tilde{W} \in \mathcal{C}(D; \mathbb{R}_{\geq 0})$ be such that, for $i(x) = [\hat{x}^T, R(x)^T]^T$, one has

$$D = i(\mathbb{R}^{n}), \quad V = \tilde{V} \circ i, \text{ and } \quad W = \tilde{W} \circ i.$$

Let $A = \text{diag}(A_1, A_2)$ and $\tilde{A} = \text{diag}(\tilde{A}_1, \tilde{A}_2)$ be block-diagonal matrices, with $A_1, \tilde{A}_1 \in \mathbb{R}^{n \times n}$ and $A_2, \tilde{A}_2 \in \mathbb{R}^{(m-n) \times (m-n)}$, such that $\mathcal{N}(A) = \mathcal{N}(\tilde{A})$ and

$$R(x) \leq \kappa \| x \|_{A_1}^2$$

(4.13)
for some $\kappa > 0$, for all $x \in \mathbb{R}^n$. Assume that $\hat{V}$ and $\hat{W}$ satisfy the properties $\{P_i\}_{i=0}^4$ with respect to $\|\cdot\|_A$ and $\|\cdot\|_{A_1}$, respectively, for some $p > 0$. Then the system (2.1) is NSS in probability and in $p$th moment with respect to $\mathcal{N}(A_1)$.

**Proof.** By Corollary 4.8, we have that

\begin{equation}
\hat{V} \sim^\kappa \hat{W}, \quad \text{and} \quad \hat{V} \sim^\kappa \|\cdot\|_{\text{diag}(A_1, A_2)}^p \quad \text{in} \quad \mathcal{D}.
\end{equation}

As explained in Remark 4.2, the first relation implies that $V \sim^\kappa W$ in $\mathbb{R}^n$. This, together with the fact that (3.4) holds, implies that $V$ is a noise-dissipative Lyapunov function for (2.1). Also, setting $\hat{x} = i(x)$ and using (4.13), we obtain that

\[ \|x\|_{A_1}^2 \leq \|\hat{x}\|_{\text{diag}(A_1, A_2)}^2 = \|x\|_{A_1}^2 + \|R(x)\|_{A_2}^2 \leq (1 + \kappa)\|x\|_{A_1}^2, \]

so, in particular, $\|[\cdot, R(\cdot)]\|_{\text{diag}(A_1, A_2)}^p \sim \|\cdot\|_{A_1}^p$ in $\mathbb{R}^n$. Now, from the second relation in (4.14), by Remark 4.2, it follows that $\hat{V} \circ i \sim^\kappa \|[\cdot, R(\cdot)]\|_{\text{diag}(A_1, A_2)}^p \sim \|\cdot\|_{A_1}^p$ in $\mathbb{R}^n$. Thus, using (4.1) and Lemma 4.4, we conclude that $V \sim^\kappa \|\cdot\|_{A_1}^p$ in $\mathbb{R}^n$. In addition, the Euclidean distance to the set $\mathcal{N}(A_1)$ is equivalent to $\|\cdot\|_{A_1}$, i.e., $\|x\|_{\mathcal{N}(A_1)} \sim \|\cdot\|_{A_1}$. This can be justified as follows: choose $B \in \mathbb{R}^{n \times k}$, with $k = \dim(\mathcal{N}(A_1))$, such that the columns of $B$ form an orthonormal basis of $\mathcal{N}(A_1)$. Then,

\begin{equation}
|x|_{\mathcal{N}(A_1)} = \|\text{I} - BB^T\|_2 = \|x\|_2 - \|BB^T\|_2 \sim \|\cdot\|_{A_1},
\end{equation}

where the last relation follows from Lemma 4.3 because $\mathcal{N}(\text{I} - BB^T) = \mathcal{N}(A_1)$. Summarizing, $V \sim^\kappa \|\cdot\|_{A_1}^p$ and $\|\cdot\|_{A_1}^p \sim \|x|_{\mathcal{N}(A_1)}^p$ in $\mathbb{R}^n$ (because the $p$th power is irrelevant for the property $\sim$). As a consequence,

\begin{equation}
V \sim^\kappa \|x|_{\mathcal{N}(A_1)}^p \quad \text{in} \quad \mathbb{R}^n,
\end{equation}

which implies condition (3.17) with convex $\alpha_1 \in \mathcal{K}_\infty$, concave $\alpha_2 \in \mathcal{K}_\infty$, and $\mathcal{U} = \mathcal{N}(A_1)$. Therefore, $V$ is a $p$th moment NSS-Lyapunov function with respect to the set $\mathcal{N}(A_1)$, and the result follows from Corollary 3.9. \Box

**5. Conclusions.** We have studied the stability properties of SDEs subject to persistent noise (including the case of additive noise). We have generalized the concept of noise-dissipative Lyapunov function and introduced the concepts of strong NSS-Lyapunov function in probability and $p$th moment NSS-Lyapunov function, both with respect to a closed set. We have shown that noise-dissipative Lyapunov functions have NSS dynamics and established that the existence of an NSS-Lyapunov function, of either type, with respect to a closed set, implies the corresponding NSS property of the system with respect to the set. In particular, $p$th moment NSS with respect to a set provides a bound, at each time, for the $p$th power of the distance from the state to the set, and this bound is the sum of an increasing function of the size of the noise covariance and a decaying effect of the initial conditions. This bound can be achieved regardless of the possibility that inside the set some combination of the states accumulates the variance of the noise. This is a meaningful stability property for the aforementioned class of systems because the presence of persistent noise makes it impossible to establish in general a stochastic notion of asymptotic stability for the set of equilibria of the underlying differential equation. We have also studied in depth the inequalities between pairs of functions that appear in the various notions of Lyapunov functions mentioned above. We have shown that these inequalities define equivalence
relations and have developed a complete characterization of the properties that two functions must satisfy to be related by them. Finally, building on this characterization, we have provided an alternative statement of our stochastic stability results. Future work will include the study of the effect of delays and impulsive right-hand sides in the class of SDEs considered in this paper.

References.
[19] L. Praly and Y. Wang, Stabilization in spite of matched unmodeled dynamics


Appendix. The next result is used in the proof of Theorem 4.6.

Lemma A.1. (Existence of bounding convex and concave functions in $K_\infty$). Let $\alpha$ be a class $K_\infty$ function. Then the following are equivalent:

(i) There exist $s_0 \geq 0$ and $\alpha_1, \alpha_2 \in K_\infty$, convex and concave, respectively, such that $\alpha_1(s) \leq \alpha(s) \leq \alpha_2(s)$ for all $s \geq s_0$, and

(ii) $\alpha(s), s^2/\alpha(s)$ are in $O(s)$ as $s \to \infty$.

Proof. The implication (i) $\Rightarrow$ (ii) follows because, for any $s \geq s_0 > 0$,

$$\frac{\alpha_1(s)}{s_0} s \leq \alpha_1(s) \leq \alpha(s) \leq \alpha_2(s) \leq \frac{\alpha_2(s_0)}{s_0} s,$$

by convexity and concavity, respectively, where $\alpha_1(s_0), \alpha_2(s_0) > 0$.

To show (ii) $\Rightarrow$ (i), we proceed to construct $\alpha_1, \alpha_2$ as in the statement using the correspondence between functions, graphs and epigraphs (or hypographs). Let $\alpha_1 : \mathbb{R}_{\geq 0} \to \mathbb{R}$ be the function whose epigraph is the convex hull of the epigraph of $\alpha$, i.e., $\text{epi} \alpha_1 \equiv \text{conv}(\text{epi} \alpha)$. Thus, $\alpha_1$ is convex, nondecreasing, and $0 \leq \alpha_1(s) \leq \alpha(s)$ for all $s \geq 0$ because $\mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \supset \text{epi} \alpha_1 = \text{conv}(\text{epi} \alpha) \supset \text{epi} \alpha$. Moreover, $\alpha_1$ is continuous in $(0, \infty)$ by convexity [20, Th. 10.4], and is also continuous at 0 by the sandwich theorem [11, p. 107] because $\alpha \in K_\infty$. To show that $\alpha_1 \in K_\infty$, we have to check that it is unbounded, positive definite in $\mathbb{R}_{\geq 0}$, and strictly increasing. First, since $s^2/\alpha(s) \in O(s)$ as $s \to \infty$, there exist constants $c_1, s_0 > 0$ such that $\alpha(s) \geq c_1 s$ for all $s > s_0$. Now, define $g_1(s) \equiv \alpha(s)$ if $s \leq s_0$ and $g_1(s) \equiv c_1 s$ for $s > s_0$. Since $g_1(s) \geq c_1 s$, we have $\alpha(s) \geq c_1 s$ for all $s \geq s_0$. Therefore, $\alpha_1(s) \geq \alpha_1(s_0) + (s - s_0) c_1$. This shows that $\alpha_1$ is strictly increasing. Since $\alpha_1(s) \to \infty$ as $s \to \infty$, we have $\alpha_1 \in K_\infty$. The proof is complete.
if \( s > s_0 \), and \( g_2(s) \triangleq -c_1s_0 + c_1s \) for all \( s \geq 0 \), so that \( g_2 \leq g_1 \leq \alpha \). Then, \( \text{epi} \alpha_1 = \text{conv}(\text{epi} \alpha) \subseteq \text{conv}(\text{epi} g_1) \subseteq \text{epi} g_2 \), because \( \text{epi} g_2 \) is convex, and thus \( \alpha_1 \) is unbounded. Also, since \( \text{conv}(\text{epi} g_1) \cap \mathbb{R}_{\geq 0} \times \{0\} = \{(0, 0)\} \), it follows that \( \alpha_1 \) is positive definite. To show that \( \alpha_1 \) is strictly increasing, we use two facts: since \( \alpha_1 \) is convex, we know that the set in which \( \alpha_1 \) is allowed to be constant must be of the form \([0, b]\) for some \( b > 0 \); on the other hand, since \( \alpha_1 \) is positive definite, it is nonconstant in any neighborhood of 0. As a result, \( \alpha_1 \) is nonconstant in any subset of its domain, so it is strictly increasing.

Next, let \( \alpha_2 : \mathbb{R}_{\geq 0} \to \mathbb{R} \) be the function whose hypograph is the convex hull of the hypograph of \( \alpha \), i.e., \( \text{hyp} \alpha_2 \triangleq \text{conv}(\text{hyp} \alpha) \). The function \( \alpha_2 \) is well-defined because \( \alpha(s) \in \mathcal{O}(s) \) as \( s \to \infty \), i.e., there exist constants \( c_2, s_0 > 0 \) such that \( \alpha(s) \leq c_2s \) for all \( s > s_0 \), so if we define \( g(s) \triangleq c_2s_0 + c_2s \) for all \( s \geq 0 \), then \( \text{hyp} \alpha_2 = \text{conv}(\text{hyp} \alpha) \subseteq \text{hyp} g \), because \( \text{hyp} g \) is convex, and thus \( \alpha_2(s) \leq g(s) \). Also, by construction, \( \alpha_2 \) is concave, nondecreasing, and \( \alpha_2 \geq \alpha \) because \( \text{hyp} \alpha_2 \supseteq \text{hyp} \alpha \), which also implies that \( \alpha_2 \) is unbounded. Moreover, \( \alpha_2 \) is continuous in \((0, \infty)\) by concavity [20, Th. 10.4], and is also continuous at 0 because the possibility of an infinite jump is excluded by the fact that \( \alpha_2 \leq g \). To show that \( \alpha_2 \in \mathcal{K}_\infty \), we have to check that it is positive definite in \( \mathbb{R}_{\geq 0} \) and strictly increasing. Note that \( \alpha_2 \) is positive definite because \( \alpha_2(0) = 0 \) and \( \alpha_2 \geq \alpha \). To show that \( \alpha_2 \) is strictly increasing, we reason by contraction. Assume that \( \alpha_2 \) is constant in some closed interval of the form \([s_1, s_2]\), for some \( s_2 > s_1 \geq 0 \). Then, as \( \alpha_2 \) is concave, we conclude that it is nonincreasing in \((s_2, \infty)\). Now, since \( \alpha_2 \) is continuous, we reach the contradiction that \( \lim_{s \to \infty} \alpha(s) \leq \lim_{s \to \infty} \alpha_2(s) \leq \alpha_2(s_1) < \infty \). Hence, \( \alpha_2 \) is strictly increasing. \( \blacksquare \)