

Nonlinear Cross Gramians

Tudor C. Ionescu and Jacquélien M. A. Scherpen

Abstract We study the notion of cross Gramians for nonlinear gradient systems, using the characterization in terms of prolongation and gradient extension associated to the system. The cross Gramian is given for the variational system associated to the original nonlinear gradient system. We obtain linearization results that correspond to the notion of a cross Gramian for symmetric linear systems. Furthermore, first steps towards relations with the singular value functions of the nonlinear Hankel operator are studied and yield promising results.

1 Introduction

In this paper, we give an extension of the cross Gramian notion for nonlinear gradient systems. The gradient systems are an important class of nonlinear systems, endowed with a pseudo-Riemannian metric on the state-space manifold, such that the drift is a gradient vectorfield with respect to this metric and a potential function and the input vectorfields are gradient with respect to the same metric and output, see e.g. [3, 15] and references therein. Examples of gradient systems include nonlinear electrical circuits and certain dissipative systems. The linear counterpart is a symmetric system. With respect to model reduction, for linear systems it is showed in [1, 4, 14] that exploiting the symmetry, model reduction becomes more efficient. This is based on the notion of cross Gramian, that is the solution of a Sylvester equation, which can be solved in an efficient way. The cross Gramian for a symmetric system contains information about both controllability and observability at the same time and moreover the squared cross Gramian is the product of the controllability

Tudor C. Ionescu, e-mail: t.c.ionescu@rug.nl
Jacquélien M. A. Scherpen, e-mail: j.m.a.scherpen@rug.nl

University of Groningen, Faculty of Mathematics and Natural Sciences ITM, Nijenborgh 4,
9747AG, Groningen, The Netherlands

and observability Gramians. Then the Hankel singular values are the eigenvalues of the cross Gramian. Moreover, the cross Gramian can be obtained using only one of the Gramians of the system and the metric.

For nonlinear systems the problem is more complicated and not yet tackled in the literature. The notion of symmetry for a nonlinear system is now best studied by considering nonlinear gradient systems. We use the associated prolongation and gradient extension and the results in [3]. A nonlinear system is gradient if the two latter systems have the same input-output behavior. Using this property and its consequences, we give the definition of the cross Gramian for the variational system (which is a gradient system, too) as the nonlinear, non-trivial extension of the concept of the cross Gramian for linear systems. Furthermore, we give a nonlinear counterpart of the Sylvester equation. Using the cross Gramian and the theory of Hankel singular values as in [6, 11], first steps towards proving that the squared eigenvalues of the nonlinear cross Gramian are directly related to the Hankel singular values of the system, are set. In this case, instead of balancing, only solving a nonlinear Sylvester equation, a metric and an eigenvalue decomposition suffice for obtaining the Hankel singular values of the gradient system.

The paper is outlined as follows. In Section 2 we give an overview of the cross Gramian technique for linear systems. To show the line of thinking in the nonlinear case, in Section 3, we give a review of the definitions of the prolongation and gradient extension and the property of a nonlinear system being gradient itself, this being a natural extension of the linear symmetric system notion, to the nonlinear case. In Section 4, we analyze some linearization results which motivate the reasoning in Section 5, where the definition of the nonlinear Gramian is presented and the conjecture about the relation for singular value functions is stated. Finally an example is given in Section 6 and in Section 7 some conclusions end this paper.

A nonlinear system is defined here as:

$$\begin{cases} \dot{x} = f(x) + g(x)u \\ y = h(x) \end{cases}, \quad (1)$$

where $x \in \mathcal{M}$ is the state vector, $u \in \mathbb{R}^m$ is the vector of inputs and $y \in \mathbb{R}^p$ is the output. \mathcal{M} is a smooth manifold, of dimension n . We make the following assumptions:

Assumption 1 $f(x), g(x), h(x)$ are smooth vectorfields;

Assumption 2 The system is square, i.e. $m = p$;

Assumption 3 x_0 is an asymptotically stable equilibrium point of the system and $h(x_0) = 0$;

Assumption 4 System (1) asymptotically reachable from x_0 (i.e. for any x , there exists an input u and $t \geq 0$, such that $x = \phi(t, 0, x_0, u)$, with ϕ being the trajectory obtained by integrating the first equation in (1)).

Assumption 5 System (1) is zero-state observable (i.e. if $u(t) = 0$, $y(t) = 0$ then $x(t) = 0$).

Assumptions 4 and 5 are related to the minimality of the system, see [12].

Notation:

Let \mathcal{M} be a smooth manifold and $V(x)$ a smooth vectorfield, $x \in \mathcal{M}$. Then we denote by $\text{grad}_G V$ the gradient of $V(x)$ on the manifold \mathcal{M} endowed with the pseudo-Riemannian metric G . In local coordinates $\text{grad}_G V = -G^{-1}(x) \frac{\partial V(x)}{\partial x}$ (see [15] for details). $\frac{\partial V(x)}{\partial x}$ means the row vector $\left[\frac{\partial V(x)}{\partial x_1} \dots \frac{\partial V(x)}{\partial x_n} \right]$. \mathbb{R} is the set of real numbers.

2 Linear Systems Case

If the system (1) is linear, then it can be written as:

$$\begin{cases} \dot{x} = Ax + Bu \\ y = Cx \end{cases}, \quad (2)$$

where $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times m}$, $C \in \mathbb{R}^{p \times n}$ are constant matrices. In this case, Assumption 1 is automatically satisfied. We consider system (2) satisfying Assumptions 2-5. Assumptions 4 and 5 are equivalent to the minimality of the system (see e.g. [16] for more details). A linear system has a corresponding unique dual system defined as:

$$\begin{cases} \dot{z} = A^T z + C^T u_d \\ y_d = B^T z \end{cases}. \quad (3)$$

Because (1) is controllable and observable, and these properties are dual to each other (i.e. if the pair (A, B) is controllable, then (B^T, A^T) is observable), it follows immediately that the dual system (3) is controllable and observable, i.e. minimal, too.

The definition of the cross Gramian for a linear square system is:

Definition 1. [14] Let (2) be a square system. Then the cross Gramian X is defined as the solution of the Sylvester equation:

$$AX + XA + BC = 0. \quad (4)$$

If the system is asymptotically stable, then the cross Gramian can be equivalently defined as: $X = \int_0^\infty e^{At} BC e^{At} dt$.

Another important definition is the one of the Hankel operator associated to the linear system (2):

$$\mathcal{H}(u) = \int_{-\infty}^0 H(t - \tau) u(-\tau) d\tau \quad (5)$$

where $t > 0$ and H is the impulse response of the system (2). The singular values of the Hankel operator are fundamental for the balanced truncation model order reduc-

tion. Each singular value represents a measure for the importance of each state component in the output response of system (2) to a certain input (see e.g. [2] for more details). The cross Gramian possesses some interesting properties being related to the above defined Hankel operator and the Hankel singular values of a linear square system.

Theorem 1. [14] *For square linear systems the non-zero eigenvalues of the cross Gramian X are the non-zero eigenvalues of the Hankel operator associated to the system.*

However, the singular value problem is different, that is the singular values of the cross Gramian are not the Hankel singular values of the system. Still, there is a relation of majorization between the two as shown below.

Theorem 2. [14] *For a square linear system, the following relations hold: $\sum_{i=1}^k \sigma_i \geq \sum_{i=1}^k \pi_i$ and $\sum_{i=k+1}^n \sigma_i \leq \sum_{i=k+1}^n \pi_i$, where σ_i are the Hankel singular values, π_i are the singular values of X , and k is the index for which σ_k is much larger than σ_{k+1} .*

For symmetric systems, the cross Gramian X has more attractive properties, useful for model reduction.

First we give the definition of a symmetric linear system:

Definition 2. [1, 4, 14] A square linear, system $G(s) = C(sI - A)^{-1}B$, with the state-space realization (2) is called symmetric if $G(s) = G^T(s)$.

Proposition 1. *Assume that system (2) satisfies assumptions 2-5. Then system (2) is symmetric if and only if there exists an invertible symmetric matrix T such that $A^T T = TA$, $C^T = TB$, i.e. the system and its dual are input-output (externally) equivalent.*

In, for instance [1, 14], model reduction based on the balancing procedure, for this type of systems is considered. The symmetry property is exploited, making the procedure more efficient. Basically, the Sylvester equation from Definition 1 is solved and the cross Gramian is obtained. It will directly provide the Hankel singular values of the system. We refer to the results presented in [14, 1, 4], which are summarized in the sequel.

Defining the controllability Gramian as W and the observability Gramian as M , they are the solutions of the following Lyapunov equations, respectively:

$$AW + WA^T + BB^T = 0 \quad (6)$$

$$A^T M + MA + C^T C = 0. \quad (7)$$

The following theorem summarizes the properties of X in relation with W and M .

Theorem 3. [14, 4] *Let (2) be a square asymptotically stable symmetric system in the sense of Definition 2. If X is the solution of (4) then the following relations are equivalent:*

1. $X^2 = WM > 0$;
2. *If T is the symmetry transformation as in Proposition 1, then $X = WT = T^{-1}M$;*
3. *The Hankel singular values of (2) are the absolute values of the eigenvalues of X .*

For symmetric systems, when compared to the classical balancing procedure, there are two advantages: the first is that instead of solving two Lyapunov equations, whose computational complexity is known to be a drawback, only one Sylvester equation is solved. The second advantage consists of avoiding in this way the balancing procedure. Since the Hankel singular values satisfy $\sigma_i = \sqrt{\lambda_i}$, $\lambda_i \in \lambda(WM)$, $i = 1, \dots, n$, the problem of finding them turns into an eigenvalue problem of the cross Gramian X .

Remark 1 *There exists a relation between the controllability and observability operators, and the cross Gramian. Define by $x = \mathcal{C}(u) = \int_0^\infty e^{At}Bu(t)dt$, the controllability operator and by $y = \mathcal{O}(x) = Ce^{At}x$ the observability operator of the system (A, B, C) . Then, by the definition of the cross Gramian, we have: $Xx = \mathcal{C}\mathcal{O}(x)$. It can be proven that, under minimality and symmetry assumptions as in the definitions presented here, the eigenvalues of the $\mathcal{C}\mathcal{O}\mathcal{C}\mathcal{O}$ operator are the squared Hankel singular values of the system, i.e. the eigenvalues of $\mathcal{H}^*\mathcal{H}$.*

3 Review of Gradient Systems

The nonlinear extension of the notion of symmetric systems is the gradient systems. The property of a system being gradient is described in terms of necessary and sufficient conditions satisfied by the prolongation (variational) system and the gradient extension associated with (1). We will give a brief overview of the results in [3, 15].

Definition 3. [3, 15] A nonlinear system (1) is called a gradient system if:

1. There exists a pseudo-Riemannian metric G , on the manifold \mathcal{M} , given as

$$\sum_{i,j=1}^m g_{ij}(x)dx_i \otimes dx_j, \text{ with } g_{ij}(x) = g_{ji}(x) \text{ smooth functions of } x, \text{ and the matrix}$$

$$G(x) = [g_{ij}(x)]_{i,j=1\dots n} \text{ invertible, for all } x \text{ and } \otimes \text{ representing the tensor product.}$$

2. There exists a smooth potential function $V : \mathcal{M} \rightarrow \mathbb{R}$,

such that the system (1) can be written as:

$$\begin{cases} \dot{x} = \text{grad}_G V(x) - \sum_{i=1}^m u_i \text{grad}_G h_i(x), & x \in \mathbb{R}^n \\ y_i = h_i(x), & i = 1, \dots, m \end{cases} \quad (8)$$

In local coordinates $x = [x_1 \ x_2 \ \dots \ x_n]^T \in \mathcal{M}$, the system can be written as:

$$\begin{cases} \dot{x} = -G^{-1}(x) \frac{\partial^T V}{\partial x}(x) + G^{-1}(x) \frac{\partial^T h}{\partial x}(x)u \\ y = h(x) \end{cases} \quad (9)$$

Next, we present the definition of the prolonged system associated with (1).

Definition 4. [3] The prolongation Σ_p of (1) is defined by:

$$\begin{cases} \dot{x} = f(x) + g(x)u \\ \dot{v} = \frac{\partial f(x)}{\partial x}v + \sum_{j=1}^m u_j \frac{\partial g_j(x)}{\partial x}v + g(x)u_p \\ y = h(x), y_p = \frac{\partial h(x)}{\partial x}v \end{cases} \quad (10)$$

where $v \in T\mathcal{M}$, the tangent bundle of the manifold \mathcal{M} .

3.1 The Riemannian Metric on $T^*\mathcal{M}$

Since a canonical pseudo-Riemannian structure on the cotangent bundle $T^*\mathcal{M}$ of the manifold \mathcal{M} does not exist, a pseudo-Riemannian metric cannot be defined directly. In this case a torsion-free affine connection defined on the manifold \mathcal{M} and its Riemannian extension G^C to $T^*\mathcal{M}$ are used.

Definition 5. An affine connection on a manifold \mathcal{M} is defined as an assignment $\nabla : (X, Y) \rightarrow \nabla_X Y$, where X, Y and $\nabla_X Y$ are vectorfields on \mathcal{M} , satisfying the following properties: it is \mathbb{R} -bilinear, $\nabla_{fX} Y = f\nabla_X Y$ and $\nabla_X(fY) = f\nabla_X Y + X(f)Y$, for every $f \in C^\infty(\mathcal{M})$.

Let X and Y be any two vectorfields on \mathcal{M} . Their symmetric product is given as: $\langle X : Y \rangle = \nabla_X Y + \nabla_Y X$. We introduce the construction that associates to each vectorfield X a function V^X on $T^*\mathcal{M}$, given by $V^X(x, p) = \langle p, X(x) \rangle$, $x \in \mathcal{M}$, $p \in T^*\mathcal{M}$.

If ∇ is a torsion-free affine connection (see [3] and references therein for more details) then it defines a pseudo-Riemannian metric G^C as a unique (0,2)-tensor on $T^*\mathcal{M}$ which satisfies:

$$G^C(X, Y) = -V^{\langle X : Y \rangle} \quad (11)$$

Now the gradient vectorfield associated with the function $V^X \in C^\infty(T^*\mathcal{M})$, X vectorfield on \mathcal{M} , can be expressed locally as:

$$\text{grad}_{G^C} V^X = X_i \frac{\partial}{\partial x_i} + p_i \left(\frac{\partial X_i}{\partial x_j} + 2\Gamma_{jk}^a X_k \right) \frac{\partial}{\partial p_j}, \quad (12)$$

where X is a vectorfield on \mathcal{M} , $i, j, k = 1, \dots, n$ and Γ_{jk}^i represent the Christoffel symbols of the affine connection ∇ (relation (2.8) in [3]).

For our purpose, we assume that G^C is properly defined ([3]) and we will use the local expression from (12) to express the gradient extension of (1), comprising all the terms $2p_i \Gamma_{jk}^a X_k \frac{\partial}{\partial p_j}$ in a function \mathcal{F} .

3.2 The Gradient Extension of a Nonlinear System

Definition 6. The gradient extension of (1) is defined by:

$$\begin{cases} \dot{x} = f(x) + g(x)u \\ \dot{p} = \frac{\partial^T (f(x) + g(x)u)}{\partial x} p + \mathcal{F}(g_{ij}(x), \frac{\partial g_{ij}(x)}{\partial x_k}, f_k(x), u, g(x), p) + \frac{\partial h(x)}{\partial x} u_g, \\ y = h(x), y_g = g^T(x)p, i, j, k = 1, \dots, n. \end{cases} \quad (13)$$

Remark 2 Notice that for the linear system (2) the prolongation is the system itself written twice and the gradient extension contains the system itself and the dual of the prolonged variable part, yielding, respectively:

$$\begin{cases} \dot{x} = Ax + Bu \\ \dot{v} = Av + Bu_p \\ y = Cx, y_p = Cv \end{cases}, \begin{cases} \dot{x} = Ax + Bu \\ \dot{p} = A^T p + C^T u_g \\ y = Cx, y_g = B^T p \end{cases}. \quad (14)$$

Remark 3 According to [3, Corollaries 3.3, 3.6] (1) is zero-state observable if and only the prolonged system is zero-state observable and the zero-state observability of (1) implies the zero-state observability of the gradient extension.

The main result, useful for our purpose, is:

Theorem 4. [3, Theorem 5.4, Corollary 4.4] Let (1) be as in Assumption 4. Assume that there exists a torsion-free affine connection on \mathcal{M} with which the system is compatible, and that the system is observable with its observability distribution having constant dimension. Then, system (1) is a gradient control system, as in Definition 3, if and only if the prolonged system Σ_p and the gradient extension Σ_g have the same input-output behavior.

Remark 4 In the linear systems case, this result becomes a property between the system itself and its dual counterpart, which immediately leads to the definition of symmetric systems. The metric is given by the matrix T , showing that a linear symmetric system is a particular case (linear version) of the gradient system.

Lemma 1. [3, Lemmas 5.5, 5.6] If (1) is a gradient control system, then there exists a diffeomorphism $\phi(x, v) = (x, G(x)v)$, such that $(x, p) = \phi(x, G(x)v)$, where v and p satisfy (14), and $G(x)$ is the matrix associated to the metric.

Remark 5 For linear systems this means, indeed that $p = Tv$.

4 Linearization Results

For (1) satisfying Assumptions 1 and 3 we define the observability function ([9])

$$L_o(x) = \frac{1}{2} \int_0^\infty \|y(t)\|_{L_2}^2 dt, \quad x(0) = x, \quad x(\infty) = x_0 \quad (15)$$

and the controllability function ([9])

$$L_c(x) = \min_{u \in L_2^-, x(0)=x, x(-\infty)=x_0} \frac{1}{2} \int_{-\infty}^0 \|u(t)\|_{L_2}^2 dt \quad (16)$$

If the system satisfies Assumption 4 as well, then $L_c(x)$ exists, is finite, $L_c(x) > 0$, $L_c(x_0) = 0$ and satisfies the Hamilton-Jacobi equation ([9]):

$$\frac{\partial L_c}{\partial x} f(x) + \frac{1}{2} \frac{\partial L_c}{\partial x} g(x) g^T(x) \frac{\partial^T L_c}{\partial x} = 0 \quad (17)$$

such that $-\left(f(x) + g(x)g^T(x) \frac{\partial L_c(x)}{\partial x}\right)$ is asymptotically stable. If the system also satisfies Assumption 5, then $L_o(x)$ exists, is finite, $L_o(x) > 0$, $L_o(x_0) = 0$, and satisfies the nonlinear Lyapunov equation ([9]):

$$\frac{\partial L_o}{\partial x} f(x) + \frac{1}{2} h^T(x) h(x) = 0. \quad (18)$$

Suppose $x_0, u = 0$ is an equilibrium point and assume that $h(x_0) = 0$. Then $-G^{-1}(x_0) \frac{\partial^T V}{\partial x}(x_0) = 0$. Taking Taylor series expansion in system (8), we can write:

$$\dot{x} = G^{-1}(x_0) \frac{\partial^2 V}{\partial x^2}(x_0)(x - x_0) + \left[\sum_{i,j=1}^n \frac{\partial g_{ij}}{\partial x_i}(x_0) \frac{\partial V}{\partial x_j}(x_0) \right]_{i,j=1\dots n} (x - x_0) + \dots \quad (19)$$

Since $\frac{\partial V}{\partial x_j}(x_0) = 0$, $j = 1, \dots, n$, then the linearization of the gradient system (8) yields:

$$\begin{cases} \dot{\bar{x}} = -G^{-1}(x_0) \frac{\partial^2 V}{\partial x^2}(x_0) \bar{x} + G^{-1}(x_0) \frac{\partial^T h}{\partial x}(x_0) u \\ \bar{y} = \frac{\partial h}{\partial x}(x_0) \bar{x} \end{cases} \quad (20)$$

Lemma 2. The system (20) is a gradient (symmetric) system with the metric $T = G(x_0)$.

Proof. Denote $G = G(x_0)$, $Q = \frac{\partial^2 V}{\partial x^2}(x_0)$. Since V is smooth, Q is symmetric. G , by definition is symmetric and invertible. Then:

$$\begin{aligned} H(s) &= C(sI + G^{-1}Q)^{-1}G^{-1}C^T = C(sG^{-1}G + G^{-1}Q)^{-1}G^{-1}C^T \\ &= C[G^{-1}(sI + QG^{-1})G]^{-1}G^{-1}C = CG^{-1}(sI + QG^{-1})C^T = H^T(s). \end{aligned} \quad (21)$$

□

Let W and M be the controllability and the observability Gramians of (20), respectively, and assume $W > 0$, $M > 0$, i.e. (20) is controllable and observable. Then:

$$M = \frac{\partial^2 L_o}{\partial x^2}(x_0), \quad W^{-1} = \frac{\partial^2 L_c}{\partial x^2}(x_0). \quad (22)$$

The asymptotic reachability of the nonlinear system implies its accessibility and this implies the controllability of the linear system, see [12]. Since the linearized system is assumed symmetric, controllability implies observability, and this implies the local zero-state observability of the nonlinear system. So, locally there exists a duality of the controllability and observability property, which motivates the search for a cross Gramian for the nonlinear gradient system.

The linearized system is gradient and then, according to Theorem 3, statement 2, we have that near x_0 :

$$\left(\frac{\partial^2 L_o}{\partial x^2}(x) \right)^{-1} G(x) = G^{-1}(x) \frac{\partial^2 L_c}{\partial x^2}(x). \quad (23)$$

Remark 6 Given a system (1), the linearization of the prolonged system Σ_p around $x_0, v = 0, u = u_p = 0$ and of the gradient extension Σ_g around $x_0, p = 0, u = u_g = 0$, respectively, we obtain the linear systems (14). If the system is symmetric then $p = Tv$, $G(x_0) = T$.

Since the duality in properties takes place between the v part and the p part of the two systems, we are going to extract these parts from the nonlinear system and study them.

4.1 The Isomorphic Case

Another case related to linearization is that when the system is equivalent to a linear system, as treated in [15]. This means that there exists a coordinate transformation $x' = \eta(x)$, such that in the new coordinates, the system is described by a linear state-space realization. If the equivalent linear system is a gradient system, as well, and the transformation η is an isometry (see e.g. [8]), then the gradient system is said to be isomorphic to the linear symmetric system. Then, the linear idea of cross Gramian can be extended to the nonlinear gradient system via the diffeomorphism η and the isometry relation, as follows.

Denote by $\bar{L}_o(x') = \frac{1}{2}x'^T \bar{M}x'$ the observability function and by $\bar{L}_c(x') = \frac{1}{2}x'^T \bar{W}^{-1}x'$ the controllability function of the linear system, where

$$\bar{M} = \frac{\partial^2 \bar{L}_o}{\partial x'^2}(x'), \quad \bar{W}^{-1} = \frac{\partial^2 \bar{L}_c}{\partial x'^2}(x') \quad (24)$$

are the constant Gramians of the linear system. Moreover

$$\bar{L}_o(x') = L_o(\eta^{-1}(x)), \quad \bar{L}_c(x') = L_c(\eta^{-1}(x)).$$

This leads to the following relation:

$$\frac{\partial \bar{L}_o}{\partial x'}(x') = \frac{\partial L_o}{\partial x}(\eta^{-1}(x')) \left(\frac{\partial \eta}{\partial x}(\eta^{-1}(x')) \right)^{-1} \quad (25)$$

Let T be the matrix associated with the metric for the symmetric linear system. Then, according to Theorem 3, statement 2, $\bar{W}T = T^{-1}\bar{M}$ that can be rewritten as $\bar{M}^{-1}T = T^{-1}\bar{W}^{-1}$. Postmultiplying with x' we get: $(\partial^2 \bar{L}_o(x')/\partial x'^2)^{-1} T x' = T^{-1}(\partial^T \bar{L}_c(x')/\partial x')$. Using relation (25) and $x' = \eta(x)$ we can write:

$$\frac{\partial^T L_c}{\partial x}(x) = \frac{\partial^T \eta}{\partial x}(x) \cdot T \cdot \left(\frac{\partial^2 \bar{L}_o}{\partial x'^2}(\eta(x)) \right)^{-1} \cdot T \cdot \eta(x) \quad (26)$$

which shows that the observability function, the metric, and the isomorphism between the systems give the controllability function of the gradient system. This can be called a nonlinear version of the cross Gramian idea for this particular case, and it motivates the search for the nonlinear cross Gramian in the general case.

5 Nonlinear Cross Gramian

In this section, we will make an analysis of the variational part of the prolonged system. Denote by:

$$\Sigma'_p : \begin{cases} \dot{v} = \frac{\partial(f(x) + g(x)u)}{\partial x} v + g(x)u_p \\ y_p = \frac{\partial h(x)}{\partial x} v \end{cases}, \quad (27)$$

where x is considered a parameter varying according to (1).

Since the system is asymptotically stable, by the definition of its variational associated system, the latter is also asymptotically stable. By Theorem 4, Σ'_p has the same input-output behavior as the system Σ'_g , given by:

$$\begin{cases} \dot{p} = \frac{\partial^T(f(x) + g(x)u)}{\partial x} p + \mathcal{F}(g_{ij}(x), \frac{\partial g_{ij}(x)}{\partial x_k}, f_k(x), u, g(x), p) + \frac{\partial^T h(x)}{\partial x} u_g \\ y_g = g^T(x)p \end{cases}, \quad (28)$$

where x again is a parameter varying as in (1). According to Lemma 1, there exists a coordinate transformation such that $p = \psi(x, v)$, given by $\psi(x, v) = G(x)v$, where $G(x)$ is symmetric and invertible (as in the definition of (8)) and is given by the pseudo-Riemannian metric. Applying the coordinate transformation on Σ'_p , we get:

$$G(x)g(x) = \frac{\partial^T h(x)}{\partial x} \quad \text{and} \quad \frac{\partial h(x)}{\partial x} G^{-1}(x)p = g^T(x)p. \quad (29)$$

Remark 7 *In the linear systems case, everything fits with the definition and characterization of the property of symmetry. Moreover, the linearization of Σ'_p and Σ'_g around an equilibrium point $(x_0, 0, 0, 0)$ yields the v part and p part of (14), with $p = Tv$, with T invertible and symmetric.*

Based on the local existence of the cross Gramian, we make an analysis of the observability function of Σ'_p . In this case, $u = 0$, $u_p = 0$ and Σ'_p becomes:

$$\begin{cases} \dot{v} = \frac{\partial f(x)}{\partial x} v \\ y_p = \frac{\partial h(x)}{\partial x} v \end{cases}. \quad (30)$$

Assuming the zero-state observability combined with the asymptotic stability of Σ'_p implies the existence of the observability function $L_o^p(x, v) > 0$, $L_o^p(x_0, 0) = 0$, defined as:

$$L_o^p(x, v) = \frac{1}{2} \int_t^\infty y_p^T(\tau) y_p(\tau) d\tau \quad (31)$$

and satisfying the nonlinear Lyapunov equation:

$$\frac{\partial L_o^p(x, v)}{\partial v} \frac{\partial f(x)}{\partial x} v + \frac{1}{2} v^T \frac{\partial^T h(x)}{\partial x} \frac{\partial h(x)}{\partial x} v = -\frac{\partial L_o^p(x, v)}{\partial x} f(x). \quad (32)$$

Since the system is linear in v , without loss of generality, we can write $L_o^p(x, v)$ as:

$$L_o^p(x, v) = \frac{1}{2} v^T \mathcal{L}(x) v \quad (33)$$

with $\mathcal{L}(x)$ symmetric, positive definite and with smooth elements.

In the sequel, we determine the nonlinear counterpart of the Sylvester equation which in the linear case gives the cross Gramian. Taking the derivative with respect to v and using (29), we get:

$$\frac{\partial^2 L_o^p(x, v)}{\partial v^2} \frac{\partial f(x)}{\partial x} v + \frac{\partial^T f(x)}{\partial x} \frac{\partial^T L_o^p(x, v)}{\partial v} + G(x)g(x) \frac{\partial h(x)}{\partial x} v = -\frac{\partial^2 L_o^p(x, v)}{\partial v \partial x} f(x). \quad (34)$$

Applying the coordinate transformation, $p = G(x)v$, on (30) we get:

$$\frac{\partial f}{\partial x}v = G^{-1}(x)\frac{\partial^T f}{\partial x}G(x)v + \overline{\mathcal{F}}(g_{ij}(x), \frac{\partial g_{ij}(x)}{\partial x_k}, f_k(x), u, g(x), p), \quad (35)$$

where $\overline{\mathcal{F}} = \mathcal{F} - G^{-1}(x)\dot{G}(x)v$.

Premultiplying the equation with v^T and using (35) we obtain:

$$\begin{aligned} p^T G^{-1}(x)\mathcal{L}(x)\frac{\partial f(x)}{\partial x}v + p^T \frac{\partial f(x)}{\partial x}G^{-1}(x)\mathcal{L}(x)v + p^T g(x)\frac{\partial h}{\partial x}v = \\ -v^T \frac{\partial^2 L_o^p(x, v)}{\partial v \partial x}f(x) - \overline{\mathcal{F}}^T \mathcal{L}(x)v. \end{aligned} \quad (36)$$

Using the coordinate transformation (29) and equation (32) we get:

$$\begin{aligned} p^T G^{-1}(x)\mathcal{L}(x)\frac{\partial f}{\partial x}v + \frac{1}{2}p^T g(x)\frac{\partial h}{\partial x}v = \\ -v^T \frac{\partial^2 L_o^p(x, v)}{\partial v \partial x}f(x) + \frac{\partial L_o^p(x, v)}{\partial x}f(x) - \overline{\mathcal{F}}^T \mathcal{L}(x)v. \end{aligned} \quad (37)$$

Remark 8 In the linear systems case, (34) becomes: $v^T MAv + v^T A^T Mv + p^T BCv = 0$. Since $v = T^{-1}p$, we get:

$$p^T T^{-1}MAv + p^T AT^{-1}Mv + p^T BCv = 0. \quad (38)$$

Using the symmetry property, this immediately leads to the Sylvester equation (4). Moreover, the relation $X = T^{-1}M$ is satisfied as in Theorem 3. Equation (38) becomes

$$XA + \frac{1}{2}BC = 0.$$

We are now ready to define the cross Gramian for a nonlinear gradient system.

Definition 7. We call

$$\mathcal{X}(x) = G^{-1}(x)\mathcal{L}(x) \quad (39)$$

the *cross-Gramian* matrix associated to Σ'_p and it satisfies (37).

This is an extension of statement 2 in Theorem 3, i.e. the cross Gramian is given by the gradient metric and the observability Gramian. In order to explain the cross Gramian and its importance we present in a nutshell the study of Hankel singular values for a nonlinear system (1) as in [6, 11]. Suppose that (1) is asymptotically reachable from $x(0)$, then the controllability function $L_c(x)$ exists and is positive definite, with $L_c(x_0) = 0$.

If $\mathcal{H}(u)$ is the Hankel operator of the system then for finding out the Hankel singular values of the system the differential problem is solved: $(d\mathcal{H}(u))^* \mathcal{H}(u) = \lambda u$, where $(d\mathcal{H}(u))^*$ represents the adjoint of the $(d\mathcal{H}(u))$ operator (see [6] for further details). A solution for this problem is given by the following result:

Lemma 3. [6] *If there exists $\lambda \neq 0$ such that*

$$\frac{\partial L_o}{\partial x}(x(0)) = \lambda \frac{\partial L_c}{\partial x}(x(0)), \quad (40)$$

then λ is an eigenvalue of the operator $(d\mathcal{H}(u))^ \mathcal{H}(u)$, with the corresponding eigenvector $u = \mathcal{C}^\dagger(x(0))$, where $\mathcal{C}(u)$ is the controllability operator associated to (1).*

Remark 9 *In the linear case, this problem becomes: $Mx(0) = \lambda W^{-1}x(0)$. Since $W > 0$, we can write $WMx(0) = \lambda x(0)$ and if, moreover, the system is gradient, then, according to Theorem 6 we have: $X^2x(0) = \lambda x(0)$, X being the cross Gramian. This means that λ is the squared Hankel singular value σ , which for a symmetric system is an eigenvalue of X .*

Still, in order to make the connection between λ 's and the Hankel singular values of (1) the Hankel norm is involved. The following results give the relation:

Theorem 5. [6] *Suppose that the linearization of (1) has non-zero distinct Hankel singular values. Then, there exists a neighborhood $U \subset \mathbb{R}$ of 0 and $\rho_i(s) > 0$, $i = 1, \dots, n$ such that: $\min\{\rho_i(s), \rho_i(-s)\} \geq \max\{\rho_{i+1}(s), \rho_{i+1}(-s)\}$ holds for all $s \in U$, $i = 1, \dots, n-1$. Moreover, there exist $\xi_i(s)$, satisfying the following:*

$$\begin{aligned} L_c(\xi_i(s)) &= s^2/2, L_o(\xi_i(s)) = \rho_i(s)s^2/2, \frac{\partial L_o}{\partial x}(\xi_i(s)) = \lambda_i(s) \frac{\partial L_c}{\partial x}(\xi_i(s)), \\ \lambda_i(s) &= \rho_i^2(s) + \frac{s}{2} \frac{d\rho_i^2(s)}{ds}. \end{aligned} \quad (41)$$

Even more, if $U = \mathbb{R}$, the Hankel norm of the system is $\sup_s \rho_1(s)$.

The $\rho_i(s)$ are a clear extension of the Hankel singular values for a nonlinear system and they can be obtained from the Hankel singular value functions of the nonlinear system, as defined in [9]. The following result establishes this link:

Theorem 6. [11] *If (1) is in input-normal, output-diagonal form, i.e. $L_c(x) = x^T x/2$, $L_o(x) = x^T \text{diag}(\tau_1(x), \dots, \tau_n(x))x/2$, then*

$$\begin{aligned} \rho_i^2(x_j) &= \tau_i(0, \dots, x_j, \dots, 0), \quad i \neq j, \\ \rho_j^2(x_j) &= \tau_j(0, \dots, x_j, \dots, 0) + \frac{1}{2} \frac{\partial \tau_j}{\partial x_j}(0, \dots, x_j, \dots, 0)x_j. \end{aligned} \quad (42)$$

Returning to our case, we state the following

Conjecture 1 *Let (1) be a nonlinear gradient system with the associated variational system Σ'_p . If λ_i , $i = 1, \dots, n$, satisfy Theorem 5, then they are the squared eigenvalues of $\mathcal{X}(x)$.*

We aim at proving this conjecture by finding the meaning of the gradient extension in the context of the balancing procedure (following the reasoning in e.g. [6]),

in order to be able to obtain an equivalent of equation (40) written in terms of the cross Gramian. In this way, the λ 's in (40) associated to Σ_p , are related to the eigenvalues of the cross Gramian and thus, the Hankel singular value functions can be obtained from solving an eigenvalue problem for the cross Gramian.

Remark 10 *For linear systems this falls into place with the theory for symmetric systems, see Remark 9.*

Then using Theorem 5, the Hankel singular values of the original system are obtained, avoiding the balancing procedure.

6 Example

Given a double mass double spring system (see Figure 1), we compute the cross Gramian of the gradient system associated to it.

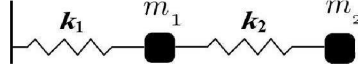


Fig. 1 Double mass double spring system

The system is given by:

$$\begin{cases} m_1 \ddot{x}_1 + k_1(x_1) + k_2(x_1, x_2) = 0 \\ m_2 \ddot{x}_2 - k_2(x_1, x_2) + u = 0 \end{cases}, \quad (43)$$

where x_1, x_2 are the displacements, $m_1, m_2 > 0$ are the masses and $k_1(x_1), k_2(x_1, x_2)$ are the corresponding elastic forces, with the initial conditions $x_1(0) = 1, x_2(0) = 0$. The potential energy of the system is given by $V(x)$, smooth, such that $\frac{\partial V(x)}{\partial x_1} = k_1(x_1)$, $\frac{\partial V(x)}{\partial x_2} = k_2(x_1, x_2)$. We choose $k_1(x_1) = -x_1^3$ and $k_2(x_1, x_2) = x_1 - x_2$ (elastic coefficients constant and equal to 1). We take $m_1 = m_2 = 1$. The Hamiltonian of the system is $H(x) = \frac{1}{2} \dot{x}^T M^{-1}(x) \dot{x} + V(x)$, with $M(x) = I_2$. Since $M > 0$, $G(x) = M^{-1}(x) = I_2$ can define a Riemannian metric on \mathbb{R}^n (e.g. see [10, Chapter 6, Section 6.1]). The associated gradient system, of the form (8), is:

$$\dot{x} = \begin{bmatrix} -x_1^3 \\ x_1 - x_2 \end{bmatrix} + \begin{bmatrix} 2 \\ 1 \end{bmatrix} u, \quad y = x_1. \quad (44)$$

Denote $\mathcal{L}(x(t)) = [l_{ij}(x(t))]_{i,j=1,2} = [\mathbf{l}_{ij}(t)]_{i,j=1,2}$. Solving equation (37) associated to (44), for all $v \in T\mathcal{M}$, yields the following parameter-varying system to be solved:

$$\begin{cases} \frac{d\mathbf{I}_{11}(t)}{dt} = 3x_1^2(t)\mathbf{I}_{11}(t) - \mathbf{I}_{12}(t) - 1 \\ \frac{d\mathbf{I}_{12}(t)}{dt} = \left(\frac{3}{2}x_1^2(t) + 1\right)\mathbf{I}_{12}(t) + \mathbf{I}_{22}(t) \\ \frac{d\mathbf{I}_{22}(t)}{dt} = \mathbf{I}_{22}(t) \end{cases} \quad (45)$$

Solving system (44) for $u(t) = 0$, $t > 0$, $x_1(0) = 1$ we get $x_1(t) = \frac{1}{\sqrt{2t+1}}$. Substituting in (45) we obtain a time varying system. We solve it using approximation of 3rd order and obtain:

$$\mathcal{L}(t) = \mathcal{X}(t) = \begin{bmatrix} 3 + 10t + 9t^2 + 2t^3 & -t - \frac{3}{2}t^2 - \frac{1}{6}t^3 \\ -t - \frac{3}{2}t^2 - \frac{1}{6}t^3 & 1 + t + \frac{1}{2}t^3 + \frac{1}{6}t^3 \end{bmatrix} \quad (46)$$

The eigenvalue functions of the cross Gramian are given as:

$$\begin{aligned} \lambda_1(t) &= 3 + 10t + 10t^2 - 3t^3 + O(t^4) \\ \lambda_2(t) &= 1 + t + 0.9t^3 + O(t^4) \end{aligned} \quad (47)$$

7 Conclusions and Future Work

We present here the nonlinear counterpart of the cross Gramian for gradient systems. We do this in terms of the variational system. The reason is that in the next step we want to prove that the eigenvalues obtained from the cross Gramian are related in a direct manner to the Hankel singular values of the system. For later concern we will also take into account the computational aspect of solving equation (37).

Acknowledgements The authors gratefully acknowledge the contribution of National Research Organization and reviewers' comments.

References

1. R. W. Aldaheri: Model order reduction via real Schur-form decomposition. *Int. J. Control* **53**(3) (1991) pp. 709–716
2. A. C. Antoulas: *Approximation of Large-scale Systems*. SIAM, 2005.
3. J. Cortes, A. J. van der Schaft, P. E. Crouch: Characterization of gradient control systems. *SIAM J. Contr. & Opt.* **44** (2005) pp. 1192–1214
4. K. V. Fernando, H. Nicholson: On the structure of balanced and other principal representations of SISO systems. *IEEE Trans. Automat. Contr.* **28**(2) (1983) pp. 228–231
5. K. Fujimoto, J. M. A. Scherpen: Singular Value Analysis of Hankel Operators for General Nonlinear Systems. *Proceedings of ECC* (2003)

6. K. Fujimoto, J. M. A. Scherpen: Nonlinear Input-Normal Realizations Based on the Differential Eigenstructure of Hankel Operators. *IEEE Trans. on Aut. Control* **50** (2005)
7. W. S. Gray, J. P. Mesko: Observability functions for linear and nonlinear systems. *System and Control Letters* **33** (1999) pp. 99–113
8. J. E. Marsden, T. Ratiu, R. Abraham: *Manifolds, Tensor Analysis and Applications*. Springer-Verlag, NY (2001)
9. J. M. A. Scherpen: Balancing for nonlinear systems, *System & Control Letters* **21** (1993) pp. 143–153
10. J. M. A. Scherpen: *Balancing for Nonlinear Systems*. PhD Thesis, Univ. of Twente, Netherlands (1994)
11. J. M. A. Scherpen, K. Fujimoto, W. S. Gray: Hamiltonian extensions, Hilbert adjoints and singular value functions for nonlinear systems. *Proc. of the 39th IEEE CDC 2000* **5** (2000) pp. 5102–5107
12. J. M. A. Scherpen, W. S. Gray: Minimality and local state decompositions of a nonlinear state space realization using energy functions, *IEEE Trans. on Aut. Contr.* **45**(11) 2000 pp. 2079–2086
13. J. M. A. Scherpen: Duality and singular value functions of the nonlinear normalized right and left coprime factorizations. *Proceedings of the 44th IEEE CDC-ECC* (2005) pp. 2254–2259
14. D. C. Sorensen, A. C. Antoulas: The Sylvester equation and approximate balanced truncation. *Linear Algebra and Its Applications* **351-352** (2002) pp. 671–700
15. A. J. van der Schaft: Linearization of Hamiltonian and gradient systems. *IMA J. Math. Contr. & Inf* **1** (1984) pp. 185–198
16. K. Zhou, J. C. Doyle and K. Glover: *Robust Optimal Control*. Prentice-Hall, 1996