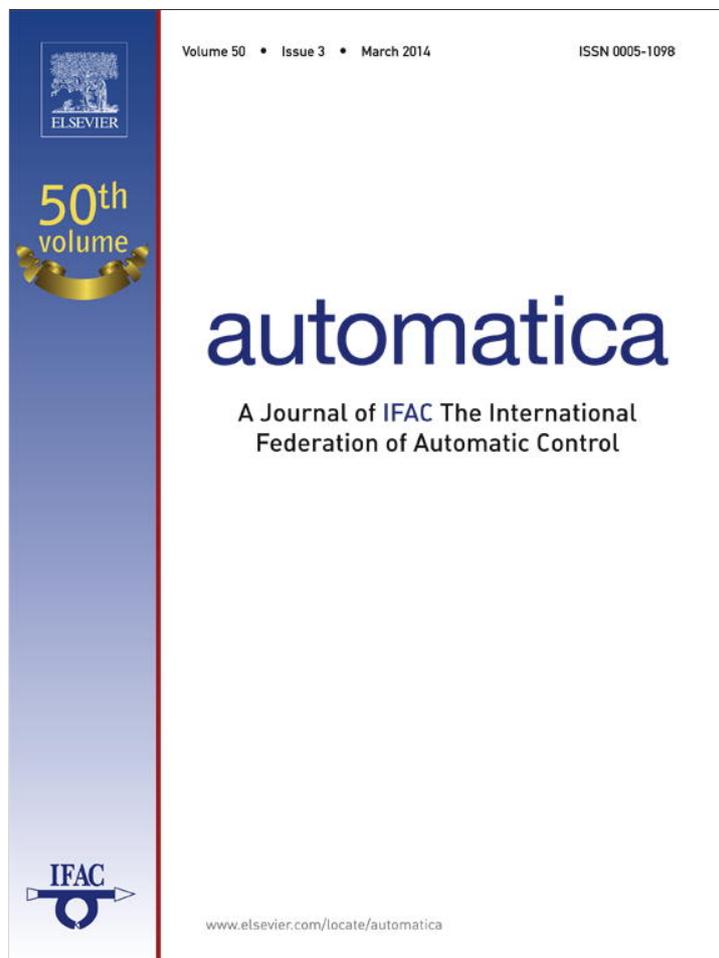


Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/authorsrights>



A variational integrators approach to second order modeling and identification of linear mechanical systems[☆]



Mattia Bruschetta^a, Giorgio Picci^{a,1}, Alessandro Saccon^b

^a Department of Information Engineering, Università di Padova, via Gradenigo 6/B, 35131 Padova, Italy

^b Department of Mechanical Engineering, Eindhoven University of Technology, The Netherlands

ARTICLE INFO

Article history:

Received 13 April 2011

Received in revised form

12 September 2013

Accepted 20 November 2013

Available online 26 December 2013

Keywords:

Discrete mechanical systems

Variational integrators

Second order models

Subspace identification

ABSTRACT

The theory of variational integration provides a systematic procedure to discretize the equations of motion of a mechanical system, preserving key properties of the continuous time flow. The discrete-time model obtained by variational integration theory inherits structural conditions which in general are not guaranteed under general discretization procedures. We discuss a simple class of variational integrators for linear second order mechanical systems and propose a constrained identification technique which employs simple linear transformation formulas to recover the continuous time parameters of the system from the discrete-time identified model. We test this approach on a simulated eight degrees of freedom system and show that the new procedure leads to an accurate identification of the continuous-time parameters of second-order mechanical systems starting from discrete measured data.

© 2013 Elsevier Ltd. All rights reserved.

1. Introduction and motivations

The identification of linear second order models of mechanical systems has been the object of intensive research and of several papers in the past decade (De Angelis, Lus, Betti, & Longman, 2002; Lus, De Angelis, Betti, & Longman, 2002, 2003). Of particular interest are systems which can be described by a second order vector model of the following form:

$$M\ddot{q} + D\dot{q} + Kq = f \quad (1.1)$$

where M and K , both symmetric positive definite matrices in $\mathbb{R}^{n \times n}$, have the interpretation of generalized mass (or inertia) and generalized stiffness coefficient matrices respectively, while $D \in \mathbb{R}^{n \times n}$, $D = D^T$ is a linear (viscous) damping coefficient which is at least positive semidefinite. The generalized forces f acting on the system can be expressed as a linear function of a vector of independently assignable generalized input forces u of dimension

$m \leq n$; namely

$$f = Lu \quad (1.2)$$

where the matrix L , which will be assumed to be known, describes the physical locations at which the input forces u act on the system. Without loss of generality it may be assumed that L is of full column rank; i.e. $\text{rank } L = m$.

For simplicity and for mathematical convenience we shall assume that a full set of linear sensors is available to the experimenter; i.e., that all n degrees of freedom are measured via linear sensors. This is obviously equivalent to assume that the measurement equation is $y = q$. The system (1.1) can also be represented in state space form; for example, defining $x := [q, \dot{q}]^T$, one gets

$$\dot{x} = \begin{bmatrix} 0 & I \\ -M^{-1}K & -M^{-1}D \end{bmatrix} x + \begin{bmatrix} 0 \\ M^{-1}L \end{bmatrix} u \quad (1.3)$$

which should be coupled with the measurement (output) equation $y = [I \ 0]x$. Note that under our assumptions the system is automatically controllable and observable and hence minimal. This is a necessary condition for parameter identifiability. See Laub and Arnold (1984) for a direct test of controllability/observability of second order models of the type considered in this paper.

Now, in several areas of engineering, such as mechanical or structural engineering, an accurate estimation of the parameters (M , K , D) of the physical *continuous time* model (1.1) is often required. A typical example being the estimation of deformations

[☆] This work was supported in part by the National Project *Identification and Adaptive control of Industrial Systems* funded by the Italian Ministry of higher education. The material in this paper was partially presented at the 15th IFAC Symposium on System Identification (SYSID), July 6–8, 2009, St Malo, France. This paper was recommended for publication in revised form by Associate Editor Wolfgang Scherrer under the direction of Editor Torsten Söderström.

E-mail addresses: bruschet@dei.unipd.it (M. Bruschetta), picci@dei.unipd.it (G. Picci), asaccon@tue.nl (A. Saccon).

¹ Tel.: +39 049 8277705; fax: +39 049 8277699.

at points which are not monitored or the estimation of proper modes of vibration of a mechanical structure.

Mainstream system identification theory deals with discrete-time data and discrete-time models and normally the recovery of the continuous-time parameters involves a conversion step from discrete to continuous time (the so-called *indirect approach*). The problem of reconstructing a continuous-time model from an identified discrete-time model has a long history and has been discussed in several places, see e.g. Söderström (1991) and the reference list in the more recent paper Mahata and Fu (2007). The conversion step from discrete to continuous may sometimes be cause of troubles. It is a commonly experienced fact that for multivariable systems of moderate/large dimension, accurate values of the continuous-time parameters may be hard to recover from the estimated discrete-time system, no matter how accurate the estimates of the latter may be. One reason of this difficulty may be attributed to the ill-conditioning of the discrete-to-continuous conversion, which involves, in the zero-(or first-order)-hold (ZOH) discretizations,² inverting the exponential relation $F = \exp Ah$, $G = \int_0^h \exp As ds B$ to recover the matrices (A, B) of the continuous-time model from estimates (F, G) of an identified discrete time model

$$x_{k+1} = Fx_k + Gf_k, \quad y_k = Hx_k + Jf_k. \quad (1.4)$$

The default option in the discrete-to-continuous (d2c) routine in MATLAB is ZOH. It is well-known that this operation may turn into an ill-conditioned problem since the recovery of matrix A involves the computation of the logarithm of F which may be a complex matrix or, may be undefined as requiring the inversion of the exponential map in a region of the complex plane where it is not invertible. We would like to point out that the common belief that this problem should be solvable by choosing a high sampling frequency may actually worsen the problem. Consider the trivial case of a scalar discrete F subject to a perturbation δF . The relative error incurred when computing $A + \delta A := \frac{1}{h} \log(F + \delta F)$ is

$$\frac{\delta A}{A} = \frac{1}{\log F} \frac{\delta F}{F}$$

a more complicated formula holding in the matrix case, see Dieci and Papini (2000, Formula 2.3). Since for $h \rightarrow 0$, $F \rightarrow I$, the condition number of computing $A = \frac{1}{h} \log F$ tends to infinity when $h \rightarrow 0$. This means that at high sampling frequency, the effect of unavoidable random errors on the estimates of F (and G) could be largely amplified when computing A by the logarithmic transformation. See Dieci and Papini (2000) and the references therein.

A possible option in the Matlab d2c routine is the so-called *Tustin transform*. Since this discretization scheme has superficial similarities with the approach proposed in this paper and deserves an accurate analysis, we shall postpone a detailed discussion of this option to Section 3.

Now, since the problem we are discussing is a specific parameter estimation problem, the continuous model structure obtained from the discrete-time identified model should be easily transformable into the form (1.1) or (1.3) in that particular basis. In general however, an identified discrete model will just have a generic structure (1.4) with full matrices (F, G, H, J) and need not have

any of the structural properties of a mechanical system. In particular, the continuous state-space realization obtained by the inverse of the ZOH or FOH discretizations does not lead to an input–output relation of the special second-order form (1.1). This means that the recovery of the physical parameters M, D, K may in general be ill-defined or impossible. That this is not of purely academic interest is witnessed by the interest in this problem in the recent mechanical engineering literature, see e.g. De Angelis et al. (2002), Lus, De Angelis, and Betti (2003), Lus, De Angelis, Betti, Longman (2003) and the references therein. Ideally, we would like to use discretization schemes which preserve the second order input–output structure of the type (1.1), which is a basic characteristic of linear models of fully observed mechanical systems (Newton law).

In addition, besides the previous difficulties, since the inverse discretization transform is generally non-linear it does introduce bias in the estimates of the continuous-time parameters, even when the estimates of the discrete-time parameters are unbiased and accurate. For this reason a linear (or “approximately linear”) discrete-to-continuous conversion would be highly desirable.³

1.1. On continuous-time identification

An alternative approach could be to identify the continuous-time parameters directly (the so-called *direct approach*). This could be done in several ways. One may attempt to identify the parameters of the model (1.1) or (1.3) from (discrete) noisy observations directly, by using a continuous-time PEM method. However continuous-time iterative optimization methods are especially sensitive to the choice of good initial estimates of the parameters, particularly when the data sampling frequency may not be suitable for a reasonable numerical approximation of derivatives and gradients. The sensitivity to initial estimates is a serious difficulty especially for multivariable (multi input/multi output) models like the mechanical systems we are dealing with, where good quality initial parameter estimates may be hard to obtain. The problem of getting reasonable initial continuous-time parameter estimates seems indeed to be a non trivial one.

Correlation methods, say by replacing the differentiation operator with the so-called delta operator (Feuer & Goodwin, 1996) or by various approximations of the continuous derivative operator have been proposed (Söderström, Fan, Carlsson, & Bigi, 1997). The approach seems to be advantageous only if the underlying continuous system is scalar autoregressive. The accuracy of the approach depends on the particular approximation being used and does not seem to be easy to assess, especially when the method should be applied to multivariable continuous-time models of moderate or high dimension. Other continuous-time identification algorithms eventually end up to rely on logarithmic transforms, like inverting the relation $z = \exp\{sh\}$ which turns out to be equivalent to the logarithmic d2c transformation. For the reasons given above, these methods are not always reliable. There is also a quite popular approach based on filtering the continuous-time data by a family of test functions (Heuberger, de Hoog, van den Hof, & Wahlberg, 2003; Ohta & Kawai, 2004), which may or may not be orthonormal. This approach needs extensive numerical integration for computing the inner products over a long period of time, since in order to reach a reasonable accuracy the computation of many inner products of the measured signal with a large number of test functions is needed. In a sense each inner product plays eventually the role of a single discrete-time sample value of the signal.

Unfortunately according to the current literature on continuous time identification, see e.g. Garnier and Wang (2008), Sinha and Rao (1991) and the references therein, the existing

² The ZOH sampler transforms continuous-time into discrete-time by synchronously sampling the output of the continuous system once the input signal is approximated by a piecewise constant function on each sampling interval. There are more refined schemes, such as the first-order-hold (FOH) which assumes instead the input to be piecewise linear. The discussion which follows applies also to FOH, modulo notational complications which we choose to avoid.

³ One may argue that the Euler discretization is a well known instance of linear conversion map but unfortunately the Euler discretization is too primitive to be of use in most situations.

continuous-time algorithms do not seem to be of much help for an accurate identification of the physical parameters of multi-variable mechanical systems. To our knowledge, reliable one shot continuous-time identification methods which can be applied to concrete multivariable real-world problems seem still to be on the way. There seem to be progress still to be made in this area.

1.2. Contributions of the paper

After this introduction, we shall describe the content of the paper. The layout is as follows.

- (1) We shall introduce a discretization technique of system (1.1) or (1.3) based on the idea of variational integrators. This technique leads to conversion formulas from a discrete identified model to the corresponding continuous *input-output* model which are simple and generally better conditioned.
- (2) We shall show that in an important special case the variational discretization leads to a well-known continuous-to-discrete transformation used in system and control, namely the Cayley–Tustin discretization. The model discretization is different from the traditional Cayley–Tustin discretization formula since it involves uniformly sampled values of the input function available at the outset. We shall discuss this peculiar discretization technique and we shall address some related approximation problems.
- (3) We shall use the variational discretization to attack the mechanical system identification from noisy sampled input–output data. As a preliminary step we shall discuss a standard discrete-time subspace identification technique which is used in order to supply good starting values to a successive Prediction Error optimization-based algorithm.
- (4) We shall describe a refinement of the subspace identification estimates by a discrete-time Prediction Error algorithm which updates the Variational model directly and complies with the constraints of second order mechanical structure. The inverse Cayley–Tustin discretization formulas then provide the continuous-time estimates (M, D, K) . This is the final step of the procedure.
- (5) Finally, simulation results are shown and compared with the results obtained by state of the art identification methods.

2. The variational integrators approach to discretization

A novel twist to the discretization of mechanical systems has been provided by the theory of *variational integrators*, see Veselov (1988), and the extensive work of J. Marsden and co-workers, see e.g. Marsden and West (2001). These techniques seem to be fairly well known to numerical analysts working with mechanical models but not so familiar to the system and control community. The key idea is that the discrete equations of motion should not be derived by attempting a direct discretization of Eqs. (1.1) or (1.3) but rather derived by paraphrasing what happens in continuous time; i.e. by making stationary a *discrete action integral* defined in terms of a suitable *discrete Lagrangian function*. The (discrete) equations of motion should then follow just like the Euler–Lagrange equations in continuous time. In short, the variational integrators paradigm is to build from scratch a theory of Lagrangian *Discrete Mechanics*.

In (continuous-time) Lagrangian mechanics, given a Lagrangian function $L(q(t), \dot{q}(t))$ and external forces $f_L(q(t), \dot{q}(t), t)$, the equations of motion follow from the *Lagrange–d’Alembert principle*, stating that the trajectory of a mechanical system starting at time t_0 in position q_0 and arriving at time t_1 at position q_1 must satisfy the variational principle

$$\delta \int_{t_0}^{t_1} L(q, \dot{q}) dt + \int_{t_0}^{t_1} f_L(q, \dot{q}, t) \delta q(t) dt = 0 \quad (2.1)$$

for arbitrary variations $\delta q(t)$, while holding the endpoints q_0 and q_1 of the curve $t \mapsto q(t)$ fixed. This leads to the well-known forced *Euler–Lagrange equations*

$$\frac{\partial L}{\partial q}(q, \dot{q}) - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}}(q, \dot{q}) \right) + f_L(q(t), \dot{q}(t), t) = 0$$

which for the quadratic Lagrangian $L(q(t), \dot{q}(t)) = \frac{1}{2} \dot{q}^T M \dot{q} - \frac{1}{2} q^T K q$ and the external force f_L composed by a dissipation force $-D\dot{q}$ and the actual (generalized) external force $f(t)$, i.e.,

$$f_L(q(t), \dot{q}(t), t) := -D\dot{q}(t) + f(t), \quad (2.2)$$

leads to the linear second order vector differential equation (1.1).

In discrete Lagrangian mechanics, one first considers a discretization $\{q_k = q(kh); k = 0, \dots, N\}$ and curve segments $\{q_{k,k+1}(t); t \in [kh, (k+1)h)\}$ between two configuration points $q_k = q(kh)$ and $q_{k+1} = q((k+1)h)$ in the configuration space, placed h units of time apart. One defines an *exact (forced) discrete Lagrangian* and the *exact discrete forces* on that curve segment as:

$$\begin{aligned} L_d^E(q_k, q_{k+1}, h) &:= \int_{kh}^{(k+1)h} L(q_{k,k+1}(t), \dot{q}_{k,k+1}(t)) dt, \\ f_d^{E-}(q_k, q_{k+1}, h) &:= \int_{kh}^{(k+1)h} f_L(q_{k,k+1}(t), \dot{q}_{k,k+1}(t), t) \frac{\partial q_{k,k+1}(t)}{\partial q_k} dt, \\ f_d^{E+}(q_k, q_{k+1}, h) &:= \int_{kh}^{(k+1)h} f_L(q_{k,k+1}(t), \dot{q}_{k,k+1}(t), t) \frac{\partial q_{k,k+1}(t)}{\partial q_{k+1}} dt, \end{aligned}$$

and then the following *Discrete Lagrange–d’Alembert principle*

$$\begin{aligned} \delta \sum_{k=0}^{N-1} L_d^E(q_k, q_{k+1}, h) \\ + \sum_{k=1}^{N-1} \left(f_d^{E+}(q_{k-1}, q_k, h) + f_d^{E-}(q_k, q_{k+1}, h) \right) \delta q_k = 0 \end{aligned}$$

where the variation $\delta q(t)$ of a continuous curve is replaced by a discrete (finite) sequence of variations $\{\delta q_k\}_{k=0, \dots, N}$, for arbitrary δq_k 's. The variation is computed with *fixed* end points. This leads to the (*Exact*) *Discrete Euler–Lagrange equations*

$$\begin{aligned} D_2 L_d^E(q_{k-1}, q_k, h) + D_1 L_d^E(q_k, q_{k+1}, h) \\ + f_d^{E+}(q_{k-1}, q_k, h) + f_d^{E-}(q_k, q_{k+1}, h) = 0 \end{aligned} \quad (2.3)$$

where D_i stands for the partial derivative operator applied to the i -th argument of the function on which it is acting. These equations should be interpreted as an algorithm mapping the pair (q_k, q_{k+1}) to the next configuration pair (q_{k+1}, q_{k+2}) . See Marsden and West (2001, p. 427) for details.

If it were possible to compute the integrals explicitly, we would obtain a discrete model which describes exactly the continuous dynamic sampled at the discrete time instants $t = kh$. In general this is not possible and one needs to use approximations both for the discrete Lagrangian and for the discretized external forces. These approximations we denote $L_d(q_k, q_{k+1})$, $f_d^+(q_k, q_{k+1}, k)$, $f_d^-(q_k, q_{k+1}, k)$. It is remarkable that although many approximations are possible, one of which will be described below, the “stationary action” principle leads in any case to approximate Discrete Euler–Lagrange (DEL) equations of the standard form (2.3).

The variational midpoint discretization: One of the simplest and most commonly used methods to approximate the Lagrangian and the external forces is by the so-called “midpoint rule”, whereby

$$L_d(q_k, q_{k+1}, h) := hL \left(\frac{q_{k+1} + q_k}{2}, \frac{q_{k+1} - q_k}{h} \right), \quad (2.4)$$

$$f_d^-(q_k, q_{k+1}, h) := \frac{h}{4} \left[f_L \left(\frac{q_{k+1} + q_k}{2}, \frac{q_{k+1} - q_k}{h}, hk \right) + f_L \left(\frac{q_{k+1} + q_k}{2}, \frac{q_{k+1} - q_k}{h}, h(k+1) \right) \right] \quad (2.5)$$

$$f_d^+(q_k, q_{k+1}, h) := f_d^-(q_k, q_{k+1}, h). \quad (2.6)$$

This scheme applies to an arbitrary Lagrangian and external forces and leads to a discretization of quite arbitrary nonlinear dynamics. Obviously, the solution of the approximate DEL equations derived from an approximate Lagrangian and approximate discrete forces will no longer be equal to the true configuration variable sampled at the discrete time instants $t = kh$, so that q_k will now have to be interpreted just as an approximation of $q(kh)$. For clarity we shall henceforth denote this approximation by the symbol $\mathbf{q}(k)$.

When L is the quadratic Lagrangian defined above, using the approximations (2.4)–(2.6), a straightforward computation leads to the resulting approximate DEL equation

$$M_d \mathbf{q}(k) + D_d \mathbf{q}(k-1) + K_d \mathbf{q}(k-2) = f_d(k), \quad (2.7)$$

where

$$M_d := \frac{M}{h} + \frac{hK}{4} + \frac{D}{2}, \quad (2.8a)$$

$$D_d := \frac{hK}{2} - \frac{2M}{h}, \quad (2.8b)$$

$$K_d := \frac{M}{h} + \frac{hK}{4} - \frac{D}{2}, \quad (2.8c)$$

and

$$f_d(k) := \frac{hL}{4} (u(kh) + 2u[(k-1)h] + u[(k-2)h]), \quad (2.9)$$

$u(kh)$ denoting the sampled value of the input u at time kh . Eq. (2.7) is the discrete-time counterpart to the continuous-time system (1.1). Note that the computation of the discrete forcing function $\{f_d(k)\}$ requires adjacent samples at times k , $k-1$ and $k-2$ of the sampled external force f (or equivalently u) so the input–output model (2.7) has zeros (or numerator dynamics), contrary to the continuous time model (1.1).

Since the relations (2.8) are linear and invertible, the original continuous-time parameters (M, D, K) can be recovered immediately from the parameters of the discretized model (2.7) by means of the linear transformation

$$M := \frac{h}{4} [M_d + K_d - D_d], \quad (2.10a)$$

$$D := M_d - K_d, \quad (2.10b)$$

$$K := \frac{1}{h} [M_d + K_d + D_d]. \quad (2.10c)$$

These are nice linear relations much in the spirit of what we wanted to achieve. Naturally, it must be kept in mind that the solution of (2.7) provides only an approximation of the sampled exact flow $t \mapsto q(t)$ at $t = kh$. It can be shown that the midpoint approximation is a second order method with an error of the order $O(h^2)$, see Marsden and West (2001, p. 402). The use of more sophisticated approximation schemes than (2.4) can however provide approximations of the exact flow of arbitrarily high order (see, e.g., Hairer, Lubich, & Wanner, 2005; Leokand & Shingel, 2011).

3. Equivalence with the Cayley–Tustin transform

We shall use hats to denote either Laplace or Z-transforms of time signals; specifically $\hat{q}(s)$ will be the Laplace transform of

$q(t)$; $t \in \mathbb{R}$ and $\hat{u}(z)$ the Z-transforms of the discrete sequence $\{u(kh)\}$. Accordingly, the Z-transform of $\{\mathbf{q}(k)\}$ will be denoted by $\hat{\mathbf{q}}(z)$. The bilinear substitution map

$$s = \frac{2z-1}{hz+1} \quad (3.1)$$

called the *Tustin transform*, is well-known in sampled-data control (see, e.g., Havu & Malinen, 2007; Jury, 1977). Given a continuous-time signal $f(t)$, $t \in \mathbb{R}$, having Laplace transform $\hat{f}(s)$, let us denote by f_T the discrete-time signal whose Z-transform, say $\hat{f}_T(z)$, is the Tustin transform obtained by the substitution (3.1) on $\hat{f}(s)$, namely

$$\hat{f}_T(z) = \hat{f} \left(\frac{2z-1}{hz+1} \right)$$

and call f_T simply the *Tustin transform of the continuous-time signal* f .

Proposition 1. *The discrete transfer function of the variational integrator defined by the midpoint rule (2.4)–(2.6), applied to the linear mechanical system (1.1), coincides with the Tustin transform of the transfer function:*

$$G(s) := [Ms^2 + Ds + K]^{-1} \quad (3.2)$$

of the continuous time system (1.1).

In other words, the discrete transfer function of the system described by the difference equation (2.7) with numerator polynomial defined by (2.9), having as input the signal $\{u(kh)\}$ and as output the signal $\{\mathbf{q}(k)\}$, is the Tustin transform of the transfer function of the system (1.1).

Proof. From the expression (3.2) it is immediate to check that

$$\begin{aligned} G(s)^{-1} \Big|_{s=\frac{2z-1}{hz+1}} &= M \frac{4(z-1)^2}{h^2(z+1)^2} + D \frac{2z-1}{hz+1} + K \\ &= [M_d z^2 + D_d z + K_d] \\ &\quad \times [(h/4)I_n(z^2 + 2z + 1)]^{-1}, \end{aligned} \quad (3.3)$$

where I_n is the $n \times n$ identity matrix. This is precisely the inverse transfer function of the model (2.7). \square

Hence, denoting, as in the proof, by $G(s)$ the transfer function of (1.1), we have,

$$\hat{\mathbf{q}}(z) = G_T(z) \hat{u}(z) \quad (3.4)$$

where $G_T(z)$ is the Tustin transform of $G(s)$. Therefore given the signals $\{\mathbf{q}(k)\}$ and $\{u(kh)\}$ one could in principle recover $G_T(z)$ and obtain the continuous time parameters of the system by inverting the Tustin transform, using the relations (2.10).

Remark. We would like to draw the reader's attention to the fact that formula (3.4) is quite different from the obvious relation obtainable by taking the Tustin transform of both members of the input–output relation (1.1), which would yield instead

$$\hat{q}_T(z) = G_T(z) \hat{u}_T(z)$$

a relation necessarily involving the Tustin transforms, $\hat{q}_T(z)$, $\hat{u}_T(z)$ of the continuous time signals $u(t)$ and $q(t)$, as defined earlier. Unfortunately a relation of this kind is quite useless for recovering models from realistic data. Since sampled input/output data are the only thing that can be acquired with finite speed devices available in reality, q_T and u_T should be computable starting from discrete time samples, $u(kh)$ and $q(kh)$, $k \in \mathbb{Z}$, of the continuous-time input and output functions $u(t)$ and $q(t)$, a problem which we insist in formalizing below,

Problem. From sampled values, say $f(kh)$, $k \in \mathbb{Z}$, of a continuous-time function $f(t)$; $t \in \mathbb{R}$, compute the discrete-time sequences f_T .

To our best knowledge this problem, elementary and simple to state as it may look, is unsolved. It is touched upon in Ref. Oppenheim and Johnson (1972) but, for several reasons, the approximation procedures described in this paper have revealed to be extremely difficult (if not practically impossible) to implement in practice. Of course one may well resort to “engineering common sense approximation” and approximate q_T and u_T by sampled data $q(kh)$ and $u(kh)$, or perhaps by sampling linear interpolation of these signals. This may mean approximating $G_T(z)$ by something close to the ordinary ZOH Z-transform of $G(s)$. These approximations may sometimes be acceptable, but there are no known bounds on the errors and the range of validity vis a vis the sampling process is unknown. In this respect, it is a common suggestion that a fast sampling rate should do. However as discussed in the introduction, this approach may lead to aliasing errors and ill-conditioning. In the same vein, obtaining a continuous-time model by inverse Tustin transform of a discrete one, as mentioned e.g. in Unbehauen and Rao (1990) is, as explicitly declared, an “indirect approach method” which is applied to a discrete-time model which is actually identified by standard discrete-time techniques based on ordinary sampled input–output signals. This may be seen as one of the common sense approximations mentioned above.

Since the accuracy of the parameter estimates is a crucial factor for the applications we have in mind, we insist in a, as far as possible, rigorous procedure.

Yet it is still not obvious how to get $\{q(k)\}$. Again, since the discrete system (2.7) is driven by the ordinary sampled input $u(kh)$ (and not by the Tustin transform of $u(t)$) the output signal $q(k)$ cannot be identified with the Tustin transform of the continuous-time response $G(s)\hat{u}(s)$. Below we shall provide an approximation result which can be useful for fast sampling rates.

Let $C_m(\mathbb{R}_+)$ denote the m -dimensional continuous functions on $\{t > 0\}$. Assume the input signal u belongs to $C_m(\mathbb{R}_+) \cap L_m^2(\mathbb{R}_+)$ and let $S_h : u \mapsto \{u(kh)\}_{k=1}^{+\infty}$ denote the sampling operator. Define G_T and G to be the time domain convolution operators corresponding, on \mathbb{Z}_+ , to the multiplication $G_T(z)\hat{u}(z)$, and on \mathbb{R}_+ , to $G(s)\hat{u}(s)$. The following proposition is adapted from Havu and Malinen (2007). It says that for fast sampling rate ($h \rightarrow 0$), the ordinary sampling of the continuous-time output $\hat{q}(s) = G(s)\hat{u}(s)$, provides a good approximation of the signal $\{q(k)\}$.

Proposition 2. *Let $G(s)$ be a stable transfer function. Assume the input signal u belongs to $C_m(\mathbb{R}_+) \cap L_m^2(\mathbb{R}_+)$ and that the sampled signal $\{u(kh)\}$ is a square summable sequence; i.e. belongs to $\ell_m^2(\mathbb{Z}_+)$. Then as $h \rightarrow 0$*

$$\|G_T S_h u - S_h G u\|_{\ell_m^2(\mathbb{Z}_+)} \rightarrow 0. \quad (3.5)$$

The proof will be given in the Appendix.

In our case the quality of the approximation will be enhanced by the low-pass character of the transfer function $G(j\omega)$ of the mechanical system. When the sampling frequency is chosen large enough with respect to the bandwidth of $G(j\omega)$, the product $G(j\omega)\hat{u}(j\omega)$ will be sensibly non zero only on a frequency band about $\omega = 0$ which is small relative to the sampling angular frequency $2\pi/h$. On this frequency band, by a well-known frequency domain relation between the Tustin and Laplace transforms, the difference between $G_T(e^{j\omega h})\hat{u}(e^{j\omega h})$ and $G(j\omega)\hat{u}(j\omega)$ will be negligible.⁴

⁴ In fact by a well-known frequency domain relation, $\hat{f}_T(e^{j\omega h}) = \hat{f}(j\frac{\omega}{h} \tan(\omega h/2))$ and $\omega \frac{\tan(\omega h/2)}{\omega h/2} \simeq \omega(1 + \frac{1}{3}(\frac{\omega h}{2})^2 + \frac{2}{5}(\frac{\omega h}{2})^4 + \dots) \simeq \omega$ up to third order terms in ω as $h \rightarrow 0$.

Concerning the validity of variational discretization, it should be remarked that it need not be a universally applicable discretization method as it may only be applicable to physical systems that contain energy storage elements and evolve according to a variational principle. It does not seem to be applicable to arbitrary dynamical systems. However by its very derivation, it does conserve (at least approximately) some basic physical quantities of the system. A case in point is that the discretization (2.4)–(2.6) is equivalent to the well-known Cayley transform which is known to preserve passivity (besides stability), a fact which was apparently first discovered by P. Faurre in 1973 (Faurre, 1973) and rediscovered in several later papers.

4. System identification

As discussed in the introduction, one may attempt to identify the parameters of the model (1.1) or (1.3) directly by using a continuous-time PEM method, say the continuous-time idgrey algorithm of the Matlab Sysid Toolbox (Ljung et al., 2011). However, unless very good quality initial parameter estimates are provided, this algorithm, for the particular model class under study, turns out to lack robustness and tend to behave erratically. As an alternative, we shall propose a discrete-time identification procedure based on the variational discretization described in Sections 2 and 3. Initial parameter estimates of the discrete-time variational discretization model (M_d, D_d, K_d) will be needed for this algorithm as well. These estimates are obtained by running a preliminary subspace identification method (n4sid) in discrete-time, followed by an appropriate structure transformation which converts the estimated state space model into the input–output structure of the variational discretization model (2.7). A discrete-time PEM procedure follows which computes refined estimates of the discrete-time parameters (M_d, D_d, K_d) . Finally the continuous-time estimates (M, D, K) are obtained by using the inverse (Tustin) transform (2.10). We maintain that this procedure has some advantages with respect to the classical d2c transform methods working under the assumption of ZOH (or FOH) discretization which suffer from the problems described in the introduction.

Ideal sampled data measurements from our system with sampling period h , are denoted

$$\begin{aligned} \{q(k) \in \mathbb{R}^n : k = 1, 2, \dots, N\}, \\ \{u(k) \in \mathbb{R}^m : k = 1, 2, \dots, N\} \end{aligned} \quad (4.1)$$

where $q(k) \equiv q(kh)$ and $u(k) \equiv u(kh)$ (we suppress the sampling interval h for ease of notation). The sample size N is supposed suitably large. Since the signal $\{q(k)\}$ will be always corrupted by additive measurement noise, it will be convenient to merge the measurement noise and the approximation errors of the ideal signal $\{q(k)\}$ and assume an output model of the form $y(k) = q(k) + n(k)$, where $n(k)$ is a white stationary noise. Hence we shall model our measurements as

$$y(k) = q(k) + n(k)$$

where $q(k)$ is the (noiseless) sampled output of the variational approximation system (2.7). We shall then write our noisy variational integrator model as a second order stochastic vector difference equation of the form

$$y(k) = A_1 y(k-1) + A_2 y(k-2) + B_0 f_d(k) + w(k) \quad (4.2)$$

where f_d is the filtered input defined in (2.9) which we use to avoid including the zero polynomial $h(1 + 2z^{-1} + z^{-2})/4$ in the input–output model. The noise process $\{w(k)\}$, is described by $w(k) = n(k) - A_1 n(k-1) - A_2 n(k-2)$ and is therefore colored.

The model parameters are

$$A_1 := -M_d^{-1}D_d, \quad A_2 := -M_d^{-1}K_d, \quad B_0 := M_d^{-1}. \quad (4.3)$$

Clearly the term B_0 represents the direct coupling between $f_d(k)$ and $\mathbf{q}(k)$ in the midpoint approximation model (2.7).

Note that (4.2) is a so-called *output error* (OE) model (Ljung, 1999), whose predictor depends non-linearly on the parameters and gives rise to a nonlinear estimation problem. For OE models a standard choice of identification algorithm is the “grey box” PEM algorithm `idgrey` described in the MATLAB System Identification Toolbox guide (Ljung et al., 2011), which can in principle incorporate various structural constraints on the system parameters such as, e.g., the symmetry of various matrices. Since the discrete-time PEM algorithm is also very sensitive to the choice of initial values for the parameters, good initial parameter estimates are necessary. In the next subsection we describe an initialization procedure which is tailored to the specific structure of the model (2.7).

4.1. The initialization step

In order to compute the initial estimates a rather common choice is to run a preliminary subspace algorithm, say the `n4sid` algorithm, directly on the input–output data ($y(k)$, $f_d(k)$) (or ($y(k)$, $u(k)$)). Any such subspace algorithm yields a discrete innovation model of the type

$$x(k+1) = Fx(k) + Gf_d(k) + Ke(k) \quad (4.4)$$

$$y(k) = Hx(k) + Jf_d(k) + e(k), \quad (4.5)$$

of which we are interested only in the parameters (F, G, H, J). Since for OE models the state–output dynamics of the stochastic and deterministic subsystems are the same, one can get a minimal realization of the deterministic part simply by setting $e(k) \equiv 0$ in (4.4), see Ljung et al. (2011). Also, since we have a priori knowledge of the number of degrees of freedom, order estimation is not necessary and the algorithm can be pre-set to return a $2n$ -dimensional discrete realization (F, G, H).

We shall now describe a procedure to compute a preliminary estimate of the parameters M_d, D_d, K_d from those of the model (4.4), identified by the subspace algorithm. To this end we shall need to recall the following result; see e.g. Meyer and Srinivasan (1996).

Lemma 3. *A necessary and sufficient condition for a $2n$ dimensional discrete state space model (1.4), with $\dim y(k) = n$, to have an input–output relation described by a second order vector difference equation is that*

$$\text{rank} \begin{bmatrix} H \\ HF \end{bmatrix} = 2n. \quad (4.6)$$

In practice, for our system identified by a subspace method, the matrix $\Omega := \begin{bmatrix} H \\ HF \end{bmatrix}$ will almost always be of rank $2n$ (i.e., invertible). Using Ω^{-1} as a similarity transformation, one gets the block-companion form

$$\mathbf{F} := \Omega F \Omega^{-1} = \begin{bmatrix} 0 & I \\ \mathbf{F}_{21} & \mathbf{F}_{22} \end{bmatrix}, \quad \mathbf{H} := H \Omega^{-1} = [I \quad 0],$$

$$\mathbf{G} := \Omega G = \begin{bmatrix} HG \\ HFG \end{bmatrix} := \begin{bmatrix} \mathbf{G}_1 \\ \mathbf{G}_2 \end{bmatrix},$$

where the blocks \mathbf{F}_{21} and \mathbf{F}_{22} can be computed by solving the equation

$$HF^2 = [\mathbf{F}_{21} \quad \mathbf{F}_{22}] \Omega. \quad (4.7)$$

Hence, the identified deterministic model⁵ can generically be transformed to the block-companion structure

$$x(k+1) = \begin{bmatrix} 0 & I \\ \mathbf{F}_{21} & \mathbf{F}_{22} \end{bmatrix} x(k) + \begin{bmatrix} \mathbf{G}_1 \\ \mathbf{G}_2 \end{bmatrix} f_d(k) \quad (4.8)$$

$$y(k) = [I \quad 0] x(k) + Jf_d(k) \quad (4.9)$$

which corresponds to an estimated I/O difference equation, with full numerator dynamics

$$y(k) = \mathbf{A}_1 y(k-1) + \mathbf{A}_2 y(k-2) + \mathbf{B}_0 f_d(k) + \mathbf{B}_1 f_d(k-1) + \mathbf{B}_2 f_d(k-2) \quad (4.10)$$

where $\mathbf{A}_1 = \mathbf{F}_{22}$, $\mathbf{A}_2 = \mathbf{F}_{21}$ and

$$\mathbf{B}_0 = J, \quad \mathbf{B}_1 = \mathbf{G}_1 - \mathbf{F}_{22}J, \quad \mathbf{B}_2 = \mathbf{G}_2 - \mathbf{F}_{22}\mathbf{G}_1 - \mathbf{F}_{21}J.$$

Note that, unless both estimates $\mathbf{B}_1, \mathbf{B}_2$ turn out to be zero, the identified state space model will not have the right structure to yield the AR-type input–output representation of the form (2.7). It is not hard to check that a block-companion state-space structure yielding an I/O difference equation of the form (2.7), is

$$\bar{x}(k+1) = \begin{bmatrix} 0 & I \\ \mathbf{F}_{21} & \mathbf{F}_{22} \end{bmatrix} \bar{x}(k) + \begin{bmatrix} 0 \\ G \end{bmatrix} f_d(k+2) \quad (4.11)$$

$$y(k) = [I \quad 0] \bar{x}(k),$$

where the filtered input f_d has been shifted two steps ahead for convenience, so that the direct coupling term (\mathbf{B}_0) is equal to G .

In order to get reasonable initial estimates of the parameters we shall impose equality of the gain matrices and of the AR dynamics of the two transfer functions. The model (4.11) is defined in terms of three matrix parameters ($\mathbf{F}_{21}, \mathbf{F}_{22}, G \equiv \mathbf{B}_0$) which can be computed by equating the steady state gains and the dynamics of (4.10) and of (4.11), getting

$$G^0 = \mathbf{B}_0 + \mathbf{B}_1 + \mathbf{B}_2, \quad \mathbf{F}_{21}^0 = \mathbf{F}_{21}, \quad \mathbf{F}_{22}^0 = \mathbf{F}_{22}. \quad (4.12)$$

4.2. The overall algorithm

The matrices (4.12) are used as initial values for the parameter updating recursion of a discrete-time PEM algorithm run on the block-companion structure (4.11). The final parameter estimates (M_d, D_d, K_d) can then easily be recovered from the identified state space model (4.11) by solving the one-to-one relations

$$F_{21}^{\text{PEM}} = -M_d^{-1}K_d, \quad F_{22}^{\text{PEM}} = -M_d^{-1}D_d, \quad G^{\text{PEM}} = M_d^{-1}. \quad (4.13)$$

In order to keep track of the structural properties of the parameters like the symmetry of the matrices, it is actually advisable to parametrize the model directly in terms of (M_d, D_d, K_d). In this way we need to update only $n(n+1)/2$ parameters for each matrix. The scheme in Fig. 1 summarizes the various steps.

- (1) Acquire sampled data $\{f_d(k), y(k)\}$ from the continuous time mechanical linear system (1.3). The sampling time h is assumed small enough so that $\{q(k)\}$ is a good approximation of the ideal signal $\{\mathbf{q}(k)\}$.
- (2) Perform `n4sid` identification from the data $\{f_d(k), y(k)\}$ and get the system (4.4); (Σ^d in Fig. 1). This system has no specific structure.
- (3) Impose the condition (4.12) and compute initial estimates of the matrices M_d, D_d, K_d , say $\hat{M}_d^{\text{init}}, \hat{K}_d^{\text{init}}, \hat{D}_d^{\text{init}}$, which (after symmetrization) are taken as initialization point for the constrained optimization procedure based on PEM.

⁵ In order to save notations we still denote the output of the deterministic subsystem by the symbol y . This should cause no confusion.

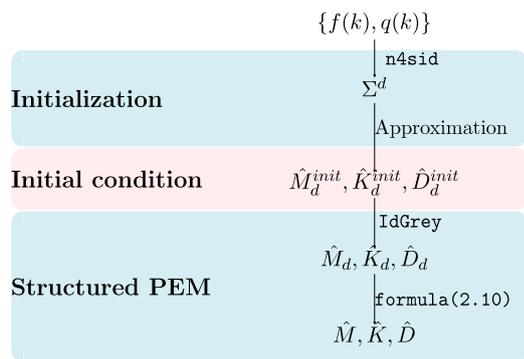


Fig. 1. Scheme of the identification procedure.

- (4) Run the PEM algorithm initialized with these parameters using the data set $\{f_d(k), y(k)\}$ imposing symmetry of the estimates \hat{M}_d, \hat{D}_d , and \hat{K}_d .
- (5) Use the inverse transformation (2.10) to convert the discrete estimates into continuous time estimates \hat{M}, \hat{D} , and \hat{K} . Positive definiteness may be checked throughout or as a last step.

Since the algorithm is essentially an implementation of a PEM method, we should expect consistency and asymptotic efficiency in case of Gaussian additive noise. Due to the linearity of relation (2.10), consistency will hold also for the continuous-time parameter estimates computed in step (5). The stochastic error on the continuous-time parameters depends, besides the intrinsic sensitivity of the model class to parameter variations (Cramèr–Rao bound), on the conditioning of the discrete-to-continuous transform. This random error naturally worsens as the width of the sampling interval tends to zero, although to a different extent than, say, for the ZOH discretization. Besides this, one key advantage of the proposed identification method is that the linear transformation (2.10) automatically transforms the symmetric matrices M_d, K_d , and D_d into their symmetric continuous counterparts, acting effectively on $3n \times (n + 1)/2$ parameters while a generic `d2c` transform would operate instead on an unstructured discrete-time 4-tuple (F, G, H, J) and would make (M, D, K) to inherit errors present in $4n^2$ parameters.

5. Simulation results

We consider an eight degrees of freedom system described in De Angelis et al. (2002) with stiffness and damping matrices K and D given in Table 1 and a diagonal mass matrix M with entries equal to 100 kg. See Table 1. Eight sensors measure and record the sampled displacements of the eight point masses with independent (white) measurement noise with SNR of 20 dB. Note that such a noise mimics a difficult realistic condition. An array of $n \times N = 8 \times 3300$ data points is collected and used for system identification. The sampling time interval is varied within the range corresponding to sampling frequencies from 100 to 1000 Hz. This seems to be the most significant frequency range in applications. The eight inputs are obtained as successive chunks of a continuous time signal obtained by cubic spline interpolation of the samples of a normalized (pseudo-random) discrete Gaussian white noise, see Fig. 2. The amplitude is varied according to the desired SNR.

5.1. Comparisons

System identification based on continuous-time optimization methods does not seem to be popular in the mechanical/civil engineering literature. What seems to be the state of the art in structural engineering identification is the procedure described in

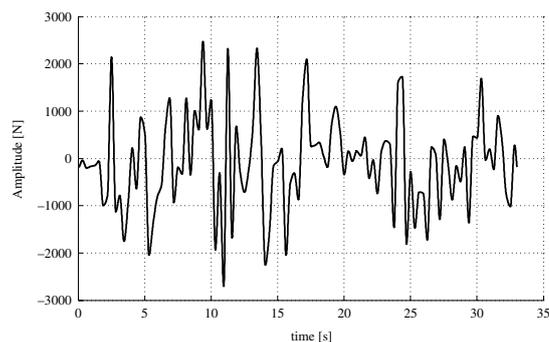


Fig. 2. Input function.

Lus et al. (2002) and Lus, De Angelis, Betti, Longman (2003), see also De Angelis et al. (2002) and Lus, De Angelis, Betti (2003) which is based on a first step of discrete subspace identification, followed by a standard `d2c` step and then by a suitable “projection” procedure in continuous time which is needed to enforce the second order structure (1.1) upon the identified continuous-time system.

We shall henceforth compare this method with our variational-integrator subspace-based algorithm described in Section 4.1, followed by the inverse transformation (2.10) to convert the discrete estimates $\hat{M}_d^{init}, \hat{K}_d^{init}, \hat{D}_d^{init}$, into continuous time estimates $\hat{M}, \hat{D}, \hat{K}$. Recall that in our algorithm this conversion step is not implemented but rather the discrete estimates $\hat{M}_d^{init}, \hat{K}_d^{init}, \hat{D}_d^{init}$ are to provide the initial conditions to the PEM refinement step.

First, a comparison of the two subspace-based identification methods is presented for the benchmark problem with 8 degrees of freedom described above. Although the projection step in Lus et al. (2002) and Lus, De Angelis, Betti, Longman (2003) is of course an essential ingredient of the method, this “state of the art” method will be simply referred to as `n4sid + d2c`.

To compare the quality of the identification, averages of 50 Monte Carlo runs of the relative errors of the estimated continuous time parameters are computed for different values of h . Fig. 3 shows the comparison of the state-of-the-art `n4sid + d2c` method with our proposed method. By looking at the magnitude of the relative errors, it is clear that there is an advantage in using the variational approach, even without the PEM refinement. In particular the estimation errors on the matrix D are always about double of those with the variational approach. For small h much higher relative errors are observed for the `n4sid + d2c` approach. The behavior for small sampling intervals is actually in accordance with the discussion in the introduction.

5.2. Effect of the PEM refinement

In Fig. 4 we compare the results of our proposed full variational procedure with the above state-of-the-art `n4sid + d2c` method. Although it may look a bit “unfair” this serves to appreciate, besides the expected improvement on the quality of the estimates, the fact that the PEM step always terminates in all of the 50 Monte Carlo runs in spite of random errors on the initial estimates.

Some referees have asked to add a PEM step to the `n4sid + d2c` approach for a “fair” comparison with the results of our full variational algorithm. This combined simulation, adding a continuous-time PEM step based on the `n4sid + d2c` estimates, has led to rather disappointing results. We have decided not to show the plots here since they leave the reader wondering how to interpret the results, if either due to larger errors in the subspace estimates of the `n4sid + d2c` approach or to the high sensitivity of the successive continuous PEM step. We have instead chosen to add a “somewhat artificial” simulation experiment which is meant to be a study of the sensitivity to errors in the initial estimates of

Table 1
Numerical example. Stiffness matrix K and damping matrix D for the 8 d.o.f. system.

$$K = \begin{bmatrix} 27071.1 & 0 & 0 & 0 & -10000 & 0 & -3535.5 & -3535.5 \\ 0 & 17071.1 & 0 & -10000 & 0 & 0 & -3535.5 & -3535.5 \\ 0 & 0 & 27071.1 & 0 & -3535.5 & 3535.5 & -10000 & 0 \\ 0 & -10000 & 0 & 17071.1 & 3535.5 & -3535.5 & 0 & 0 \\ -10000 & 0 & -3535.5 & 3535.5 & 27071.1 & 0 & 0 & 0 \\ 0 & 0 & 3535.5 & -3535.5 & 0 & 17071.1 & 0 & -10000 \\ -3535.5 & -3535.5 & -10000 & 0 & 0 & 0 & 27071.1 & 0 \\ -3535.5 & -3535.5 & 0 & 0 & 0 & -10000 & 0 & 17071.1 \end{bmatrix}$$

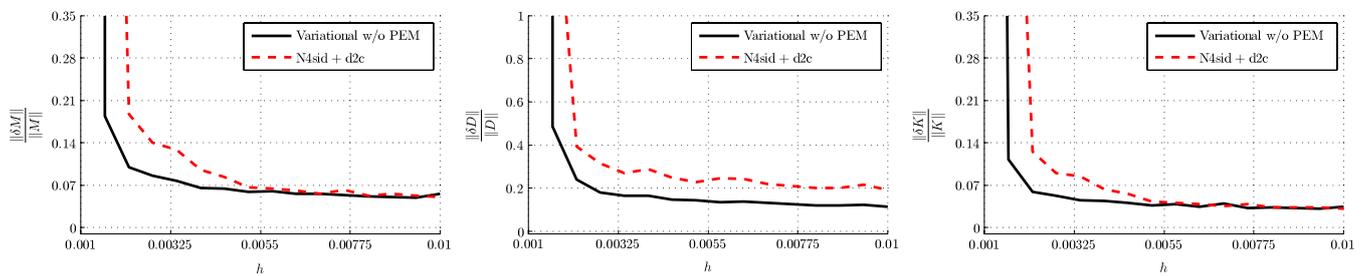
$$D = \begin{bmatrix} 136.4 & 0 & 0 & 0 & -50 & 0 & -17.7 & -17.7 \\ 0 & 86.4 & 0 & -50 & 0 & 0 & -17.7 & -17.7 \\ 0 & 0 & 136.4 & 0 & -17.7 & 17.7 & 50 & 0 \\ 0 & -50 & 0 & 86.4 & 17.7 & -17.7 & 0 & 0 \\ -50 & 0 & -17.7 & 17.7 & 136.4 & 0 & 0 & 0 \\ 0 & 0 & 17.7 & -17.7 & 0 & 86.4 & 0 & -50 \\ -17.7 & -17.7 & -50 & 0 & 0 & 0 & 136.4 & 0 \\ -17.7 & -17.7 & 0 & 0 & 0 & -50 & 0 & 86.4 \end{bmatrix}$$


Fig. 3. Comparison of relative errors on subspace estimates of M , D , and K for the 8 d.o.f. system.

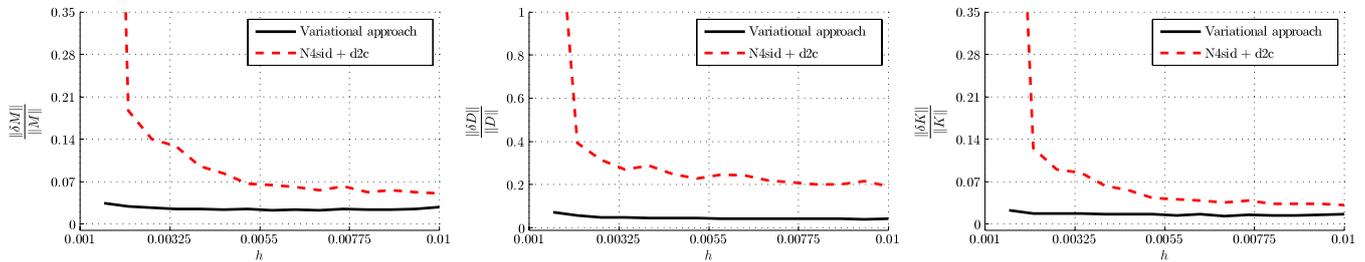


Fig. 4. Comparison of relative errors on variational estimates of M , D , and K for the 8 d.o.f. system.

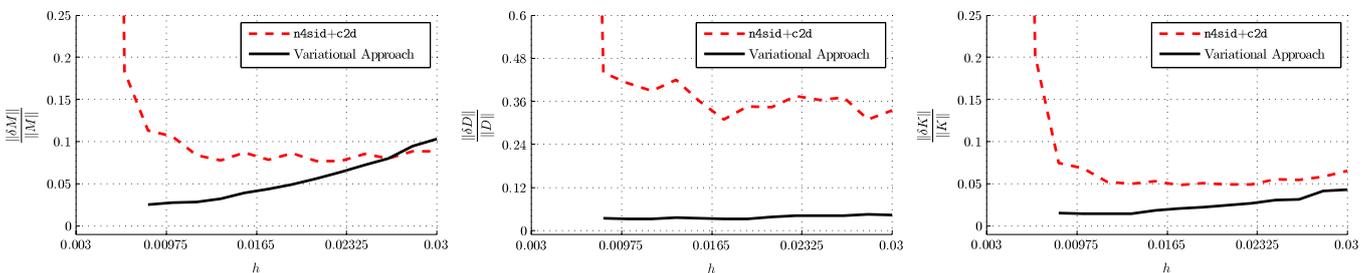


Fig. 5. Relative error on M , D , and K for the 8 d.o.f. system, large h .

the continuous-time PEM, for the particular model class under study (as of course, we are not in a position to make more general statements here).

Below are the results obtained by 30 Monte Carlo runs on 6000 data points of the continuous-time idgrey algorithm with ZOH sampling, applied to the eight degrees of freedom model, using as initial data the true parameters (M, D, K) perturbed by

Gaussian zero-mean noise of standard deviations equal to 10% and 12% of the respective true Frobenius norms. As it can be seen, even with relatively small perturbations on the true values, a good percentage of the CT idgrey runs end in local minima providing unstable or grossly wrong models. With higher standard deviation of the perturbations quite often the algorithm does not even terminate. Note from our Monte Carlo simulations in Fig. 3,

that the errors with the Lus–De Angelis–Betti–Longman method can be well above 10 or 12%, especially for small sampling period.

h	% of fails with 10% perturbation (%)	% of fails with 12% perturbation (%)
0.005	7	26
0.011	16	21
0.017	20	10
0.021	33	21

5.2.1. Algorithm behavior for increasing h

Naturally for slow sampling rates the approximation of $\{q(k)\}$ by the sampled continuous signals may be poor. Errors due to a slow sampling rate are however common to all identification schemes and may not concern us since here we are assuming fast sampling. Nevertheless, just to give an idea of the order of magnitude of these errors, in this section we shall show the effect of approximating the Tustin transformation with the sampled signals when h is not small. In Fig. 5 the relative errors for the matrices M and D are shown for a range of sampling times including larger values of h than in the previous simulation. We use a somewhat high SNR of 25 dB, as we are interested in highlighting the distortion due to the discretization rule rather than that due to the noise. A gradual increase of the relative error is evident for the matrix M up to a point when the `n4sid` and `d2c` procedure gives better results for large values of h . Interestingly, for the matrix D the good quality of the estimate seems to persist.

6. Conclusions and future work

A new identification procedure for the second order equations of a linear mechanical system has been described which appears to lead to an accurate recovery of the continuous time mechanical parameters. The proposed procedure is based on the theory of variational integrators and on a rather elementary discretization rule (the midpoint rule). Simulations (not all of which have been reported in the paper) confirm that the relative error with the proposed method is significantly smaller than that obtained by state-of-the art methods. The key point of the variational discretization technique is to enforce the mechanical structure on the discrete-time model from first principles rather than by ad hoc methods. For this reason, variational integrators theory seems to provide an interesting and rather general point of view which may even allow to attack identification of *nonlinear* mechanical systems. More elaborate variational discretizations are possible; say higher order methods which however lead to more complicated models and have not been considered here.

The theory presented here applies without difficulties, also to more realistic linear sensing schemes such as measurements of combinations of configuration q , velocities \dot{q} and acceleration, in place of the full configuration q . The analysis of these more general schemes is left to future publications.

Appendix. Proof of Proposition 2

Our main reference for the proof will be Theorem 4.3 in Havu and Malinen (2007) which is stated for more general input functions $u \in L_m^2(\mathbb{R}_+)$. On this space the pointwise sampling operator S_h is not defined and an “average” sampling operator $T_h : L_m^2(\mathbb{R}_+) \rightarrow \ell_m^2(\mathbb{Z}_+)$ defined as

$$(T_h u)(k) := \frac{1}{\sqrt{h}} \int_{(k-1)h}^{kh} u(t) dt$$

is utilized instead (equivalently, in Havu and Malinen (2007) the authors define $T_h u$ as the Z-transform of the sampled sequence). Theorem 4.3 in Havu and Malinen (2007) states that, as $h \rightarrow 0$,

$$\|T_h^* G_T T_h u - G u\|_{L_m^2(\mathbb{R}_+)} \rightarrow 0.$$

Note that the Fourier transform map \mathcal{L} on $L_m^2(\mathbb{R}_+)$ appearing in formula (4.2) of Ref. Havu and Malinen (2007) is unnecessary as G operates in the time domain. Now the adjoint, T_h^* , is $\frac{1}{\sqrt{h}}$ times the zero-order hold operator on the intervals $((k-1)h, kh]$ and for all $v \in \ell_m^2(\mathbb{Z}_+)$, we have $\|T_h^* v\|_{L_m^2(\mathbb{R}_+)} = \|v\|_{\ell_m^2(\mathbb{Z}_+)}$, so that T_h^* is isometric. Hence we also have

$$\|G_T T_h u - T_h G u\|_{\ell_m^2(\mathbb{Z}_+)} \rightarrow 0 \quad (A.1)$$

for all $u \in L_m^2(\mathbb{R}_+)$. Now, by assumption, $u \in C_m(\mathbb{R}_+) \cap L_m^2(\mathbb{R}_+)$ and $S_h u \in \ell_m^2(\mathbb{Z}_+)$ (this certainly happens if, for example, $\dot{u} \in L_m^2(\mathbb{R}_+)$ or u is band-limited) and it is trivial to see that, under these assumptions, letting f represent either u or $G u$, that

$$\|(T_h - S_h)f\|_{\ell_m^2(\mathbb{Z}_+)} \rightarrow 0,$$

as $h \rightarrow 0$. Therefore, (3.5) follows from (A.1) and the triangle inequality. \square

References

- De Angelis, M., Lus, H., Betti, R., & Longman, R. W. (2002). Extracting physical parameters of mechanical models from identified state-space representations. *Transactions of the ASME. Journal of Applied Mechanics*, 69, 617–625.
- Dieci, L., & Papini, A. (2000). Conditioning and Padé approximation of the logarithm of a matrix. *SIAM Journal on Matrix Analysis and Applications*, 21, 913–930.
- Faurre, P. (1973). *Réalisations Markoviennes de Processus Stationnaires. Rapport de recherche N. 13, INRIA, Roquencourt, France.*
- Feuer, Arie, & Goodwin, Graham (1996). *Sampling in digital signal processing and control.* Basel: Birkhäuser.
- Garnier, Hugues, & Wang, Liuping (Eds.) (2008). *Identification of continuous-time models from sampled data.* London: Springer.
- Hairer, E., Lubich, C., & Wanner, G. (2005). *Geometric numerical integration: structure preserving algorithms for ordinary differential equations* (2nd ed.). Berlin: Springer-Verlag.
- Havu, V., & Malinen, J. (2007). The Cayley transform as a time discretization scheme. *Numerical Functional Analysis and Optimization*, 28, 825–851.
- Heuberger, P. S. C., de Hoog, T. J., van den Hof, P. M., & Wahlberg, B. (2003). Orthonormal basis functions in time and frequency domain: Hambo transform theory. *SIAM Journal on Control and Optimization*, 42, 1347–1373.
- Jury, E. I. (1977). *Sampled data control systems.* Melbourne, FL, USA: Krieger Publishing Co. Inc.
- Laub, A., & Arnold, W. F. (1984). Controllability and observability criteria for linear second-order models. *IEEE Transactions on Automatic Control*, AC-29, 163–165.
- Leokand, M., & Shingel, T. (2011). Prolongation-collocation variational integrators.
- Ljung, L. (1999). *System identification-theory for the user* (2nd ed.). Upper Saddle River, NJ: Prentice-Hall.
- Ljung, L., et al. (2011). *System identification toolbox; version R2011b users guide.* Natick, MA: The Mathworks.
- Lus, H., De Angelis, M., & Betti, R. (2003). A new approach for reduced order modeling of mechanical systems using vibration measurements. *Transactions of the ASME. Journal of Applied Mechanics*, 70, 715–723.
- Lus, H., De Angelis, M., Betti, R., & Longman, R. W. (2002). Constructing second-order models of mechanical systems from identified state-space realizations, part II: numerical investigations. *ASCE Journal of Engineering Mechanics*, 129, 489–501.
- Lus, H., De Angelis, M., Betti, R., & Longman, R. W. (2003). Constructing second-order models of mechanical systems from identified state-space realizations, part I: theoretical discussions. *ASCE Journal of Engineering Mechanics*, 129, 477–488.
- Mahata, K., & Fu, M. (2007). On the indirect approaches for CARMA model identification. *Automatica*, 43, 1457–1463.
- Marsden, J. E., & West, M. (2001). Discrete mechanics and variational integrators. *Acta Numerica*, 357–514. Cambridge.
- Meyer, D. G., & Srinivasan, S. (1996). Balancing and model reduction for second order form linear systems. *IEEE Transactions on Automatic Control*, 41, 1632–1644.
- Ohta, Y., & Kawai, T. (2004). Continuous-time subspace system identification using generalized orthonormal basis functions. In *Proceedings of the 16th international symposium on the mathematical theory of networks and systems, MTNS-2004*, Leuven.
- Oppenheim, A. V., & Johnson, D. H. (1972). Discrete representations of signals. *Proceedings of the IEEE*, 60, 681–691.
- Sinha, N. K., & Rao, G. P. (1991). *Identification of continuous-time systems.* Dordrecht, Amsterdam: Kluwer Academic.
- Söderström, T. (1991). Computing stochastic continuous-time models from ARMA models. *International Journal of Control*, 53, 1311–1326.

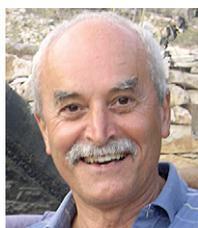
Söderström, T., Fan, H., Carlsson, B., & Bigi, S. (1997). Least squares parameter estimation of continuous-time ARX models from discrete-time data. *IEEE Transactions on Automatic Control*, 42(5), 659–673.

Unbehauen, H., & Rao, G. P. (1990). Continuous-time approaches to system identification: a survey. *Automatica*, 26(1), 23–35.

Veselov, A. P. (1988). Integrable discrete-time systems and difference operators. *Funktsional'nyi Analiz i ego Prilozheniya, Funct. Anal. Appl.*, 22, 1–13.



Mattia Bruschetta was born in Treviso, Italy on July, 2nd, 1982. He received a Ph.D. degree in control system theory in 2011 from the University of Padua, Padua, Italy. He is now research fellow at University of Padua focused on Model Predictive Control application. His research interests include Model Predictive Control, Motion Cueing, Virtual Prototyping, System Identification, Automotive.



Giorgio Picci holds a Dr. Eng. degree from the University of Padova where he graduated in 1967. Currently he is Professor Emeritus with the Department of Information Engineering, University of Padova, Italy. He has held several long-term visiting appointments with various American, Japanese and European universities among which Brown University, M.I.T., the University of Kentucky, Arizona State University, the Center for Mathematics and Computer Sciences (C.W.I.) in Amsterdam, the Royal Institute of Technology, Stockholm Sweden, Kyoto University

and Washington University in St. Louis, Mo. He has been contributing to Systems and Control mostly in the area of modeling, estimation and identification of stochastic systems and published over 100 papers and edited three books in this area. He has been involved in various joint research projects with industry and state agencies. He has been chairman of the IFAC Technical Committee on Stochastic Systems, past member of the EUCA council, project manager of the Italian team for the Commission of the European Communities Network of Excellence System Identification (ERNSI) and general coordinator of the Commission of European Communities IST project RECSYS, in the fifth Framework Program. Giorgio Picci is a Life Fellow of the IEEE, a Fellow of IFAC and a foreign member of the Swedish Royal Academy of Engineering Sciences.



Alessandro Saccon is an Assistant Professor at the Department of Mechanical Engineering, Eindhoven University of Technology, The Netherlands. He received a Ph.D. degree in control system theory in 2006 from the University of Padua, Padua, Italy, with a scholarship by MSC Software Corporation. From 2006 to 2009, he held a research and development position at the Department of Information Engineering of the University of Padua in joint collaboration with the racing motorcycle company Ducati Corse, working on control and optimization methods for the exploration of the dynamics of racing motorcycles for virtual

prototyping studies. He has then been a Postdoctoral Scholar at the Institute for Systems and Robotics, Instituto Superior Tecnico, Lisbon, Portugal from September 2009 until January 2013. There, he has focused on the development of optimal constrained motion planning strategies for multiple autonomous robotic vehicles. His research interests include modeling, analysis, identification, and control of mechanical systems, nonlinear control theory, and numerical optimal control methods for the exploration of the trajectory space of complex and highly maneuverable mechanical systems.