

Hybrid evolutionary identification of output-error state-space models

Vasilis K. Dertimanis^{*}, Eleni N. Chatzi^a and Minas D. Spiridonakos^b

*Institute of Structural Engineering, Department of Civil, Environmental and Geomatic Engineering,
ETH Zürich, Stefano-Franscini-Platz 5, 8093 Zürich, Switzerland*

(Received October 17, 2014, Revised November 20, 2014, Accepted December 8, 2014)

Abstract. A hybrid optimization method for the identification of state-space models is presented in this study. Hybridization is succeeded by combining the advantages of deterministic and stochastic algorithms in a superior scheme that promises faster convergence rate and reliability in the search for the global optimum. The proposed hybrid algorithm is developed by replacing the original stochastic mutation operator of Evolution Strategies (ES) by the Levenberg-Marquardt (LM) quasi-Newton algorithm. This substitution results in a scheme where the entire population cloud is involved in the search for the global optimum, while single individuals are involved in the local search, undertaken by the LM method. The novel hybrid identification framework is assessed through the Monte Carlo analysis of a simulated system and an experimental case study on a shear frame structure. Comparisons to subspace identification, as well as to conventional, self-adaptive ES provide significant indication of superior performance.

Keywords: Evolution strategies; Levenberg-Marquardt; hybrid optimization; system identification; state-space; modal analysis

1. Introduction

Recent developments in structural health monitoring and maintenance have advanced the promotion of inverse engineering methods and, in specific, of structural identification techniques (Friswell 2007, Fritzen *et al.* 2013, Catbas *et al.* 2013, Nagarajaiah and Basu 2009). The latter usually estimate a mathematical model of the structure and subsequently extract critical quantities (e.g., natural frequencies, mode shapes, etc.), which are then used for the formulation and monitoring of damage or nonlinearity features (Lin and Betti 2004, Papadimitriou *et al.* 2012, Wu and Kareem 2013). It is thus apparent that estimation accuracy becomes an issue of utmost importance, as it forms a major component in the process of decision making for life-cycle assessment of engineered systems.

Under this perspective, estimation accuracy has always comprised a focal point in the implementation of numerical optimization methods for parametric system identification. This is

^{*}Corresponding author, Researcher, E-mail: v.derti@ibk.baug.ethz.ch

^a Professor, E-mail: chatzi@ibk.baug.ethz.ch

^b Researcher, E-mail: spyridonakos@ibk.baug.ethz.ch

especially true for transfer function models in both SISO and MIMO representations, for the estimation of which a wide variety of deterministic (Ljung 1999, Verhaegen and Verdult 2007), stochastic (Koulocheris *et al.* 2003) and hybrid (Koulocheris *et al.* 2008) optimization algorithms have been investigated. On the contrary, numerical optimization for the fitting of state-space models to vibration data is significantly less explored in the literature. This is mainly attributed to the current practice that pertains to the implementation of subspace methods (Kim and Lynch 2012, Katayama 2005, Juang and Pappa 1985) for the identification tasks. While they are considerably simpler and faster, these methods are suboptimal and amenable to potential inaccuracies (Dahlén *et al.* 1998), especially at low signal-to-noise-ratios.

Maximum likelihood estimation is further tool that can serve in such a purpose, yet it is known that the commonly used numerical methods suffer from significant drawbacks, such as instabilities, local extrema and heavy dependence on initial guesses. In the state-space model class these problems are even more severe due to the inherent over-parameterizations that lead to surjective mappings. In an effort to face these issues, the research has focused on the development of gradient projection methods (McKelvey *et al.* 2004, McKelvey and Helmersson 1997, Gibson and Ninness 2005, Bergboer *et al.* 2002). However, the proposed algorithms inherit the limitations of the deterministic ones, while assessments in structural systems are still missing.

During the last two decades and following the seminal work of Kristinsson and Dumont (Kristinsson and Dumont 1992), an alternative approach for attaching the nonlinear estimation problem has been developed, relying on the implementation of stochastic optimization algorithms (Yu and Gen 2010, Fleming and Purshouse 2002). However, the majority of these studies are mostly limited to SISO/MIMO transfer functions. Corresponding investigations on state-space models have only recently been reported in Dertimanis (2014) where four distinct instances of ES are adapted and benchmarked. The presented analysis showed that well-reported advantages of ES (increased efficiency in continuous problems, self-adaptivity, low algorithmic complexity, etc.) can be well fitted into this type inverse problems. Following this work, Dertimanis and Chatzi (2014) subsequently performed an initial investigation on the implementation of hybrid optimization schemes that combine diverse classes of optimization algorithms, such as deterministic and stochastic (Dertimanis *et al.* 2003, Koulocheris *et al.* 2008), to the state-space model estimation problem. It is noted that, other classes of stochastic optimization algorithms that have been used in the past for inverse analysis (Casciati 2008, Tang *et al.* 2008), usually formulate an explicit optimization problem in terms of the system's physical parameters (mass, stiffness and damping). In this sense, the algorithm proposed herein provides a wider implementation framework in that it may readily be applied in the discussed optimization problems.

Following this spirit, the aim of this study is to introduce an enhanced version of the early hybrid optimization algorithm presented in Dertimanis and Chatzi (2014) for the estimation of state-space models, and to test its robustness via thorough investigation of its performance using both simulated and experimental data. This novel scheme maintains all the significant features of its initial counterpart, interconnecting stochastic and deterministic optimization algorithms, in a way that exploits the advantages of both and results into a powerful method that delivers a faster convergence rate, as well as increased reliability in the search for the global optimum. Among the class of Evolutionary Algorithms (EAs), the stochastic component has been selected to be the self-adaptive Evolution Strategy (ES), notated as $(\mu/\rho (+/,) \lambda) - ES$, in which μ , ρ and λ denote the number of parents, recombination parents and offspring, respectively. This selection is justified by existing indications concerning ES's superior performance in similar problems (Koulocheris *et al.* 2003), against Genetic Algorithms (GAs) (Dimou and Koumouis 2003) and Evolutionary

Programming (EP). In the current version of the hybrid algorithm, the original recombination and selection operators of the conventional ES remain unaltered, while the self-adaptive mutation is replaced by the Levenberg-Marquardt quasi-Newton algorithm. The deterministic mutation acts only on a subset ν of *non-privileged* individuals (e.g., the worst ν of the population during every iteration, where ν is an additional exogenous parameter of the hybrid algorithm) and contributes to the development of a robust and flexible scheme, henceforth notated as $(\mu/\rho (+/), \lambda, \nu)$ -h-ES and abbreviated as h-ES.

Accordingly, an estimation procedure is adopted that pertains to the estimation and the evaluation of successive state orders over a predefined range. Taking into account previous investigations (Dertimanis and Chatzi 2014), in which prior information has been judged as necessary, a parameter vector using subspace techniques is initially estimated and used as a centroid individual, on the basis of which the initial parental population is generated. Upon termination of the order selection process and the determination of a final candidate state-space model, the model validation process integrates a physically meaningful dispersion analysis procedure (Dertimanis 2013) that quantifies the significance of each identified vibration mode. The proposed method is assessed by a Monte Carlo study of a simulated structure, as well as an identification study of an experimental structure. In the former case, a linear model of a four-story shear building subject to uniaxial horizontal base excitation (ground acceleration) is used, while in the latter a laboratory steel frame structure is identified.

The rest of the paper is organized as follows: the structural identification problem is described in Section 2 and the proposed hybrid algorithm is expanded at Section 3. Section 4 focuses on the estimation procedure, while Section 5 contains the identification experiments. Finally, Section 6 summarizes the results and addresses future perspectives.

2. The structural identification problem

2.1 The structure in discrete-time state-space format

A structural system with n degrees of freedom (DOF) can be represented by a second-order vector differential equation in the form

$$\mathbf{M}\ddot{\mathbf{q}}(t) + \mathbf{H}\dot{\mathbf{q}}(t) + \mathbf{K}\mathbf{q}(t) = \mathbf{P}\mathbf{f}(t) \quad (1)$$

where \mathbf{M} , \mathbf{H} and \mathbf{K} are the real $[n \times n]$ mass, viscous damping and stiffness matrices, $\mathbf{q}(t)$ is the $[n \times 1]$ vibration displacement vector, $\mathbf{f}(t)$ is the $[p \times 1]$ vector of excitations and \mathbf{P} is a $[n \times p]$ coordinates matrix. By defining a $[2n \times 1]$ state vector as $\mathbf{x}(t) = [\mathbf{q}^T(t) \ \dot{\mathbf{q}}^T(t)]^T$ and assuming constant inter-sample behaviour of the input signal (e.g., zero-order hold principle for a sampling period T_s), the structural system can be described in the discrete-time state-space as

$$\mathbf{x}[t+1] = \mathbf{A}\mathbf{x}[t] + \mathbf{B}\mathbf{f}[t] \quad (2a)$$

$$\mathbf{y}[t] = \mathbf{C}\mathbf{x}[t] + \mathbf{D}\mathbf{f}[t] \quad (2b)$$

2.2 The parametric identification problem

Given a finite number of structural excitation/response samples, the parametric identification problem pertains to the estimation of a discrete-time state-space model

$$\hat{\mathbf{x}}[t+1] = \mathbf{A}\hat{\mathbf{x}}[t] + \mathbf{B}\mathbf{f}[t] \quad (3a)$$

$$\hat{\mathbf{y}}[t] = \mathbf{C}\hat{\mathbf{x}}[t] + \mathbf{D}\mathbf{f}[t] \quad (3b)$$

so that its output closely resembles the measured vibration response. This implies an optimization problem that attempts to minimize a quadratic objective function of the form

$$f(\mathbf{p}) = \frac{1}{2} \hat{\mathbf{E}}^T(\mathbf{p}) \hat{\mathbf{E}}(\mathbf{p}) \quad (4)$$

for some parameter vector \mathbf{p} , where

$$\hat{\mathbf{E}}(\mathbf{p}) = [\hat{\mathbf{e}}^T[1, \mathbf{p}] \hat{\mathbf{e}}^T[2, \mathbf{p}] \dots \hat{\mathbf{e}}^T[N, \mathbf{p}]] \quad (5)$$

N is the length of the available data samples and $\hat{\mathbf{e}}[t, \mathbf{p}] = \mathbf{y}[t] - \hat{\mathbf{y}}[t, \mathbf{p}]$. Here, \mathbf{p} is the vector that employs the full state-space parameterization, that is

$$\mathbf{p} = [\text{vec}(\mathbf{A})^T \text{vec}(\mathbf{B})^T \text{vec}(\mathbf{C})^T \text{vec}(\mathbf{D})^T]^T \quad (6)$$

where vec denotes the operator stacking the columns of a matrix in series, one underneath the other. In defining the optimization problem using the state-space formulation of Eq. (3), it is assumed that all uncertainties due to modelling errors and noise sources acting on the system are lumped together as an additive disturbance at the output vector (Verhaegen and Verdult 2007). This implies that the applied estimation procedure belongs to the class of output-error methods.

3. The hybrid algorithm

3.1 Description

Deterministic and stochastic optimization algorithms are characterized by disparate features. Specifically, deterministic methods exhibit high convergence rates and certain accuracy, provided that the objective function of Eq. (4) is regular. On the contrary, stochastic algorithms exhibit very low convergence rates, yet, they can search within a significantly broader area for the global optimum, in this way alleviating entrapment in local minima. The hybrid algorithm proposed herein attempts to combine these two seemingly contradicting features into a common platform.

The effectiveness of the $(\mu/\rho (+/,) \lambda, \nu)$ -h-ES is ensured via its partition into both a local and a global search for the optimum. This is succeeded by formulating a super-positioned stochastic global search, which is followed by an independent deterministic procedure that is conditionally activated for specific members of the involved population. This ultimately leads into a structure where every member of the population contributes in the global search, while single individuals perform the local search. Such algorithmic structures imitate insects' societies (Monmarche *et al.* 2000) and they have been presented in, for example, Colorni *et al.* (1996),

Dorigo *et al.* (2000) and Jayaraman *et al.* (2000). Refer to Yu and Gen (2010) for further details.

The ES has been selected to form the stochastic platform. The selection of ES among the other instances of EAs is justified via numerical experiments in non-linear parameter estimation problems (Schwefel 1995), which have provided significant indication that ESs perform better than the other two classes of EAs, namely GA and EP. Similar indications have been provided by Koulocheris *et al.* (2003) in a system identification problem, although only with respect to transfer function model representations. In addition to this evidence and focusing further to the ES family, late investigations that deal with the state-space estimation problem (Dertimanis 2014) show that multi-membered, self-adaptive versions outperform other ESs instances, such as the (1+1)-ES and the CMA-ES.

The conventional self-adaptive ES is based on three operators that take on the recombination, the mutation and the selection tasks. In order to maintain an adequate stochastic performance in the algorithm introduced herein, the recombination and selection tasks are retained unaltered (refer to Beyer and Schwefel (2002) for a brief discussion about the recombination phase), while its strong local topological performance is enhanced through the substitution of the original mutation operator by the LM algorithm.

3.2 The deterministic mutation

Previously developed versions of the $(\mu/\rho (+/\cdot)\lambda, \nu) - h - ES$ have integrate and test several deterministic algorithms that belong to the class of quasi-Newton methods and involve line-search, trust-region, or a combination of both, for the calculation of the parameter vector. In this study the deterministic mutation realizes the LM method, while it incorporates analytical gradient information. Very briefly, the LM algorithm updates the parameter vector by

$$\mathbf{p}_{k+1} = \mathbf{p}_k + \mathbf{s}_k \tag{7}$$

with \mathbf{s}_k denoting the direction, which is calculated as the solution of the following set of linear equations (Moré 1978)

$$(\mathbf{J}_k^T \mathbf{J}_k + \mu_k \mathbf{I}) \mathbf{s}_k = \mathbf{J}_k^T \widehat{\mathbf{E}}_k \tag{8}$$

In Eq. (8) μ_k is calculated using trust-region approaches (Moré 1978, Dennis and Schnabel 1981), while \mathbf{J}_k (the Jacobian matrix of residuals) is calculated using information from the state-space model (Verhaegen and Verdult 2007). More specifically, since

$$\mathbf{J}_k \equiv \frac{\partial \widehat{\mathbf{E}}_k}{\partial \mathbf{p}_k} = \begin{bmatrix} \frac{\partial \widehat{\mathbf{e}}_k[1]}{\partial \mathbf{p}_k} \\ \frac{\partial \widehat{\mathbf{e}}_k[2]}{\partial \mathbf{p}_k} \\ \vdots \\ \frac{\partial \widehat{\mathbf{e}}_k[N]}{\partial \mathbf{p}_k} \end{bmatrix} = - \begin{bmatrix} \frac{\partial \widehat{\mathbf{y}}_k[1]}{\partial \mathbf{p}_k} \\ \frac{\partial \widehat{\mathbf{y}}_k[2]}{\partial \mathbf{p}_k} \\ \vdots \\ \frac{\partial \widehat{\mathbf{y}}_k[N]}{\partial \mathbf{p}_k} \end{bmatrix}, \quad \text{with} \quad \frac{\partial \widehat{\mathbf{y}}_k[t]}{\partial \mathbf{p}_k} = \begin{bmatrix} \frac{\partial \widehat{\mathbf{y}}_k[t]}{\partial p_k[1]} & \dots & \frac{\partial \widehat{\mathbf{y}}_k[t]}{\partial p_k[d]} \end{bmatrix}, t = 1, \dots, N \tag{9}$$

where d is the size of the parameter vector, the calculation of the Jacobian matrix drops down to the calculation of the partial derivatives of the output, w.r.t. the parameter vector. This is very easily accomplished by differentiating the state-space model of Eq. (3) to obtain

$$\frac{\partial \hat{\mathbf{x}}_k[t]}{\partial \mathbf{p}_k} = \mathbf{A} \frac{\partial \hat{\mathbf{x}}_k[t]}{\partial \mathbf{p}_k} + \frac{\partial \mathbf{A}}{\partial \mathbf{p}_k} \hat{\mathbf{x}}[t] + \frac{\partial \mathbf{B}}{\partial \mathbf{p}_k} \mathbf{f}[t] \quad (10a)$$

$$\frac{\partial \hat{\mathbf{y}}_k[t]}{\partial \mathbf{p}_k} = \mathbf{C} \frac{\partial \hat{\mathbf{x}}_k[t]}{\partial \mathbf{p}_k} + \frac{\partial \mathbf{C}}{\partial \mathbf{p}_k} \hat{\mathbf{x}}[t] + \frac{\partial \mathbf{D}}{\partial \mathbf{p}_k} \mathbf{f}[t] \quad (10b)$$

It is thus shown that the calculation of one column of the Jacobian matrix pertains to the simulation of the state-space system described by Eq. (10), for $t = 1, \dots, N$. Due to the adopted parameterization it follows that during the calculation of specific columns of the Jacobian matrix, certain quantities of the left-hand sides of Eq. (10) are zero. Indicatively, when the columns that correspond to the entries of the state matrix are considered

$$\frac{\partial \mathbf{B}}{\partial \mathbf{p}_k^A} = \frac{\partial \mathbf{C}}{\partial \mathbf{p}_k^A} = \frac{\partial \mathbf{D}}{\partial \mathbf{p}_k^A} = \mathbf{O} \quad (11)$$

A very important matter that significantly affects the performance of the h-ES involves the members of the population that are selected for mutation: it has been reported (Koulocheris *et al.* 2003) that a possible cause for the poor performance of EA in non-linear multimodal functions is the loss of information through the non-privileged individuals of the population. Thus, the deterministic mutation is applied only to the ν worst of the population, where ν is an additional strategy parameter. This means that a sorting procedure takes place twice in every iteration step: the first time in order to form the pool of the ν worst individuals, and the second to support the selection operator, which succeeds the novel deterministic mutation one. This modification enables the strategy to yield the corresponding local optimum for each of the selected ν worst individuals in every iteration step. The advantage is reflected in terms of increased convergence rate and reliability in the search for the global optimum.

It is stressed that three further alternatives were tested in the course of this work with suboptimal performance, which is not reported herein. In these, the deterministic mutation operator was activated by:

- Every individual of the involved population: this increased the computational cost of the algorithm without the desirable effect.
- A number of privileged individuals: this led to premature convergence of the algorithm to local optima of the objective function.
- A number of randomly selected individuals: this generated unstable behavior that led to statistically low performance and insufficient consistency.

3.3 Termination criteria

The adopted termination criteria can be distinguished into both local and global, depending on whether they apply to the deterministic mutation operator, or to the $(\mu/\rho (+/,) \lambda, \nu)$ -h-ES. For the former, standard termination tests are utilized, which are described thoroughly in the literature (Dennis and Schnabel 1981, Fletcher 2000):

- Objective function value smaller than a specified tolerance,
- relative gradient norm less than a specified tolerance,
- relative distance between two successive iterations less than a specified tolerance,
- not a descent current direction, and
- maximum mutation operator iterations exceeded.

The hybrid algorithm terminates if at least one of the following occurs:

- Difference between worse and best objective function less than a specified tolerance,
- maximum function evaluations exceeded, and
- maximum iterations exceeded.

4. The structural estimation procedure

Prior investigation on pure ES (Dertimanis 2014) and early explorations of h-ES (Dertimanis and Chatzi 2014) indicated that state-space estimation using stochastic and/or hybrid algorithms may result in prohibitive computational burden if the initial cloud is allowed to vary freely. Based on these indications, the initial population (i.e., the set μ of parents) is generated around a centroid individual that is estimated using subspace techniques. While several other alternatives are possible (Katayama 2005), the estimation process realizes a deterministic subspace method that is described in Sec. 9.2.4 of Verhaegen and Verdult (2007), which is closely related to the MOESP class of algorithms. Upon estimation of the centroid individual, the initial population is generated through uniformly distributed sampling, the extrema of which constitute exogenous parameters.

Model order selection is realized by estimating successive state orders within a predefined band and by employing the *variance accounted for* (VAF) metric

$$VAF = \max \left(0, \left(1 - \frac{\frac{1}{N} \sum_{t=1}^N \|y[t] - \hat{y}[t]\|_2^2}{\frac{1}{N} \sum_{t=1}^N \|y[t]\|_2^2} \right) 100\% \right) \quad (12)$$

that is essentially a scaled version of the objective function expressed in percentage format. The adequacy of a model of certain order can be also judged by generating stabilization diagrams that plot a structural quantity (e.g., frequency) against the estimated state orders. Indications from a very wide and diverse variety of problems has shown that structural modes tend to stabilize, while false modes appear scattered on the diagram (Lau *et al.* 2007).

Once a final candidate model has been extracted and validated (using typical model validation tools, such as the auto-covariance matrix of the residuals and the cross-covariance matrix between the residuals and the excitations) the vibration modes are calculated from the eigenvalue problem of the (discrete-time) state matrix (Reynders 2012)

$$\mathbf{A} = \mathbf{\Psi} \mathbf{\Lambda} \mathbf{\Psi}^{-1} \quad (13)$$

It follows that the natural frequencies, the damping ratios and the mode shapes are calculated by

$$f_k = \frac{|\ln(\lambda_k)|}{2\pi T_s} \text{ (Hz)}, \quad \zeta_k = -\frac{\ln(\lambda_k)}{2\pi T_s f_k}, \quad \varphi_k = \mathbf{C}\Psi_{*k}, \quad (14)$$

where λ_k is the k th eigenvalue of the state matrix and Ψ_{*k} the k th column of the corresponding eigenvector matrix.

The quality of the estimated structural modes is further assessed by applying a physically meaningful dispersion analysis framework, which quantifies the individual contribution of every identified vibration mode into the total stochastic vibration energy. The relevant methodology is described in Dertimanis (2013) and pertains to a modal decomposition of the zero-lag covariance matrix of the output vector and to the definition of the *Normalized Modal Dispersion Metric* (NMDM). Under this metric, every identified mode is attributed by an additional quantity $\bar{\delta}_k$, while the mode with the greatest contribution is being normalized to have $\bar{\delta}_k = 1$. Refer to Appendix A for a brief review of the method and to Dertimanis (2013) for further details. A flowchart of the proposed structural estimation method is displayed in Fig. 1.

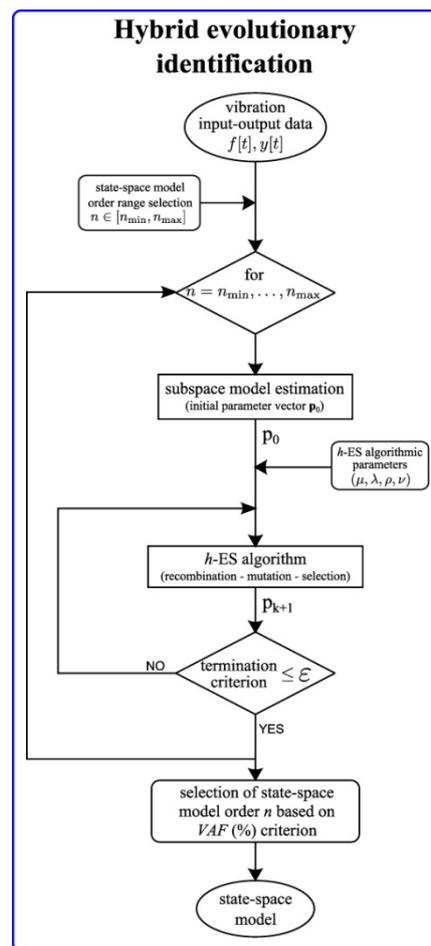


Fig. 1 Flowchart of the proposed hybrid structural identification method

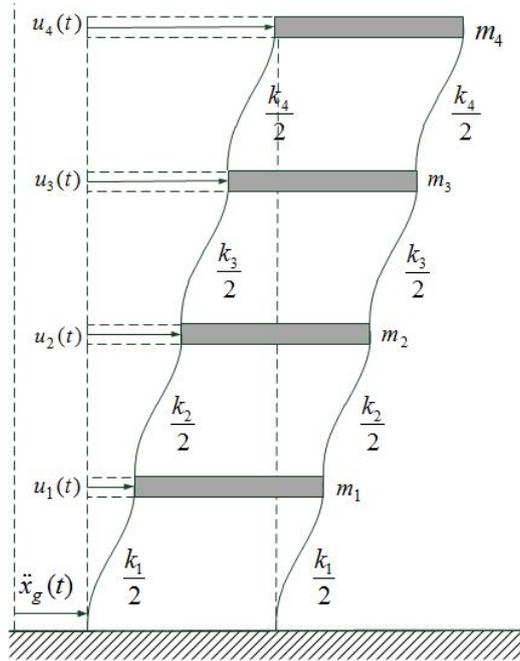


Fig. 2 Sketch of the simulated shear frame

Table 1 Physical and modal properties of the frame

Physical Space		
Story	m_j (Mgr)	k_j (kN / m)
1	100	200000
2	80	150000
3	80	150000
4	80	150000
Modal Space		
Mode	f_n (Hz)	ζ_n (%)
1	2.520	1.00
2	6.977	1.00
3	10.341	1.00
4	12.812	1.00

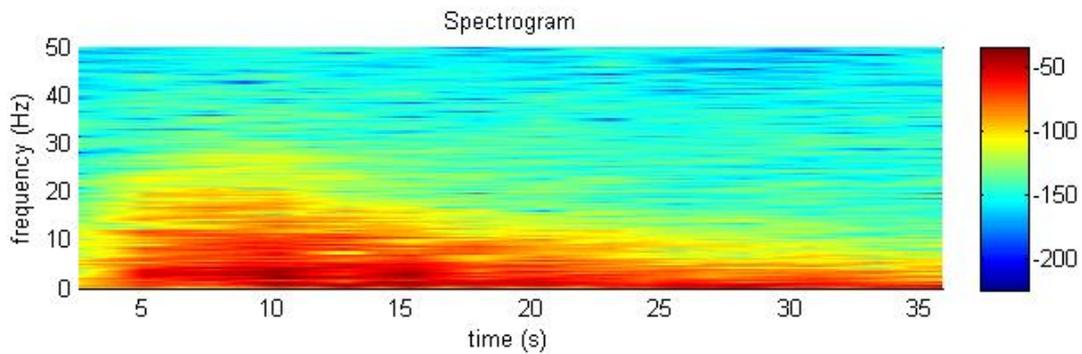


Fig. 3 Spectrogram of the Northridge earthquake (Hamming window of size equal to 512, 50% overlapping)

5. The identification experiments

5.1 Monte Carlo identification study of the simulated shear frame

The first structural identification problem considered is a simulated shear frame structure with four DOFs, a sketch of which is displayed in Fig. 2. As illustrated in Table 1, the frame has its vibration modes within the [2 13] Hz band, with each one characterized by light damping (1%). For the identification tasks, the structure is excited by the Northridge earthquake, the spectrogram of which is presented in Fig. 3 (Hamming window of size equal to 512 and 50% overlapping), where it is obvious that the frequency band of interest is sufficiently excited. The simulated data are obtained through the discretization of the structural equation (in fact its continuous state-space representation) into the state-space model of Eq. (5), at a sampling period $T_s = 0.01$ s, using the zero-order hold. Accordingly, the structural vibration outputs, i.e., the relative acceleration of every story, are mean value subtracted and the respective input-output data records have $N=4000$ data per channel.

In order to test the statistical consistency of the h-ES and perform comparisons to the standard self-adaptive ES, this series of tests presumes availability of data from all DOF, as well as known state order. A Monte Carlo analysis that involves 100 independent identification processes is incorporated. In every such individual process, the h-ES is implemented using standard values for the exogenous parameters, that is plus variant with 15 parents, 100 offspring and 2 recombination parents (panmictic intermediate scheme) (Schwefel 1995, Bäck 1996, p. 83), while in every iteration the 3 worst individuals are mutated (Koulocheris *et al.* 2008). The initial population is uniformly generated within the $\pm 10^{-4}$ boundaries around the centroid individual. The standard ES uses exactly the same exogenous parameters and initializes at the same random state, to the one of the hybrid algorithm. This implies that the initial population of both algorithms is identical during a Monte Carlo iteration. Regarding the termination of the optimization process, this occurs when both (i) the best objective function is at least 100% better to the one of the subspace estimate, and (ii) when the difference between the best and the worst objective function is less than 10^{-8} .

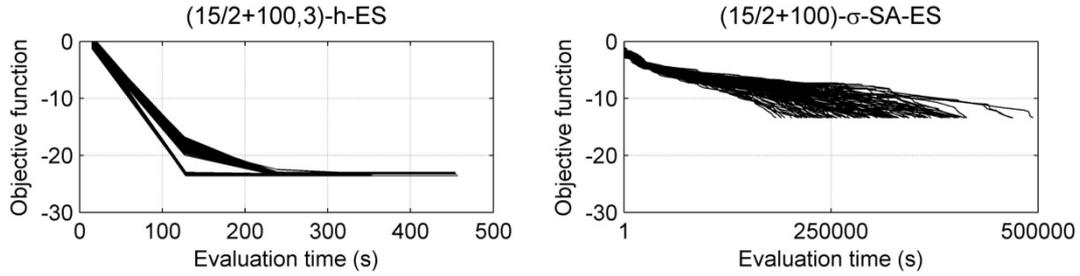


Fig. 4 Logarithmic objective function values against evaluation times (simulated frame)

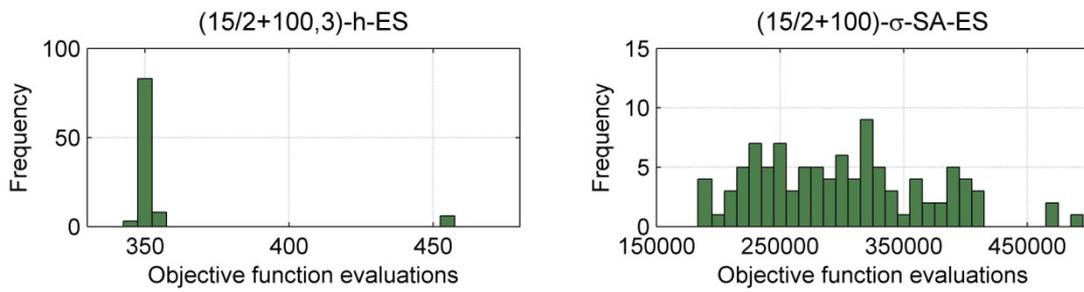


Fig. 5 Histograms for the number of objective function evaluations of each algorithm (simulated frame)

Table 2 Percentage improvement (%) from the subspace estimate and computational load in minutes. All quantities are provided in *mean* (\pm *standard deviation*) format (second row, simulated frame)*

	(15/2+100,3)-h-ES	(15/2+100)-σ-SA-ES
Improvement	250.313 (\pm 1.857)	100.116 (\pm 0.130)
CPU	0.169 (\pm 0.011)	11.030 (\pm 2.528)

Table 3 Mean values of estimated vibration modes. Standard deviations are negligible and therefore they are not shown (simulated frame)

Mode	Theoretical			(15/2+100,3)-h-ES			(15/2+100)-σ-SA-ES		
	f_n (Hz)	ζ_n (%)	$\bar{\delta}_k$	f_n (Hz)	ζ_n (%)	$\bar{\delta}_k$	f_n (Hz)	ζ_n (%)	$\bar{\delta}_k$
1	2.520	1.00	1.000	2.520	1.00	1.000	2.520	1.00	1.000
2	6.977	1.00	0.440	6.977	1.00	0.440	6.977	1.00	0.440
3	10.341	1.00	0.178	10.341	1.00	0.178	10.341	1.00	0.178
4	12.812	1.00	0.027	12.812	1.00	0.027	12.812	1.00	0.027

* Simulations were carried out on a PC with a Quad Core 3.20GHz Intel Processor and 8.00GB of RAM, on a 64-bit WINDOWS 7 operating system.

The results of the Monte Carlo simulation are expanded over Figs. 4 and 5 and Tables 2-4. Overall, both the hybrid algorithm and the conventional ES exhibit very good performance in all the examined issues and great accuracy in the estimates of the structural vibration modes (Table 4), which result identical to their theoretical counterparts, at a negligible standard deviation. Specific remarks for each competing algorithm are given below:

- **(15/2+100)- σ -SA-ES:** the conventional ES manages to return satisfying estimates in all the individual experiments, as a result of constantly decreasing iterations. However, as Fig. 4 illustrates, convergence is extremely slow and spread over a very wide range of function evaluations. Indeed, it is observed that in all individual Monte Carlo experiments the algorithm terminates at approximately the same objective function value (around -13, that corresponds to the base 10 logarithm of the final value of Eq. (4)), which implies that (i) the convergence rates return very slow and typical of stochastic optimization algorithms, (ii) the algorithm terminates only when the first criterion (at least 100% improvement from the subspace estimate) is fulfilled (see also Table 2), and (iii) the second termination criterion (convergence in population) has been already fulfilled. This performance is observed to be evolved in a quite wide range of objective function evaluations: as Fig. 4 shows, the objective function evaluations vary from around 175000, up to 500000, which is a strong indication of poor computational consistency. The very high computational costs are also reflected in the measured cpu times required for convergence, that have been reached up to 11min in the mean, followed by a deviation of approximately 2.5 min. Nonetheless, the identified structural vibration modes return identical to the true ones (see Table 3) proving an accurate methodology.
- **(15/2+100,3)-h-ES:** the proposed algorithm succeeds in returning very satisfying estimates in all the individual experiments. Starting from Fig. 4, two distinct features indicate a dramatic performance improvement, in respect to the conventional ES. The first feature refers to the induced objective function evaluations, which have been dropped down by three orders of magnitude, as compared to the ones of the self-adaptive ES. Indeed, due to the presence of the deterministic mutation, all individual experiments have been terminated in less than 400 function evaluations, while convergence is already achieved at about 100-200 objective function evaluations, which, given the number of initial population, corresponds to the first, or maybe second iteration of the hybrid algorithm. The second feature that is seen from Table 2 refers to the percentage improvement from the subspace estimate: the hybrid algorithm has resulted in a mean improvement of about 250%, in respect to the subspace estimate. This figure is quite indicative of the benefits of the deterministic mutation, which actually “pushes” a “bad” individual to become the “best”, while this is accomplished in very high rates. This performance is characterized by remarkable statistical consistency: as the histogram of Fig. 5 reveals, more than 80% of the final objective function evaluations result in the same frequency bin, while the standard deviation in the resulted objective function values is less than 2%, (Table 2). In accordance with these remarks, it is not surprising that the mean cpu load of the h-ES results two orders of magnitude smaller than the one of the standard ES and is formulated at 0.17min, with a standard deviation of 0.011min (approximately 0.6s). Naturally, the estimated structural vibration modes also return identical to the true ones (see Table 3) proving a both accurate and rapid methodology.

5.2 Structural identification of the laboratory shear frame

The algorithm and the adopted estimation procedure are now applied to the structural identification problem of a laboratory frame with lateral force resistance being provided by X-braces, which is displayed in Fig. 6 and has the geometrical and material properties of Table 4. The frame is mounted through four bolts on a single-axis shake table running on displacement control mode (frequency range 0-100 Hz, force range ± 100 kN, stroke ± 125 mm, max. velocity 0.55 m/s).

A random zero-mean Gaussian displacement signal is used as a seismic excitation. The corresponding seismic acceleration is measured at the base (actual input excitation, location 0; Fig. 6(b)), whereas the structural vibration response at each story of the frame (absolute accelerations, locations 1-4; Fig. 6(b)). To this, corresponding triaxial lightweight MEMS accelerometers (STMicroelectronics LIS344ALH, range ± 2 g, sensitivity 3.3/5 V/g, acceleration noise density $50 \mu\text{g}/\sqrt{\text{Hz}}$) are placed at the geometric center of each plate. The sampling frequency is set at $F_s = 1250$ Hz, yet, as the frequency band of interest is limited to $[0 \ 30]$ Hz, the acquired signals are low-pass filtered (16th order Chebychev Type II digital filter with 50 mHz cut-off frequency) and resampled at $F_s = 125$ Hz. The final input-output data set used for the identification contains $N = 5000$ data per channel. Fig. 7 shows the power spectral density estimates (Welch's method with Hamming window of size equal to 1024 and 50% overlapping) of the structural vibration outputs, from where four lightly damped modes can be clearly observed at around 3, 12, 20 and 27 Hz.

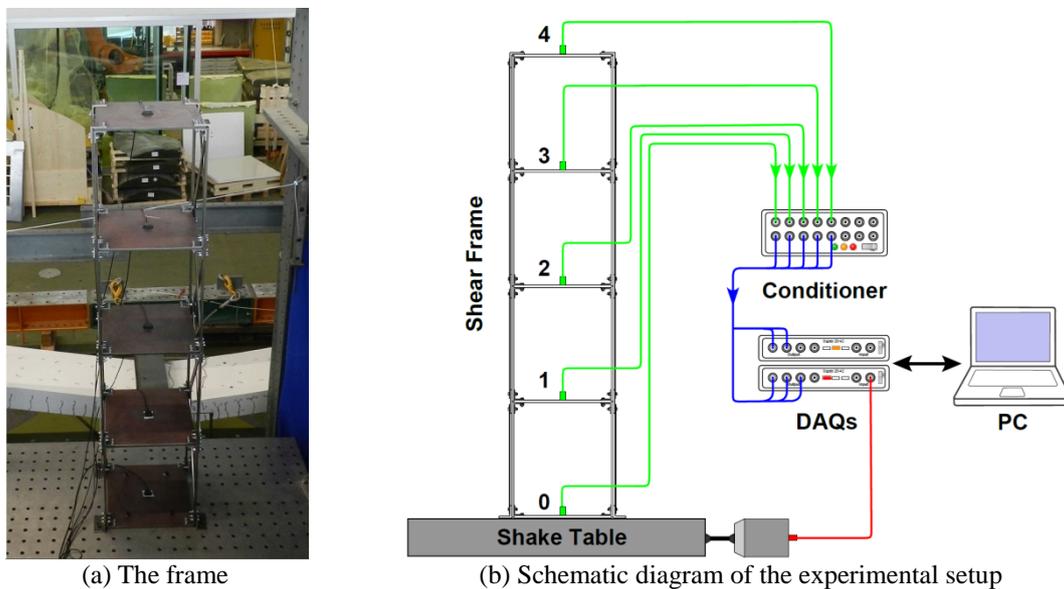


Fig. 6 The laboratory shear frame structure

Table 4 Physical properties of the laboratory frame

Component	Dimensions (L×W×H mm)	Material	Young modulus	Density
Story plate	500 × 470 × 12			
Column	550 × 45 × 10	Steel S235	210 GPa	7850 kg/m ³
Weak column	550 × 45 × 4			
X-brace	623 × 40 × 6			

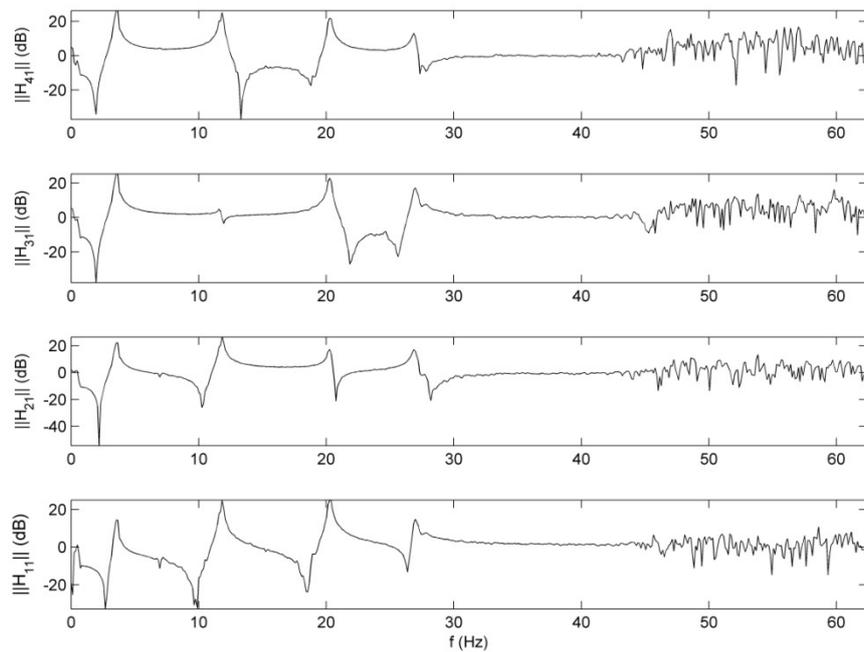


Fig. 7 Power spectral densities of the structural responses of the laboratory frame (Welch's method, Hamming window of size equal to 1024, 50% overlapping)

The Monte Carlo analysis contains two distinct cases, at two different sizes of the estimation set, that is, $N=1000$ and $N=1500$ data per channel. In order to check the proposed scheme under limited data availability, the data set used to the identification experiments contains the earthquake acceleration and the structural responses, in the form of the first and the last story absolute acceleration. During every Monte Carlo iteration, the identification is carried out for state orders $n = 8, 10, 12, 14, 16$ and 18 , which correspond to parameter vectors of lengths shown at Table 5. The hybrid algorithm is implemented using the same setting as before (plus variant; 15 parents; 100 offspring; 2 recombination parents; mutation of the three worst individuals; uniform generation of the initial population within the $\pm 10^{-4}$ boundaries around the subspace-based centroid individual). Termination of the optimization process now occurs only when the difference between the best and the worst objective function is less than 10^{-8} (convergence in population).

Table 5 State orders and lengths of the corresponding parameter vectors

n	8	10	12	14	16	18
d	90	132	182	240	306	380

Table 6 VAF (%), improvement (%) from the subspace estimate and CPU time (min) for each state order considered (laboratory frame, N=1000 case). All quantities are provided in *mean* (\pm *standard deviation*) format

n	8	10	12	14	16	18
VAF	94.90 ($\pm 2.26 \times 10^{-6}$)	95.13 ($\pm 2.57 \times 10^{-6}$)	95.15 ($\pm 2.51 \times 10^{-3}$)	95.19 ($\pm 3.48 \times 10^{-5}$)	95.19 ($\pm 2.74 \times 10^{-5}$)	95.38 ($\pm 4.44 \times 10^{-5}$)
Improvement	34.27 ($\pm 2.26 \times 10^{-6}$)	36.79 ($\pm 2.26 \times 10^{-6}$)	38.49 ($\pm 2.26 \times 10^{-6}$)	39.93 ($\pm 2.26 \times 10^{-6}$)	51.15 ($\pm 2.26 \times 10^{-6}$)	74.66 ($\pm 2.26 \times 10^{-6}$)
CPU	0.056 (± 0.008)	0.148 (± 0.024)	0.284 (± 0.046)	0.859 (± 0.311)	6.235 (± 3.530)	1.287 (± 0.808)

5.2.1 Results for N=1000

Fig. 8 and Table 6 illustrate the main results of the Monte Carlo analysis. As a general comment it can be argued that the hybrid algorithm has retained similar performance to the one observed in the numerical case study for almost all the examined state orders. In view of the critical issues assessed, the following specific remarks can be made:

- **Statistical consistency:** Fig. 8 displays the performance of the hybrid algorithm in terms of objective function values against evaluation times, as well as histograms of the latter, for each state order considered. Up to and including $n=14$, the behavior is quite similar, showing that a stationary point has already been achieved at low iterations, whereas afterwards the algorithm obviously attempts to fulfill the criterion of population convergence. This implies a greater deviation in the final number of objective function evaluations (observe the corresponding histograms), but not to a level that reduces the statistical efficiency of h-ES. In the last two orders a slight alteration from this performance is observed. Although this requires further attention and it is certainly related to the size of the parameter vector, it can be also attributed to the quality of the subspace estimate. In specific, by looking at the percentage improvement of the final objective function value in comparison to its subspace estimate (Table 6), it is obvious that at $n=16,18$ the values are sufficiently higher, reaching up to about 75% when $n=18$. This indeed means that the hybrid algorithm requires further effort to reach to the same results produced at lower orders. It is likely that this behavior is attributed to the induced numerical inefficiencies implied by the selection of such high orders. The implementation of more sophisticated subspace methods may partially resolve this issue.

The consistency of the proposed algorithm can be also positively assessed by the induced VAF values. As Table 6 displays, these have resulted very high for all the considered orders, while the corresponding standard deviations are negligible. Same levels of deviations have been also resulted for the percentage improvements from the subspace estimates, although the relative mean values are significantly different when the order increases, as commented above.

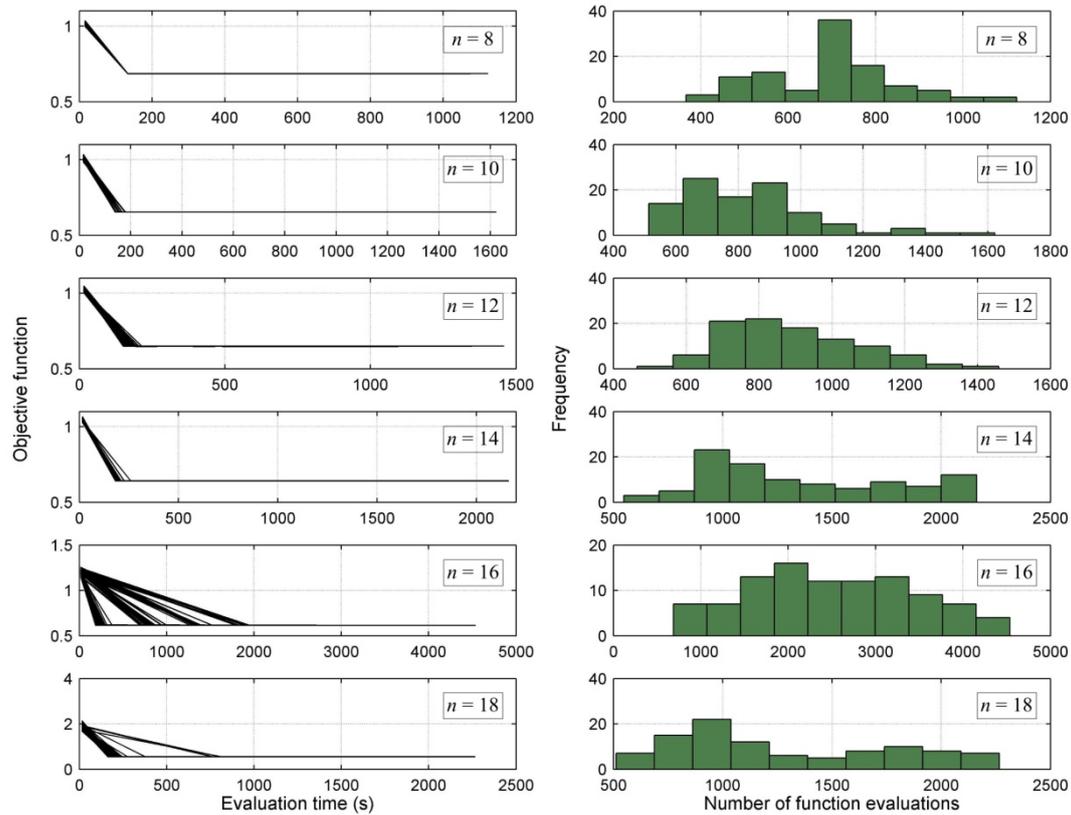


Fig. 8 Logarithmic objective function values against evaluation times and histograms for the number of function evaluations for each state order considered (laboratory frame, $N=1000$ case)

Table 7 Estimated vibration modes (laboratory frame, $N=1000$ case)

Mode	f_n (Hz)	ζ_n (%)	$\bar{\delta}_k$
1	$3.53 \pm 2.51 \times 10^{-8}$	$0.47 \pm 9.80 \times 10^{-9}$	$0.74 \pm 1.32 \times 10^{-6}$
2	$11.77 \pm 5.06 \times 10^{-8}$	$0.34 \pm 6.45 \times 10^{-9}$	$0.53 \pm 1.16 \times 10^{-6}$
3	$20.30 \pm 7.21 \times 10^{-8}$	$0.19 \pm 1.91 \times 10^{-9}$	1.00
4	$26.92 \pm 4.50 \times 10^{-6}$	$0.35 \pm 1.69 \times 10^{-7}$	$0.06 \pm 1.23 \times 10^{-6}$

- **Computational cost:** the induced CPU times in minutes are listed in the last row of Table 6. At lower state orders the computational time is comparable to that of the simulated frame and the standard deviations are one order of magnitude lower than the respective means. At higher

orders the same values increase and this is attributed to both the complexity of h-ES, yet also to the quality of the subspace estimate, which decreases with increasing state order, causing the hybrid algorithm to require more time to reaching to a stationary solution. It is reminded that the current implementation of h-ES utilizes a large number of individuals (parents and offspring), in accordance to literature suggestions. A compromise among the utilized number of individuals, the quality of the estimate and the required computational costs remains an open issue.

- **Quality of the estimated vibration modes:** the numerical results of the Monte Carlo analysis suggest that $n=8$ is a sufficient state order for the description of the laboratory frame. Taken under consideration that the structure is characterized by four modes of vibration, this selection corresponds to zero overdetermination degree and enhances the numerical efficiency of the proposed method, besides confirming the good quality of the measured data. The estimates of the four structural vibration modes are depicted in Table 8 and are in line with the nonparametric estimates (see Fig.7). The damping has resulted quite light, which may be also inferred by observing the sharp peaks of the FRFs, while the third mode has resulted the most significant in terms of its contribution to the total stochastic vibration energy. Of particular importance are the induced standard deviations of the structural estimates, which have resulted very low, again confirming the statistical consistency of h-ES and providing significant indications that the estimated quantities are close to their “actual” values.

5.2.2 Results for $N=1500$

Increasing the length of the estimation set produces the results shown at Fig. 9 and Tables 8 and 9. A general impression is that this increase does not result in significantly greater quality although it seems that it enhances the consistency of h-ES. More specifically:

- **Statistical consistency:** as Fig. 9 shows, the hybrid algorithm appears more consistent than before. Indeed, a stationary point is achieved at low iterations for all state orders and the histograms of the objective function evaluations are expanded over narrower deviations. All other monitored quantities are slightly improved, including the *VAF* values and the percentage improvements from the subspace estimates. An unexpected behavior is, however, observed for $n=12$: at this order, the results of the Monte Carlo analysis have returned quite dispersed (see Fig. 9(c) and the relative column of Table 8). Apart from numerical issues, there is no apparent explanation for this performance, especially after taking into consideration the corresponding results for $N=1000$, yet it requires a more detailed examination
- **Computational cost:** the returned CPU times at lower orders are comparable to that of the $N=1000$ case, whereas at higher orders the computational cost increases, reaching up to four times larger mean values when $n=18$. Again, for $n=12$, a discrepancy is observed in both the mean and the variance of the CPU time.
- **Quality of the estimated vibration modes:** as previously, $n=8$ is considered a sufficient state order for the description of the laboratory frame resulting in zero overdetermination degree. Table 8 illustrates the estimates of the four structural vibration modes, which are in line with both the nonparametric estimates (see Fig.7) and the parametric ones for $N=1000$ (see Table 7), apart from the damping of the first three modes which is slightly different. The induced standard deviations have returned at the same levels and provide further evidence of the algorithm’s efficiency.

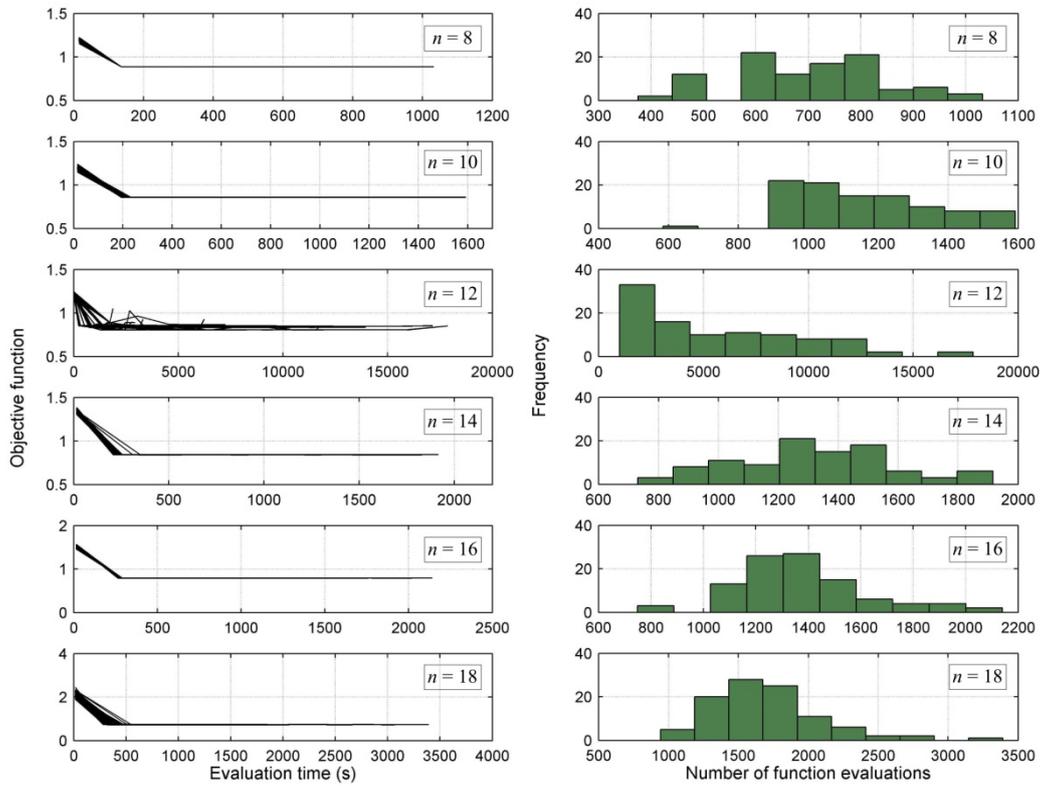


Fig. 9 Logarithmic objective function values against evaluation times and histograms for the number of function evaluations for each state order considered (laboratory frame, $N=1500$ case)

Table 8 VAF (%), improvement (%) from the subspace estimate and CPU time (min) for each state order considered (laboratory frame, $N=1500$ case). All quantities are provided in *mean* (\pm *standard deviation*) format

n	8	10	12	14	16	18
VAF	95.31 ($\pm 9.77 \times 10^{-9}$)	95.55 ($\pm 3.57 \times 10^{-6}$)	95.55 ($\pm 0.15 \times 10^{-0}$)	95.63 ($\pm 0.10 \times 10^{-3}$)	95.63 ($\pm 9.03 \times 10^{-6}$)	95.77 ($\pm 8.74 \times 10^{-3}$)
Improvement	29.02 ($\pm 5.67 \times 10^{-7}$)	31.50 ($\pm 2.81 \times 10^{-6}$)	33.21 ($\pm 2.33 \times 10^{-0}$)	39.93 ($\pm 2.60 \times 10^{-2}$)	49.90 ($\pm 1.84 \times 10^{-5}$)	71.20 ($\pm 9.54 \times 10^{-2}$)
CPU	0.055 (± 0.008)	0.289 (± 0.020)	7.312 (± 6.335)	0.955 (± 0.237)	1.484 (± 0.095)	4.302 (± 1.588)

Table 9 Estimated vibration modes (laboratory frame, N=1500 case)

Mode	f_n (Hz)	ζ_n (%)	$\bar{\delta}_k$
1	$3.53 \pm 1.03 \times 10^{-7}$	$0.49 \pm 8.91 \times 10^{-9}$	$0.78 \pm 1.84 \times 10^{-6}$
2	$11.78 \pm 7.61 \times 10^{-8}$	$0.37 \pm 4.21 \times 10^{-9}$	$0.63 \pm 1.40 \times 10^{-6}$
3	$20.30 \pm 6.35 \times 10^{-8}$	$0.22 \pm 4.07 \times 10^{-9}$	1.00
4	$26.91 \pm 8.08 \times 10^{-6}$	$0.35 \pm 1.34 \times 10^{-7}$	$0.07 \pm 1.40 \times 10^{-6}$

6. Conclusions

This study proposed a novel, hybrid optimization algorithm, followed by a corresponding estimation procedure, for the parametric, time-domain identification of structural systems via state-space models. The hybrid algorithm, h-ES, has been designed in a way that integrates the advantages of its deterministic and stochastic counterparts and combines high convergence rates and increased reliability in the search for the global optimum. This is succeeded by setting a purely stochastic platform using the conventional ES and by replacing the original mutation operator by the LM method. This operator is applied to a preselected number of the worst individuals of the population in every iteration and contributes to the convergence of the population into (at least) a stationary point.

Although the underlying mechanisms that determine the actual performance of the algorithm, in respect to the structural identification problem that is being faced, are still under careful investigation, it seems that there are two performance potentials. Specifically, h-ES may either (i) significantly improve the effects of the initial population, by providing sufficiently better estimates in every iteration, until the population converges, or (ii) converge to a solution that can already be provided by the mutation of the initial population (via the LM method) in a way that leads into a stationary point. This second potential implies that the job could be done by the LM method its self (given a good initial “guess”), yet, the hybrid algorithm ensures that this point is indeed stationary (or global optimum) by “pushing” the whole population to converge to this point.

The application of the algorithm to the test cases considered in Section 5 can be summarized by the following remarks:

- The hybrid algorithm showed increased statistical consistency. This is confirmed by a number of observed features that included the range of the final objective function values, the almost identical quality of the final estimates and the required computational costs, over the set of Monte Carlo iterations.
- The computational load of the hybrid algorithm is drastically lower than the one of its fully stochastic counterpart and can provide accurate state-space models in feasible times.
- In respect to the experimental case, the hybrid algorithm produced accurate results using data of low length and limited observations. Moreover, at this length, a zero overdetermination degree resulted (despite the fact that the latter is also due to the quality of the acquired data).

- A certain degree of robustness over the subspace estimate seems to characterize the hybrid algorithm. Indeed, as the experimental study revealed, h-ES exhibited capability to estimate state-space models of the same degree of quality, by starting from “not so good” subspace vectors. This is unavoidably achieved at a greater computational cost.
- The boundaries for the uniform generation of the initial population greatly affect the performance of the algorithm in respect to the computational cost. A systematic investigation of the relation among the subspace estimate, the adopted boundaries and the result of the process is still under development.

The encouraging results suggest further development towards this direction. In this respect, current research undertaken by the authors investigates the aforementioned critical issues, as well as the perspectives of using more sophisticated state-space representations, such as the tridiagonal one. It is believed that the latter can lead to a new establishment of the induced optimization problem and result into an even greater reduction of the computational cost, combined with increased estimation accuracy.

Acknowledgments

Rafael Dubs, Alessandra Eicher, Rebekka Habegger, and Severin Haefliger, master students of ETH Zurich, who designed, constructed and tested the shear frame structure under the supervision of the authors, are gratefully acknowledged.

References

- Bäck, T. (1996), *Evolutionary Algorithms in Theory and Practice: Evolution Strategies, Evolutionary Programming, Genetic Algorithms*, Oxford University Press.
- Bergboer, N.H., Verdult, V. and Verhaegen, M.H.G. (2002), “An efficient implementation of maximum likelihood identification of LTI state-space models by local gradient search”, *Proceedings of the 41st IEEE Conference on Decision and Control*, Las Vegas, NV, USA.
- Beyer, H.G. and Schwefel, H.P. (2002), “Evolution strategies – A comprehensive introduction”, *Natural Comput.*, **1**(1), 3-52.
- Casciati, S. (2008), “Stiffness identification and damage localization via differential evolution algorithms”, *Struct. Control Health Monit.*, **15**(3), 436-449.
- Catbas, F.N., Kijewski-Correa, T. and Aktan, A.E. (Eds.) (2013), *Structural Identification of Constructed Systems: Approaches, Methods, and Technologies for Effective Practice of St-Id*, ASCE, Reston, USA.
- Dahlén, A., Lindquist, A. and Mari, J. (1998), “Experimental evidence showing that stochastic subspace identification methods may fail”, *Syst. Control Lett.*, **34**(5), 303-312.
- Dennis, J.E. and Schnabel, R.B. (1981), *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, SIAM, Philadelphia, USA.
- Dertimanis, V.K. (2014), “Output-error state-space identification of vibrating structures using evolution strategies: a benchmark study,” *Smart Struct. Syst.*, **14**(1), 17-37.
- Dertimanis, V.K. (2013), “On the use of dispersion analysis for model assessment in structural identification”, *J. Vib. Control*, **19**(15), 2270-2284.
- Dertimanis, V.K. and Chatzi, E.N. (2014a), “A hybrid evolution strategy approach to the structural identification problem via state-space models”, *Proceedings of the 7th European Workshop on Structural Health Monitoring*, Barcelona, Spain, July.
- Dertimanis, V.K., and Chatzi, E.N. (2014b), “Dispersion-corrected stabilization diagrams for model order

- assessment in structural identification”, *Proceedings of the 6th World Conference on Structural Control and Monitoring*, Nantes, France, July.
- Dertimanis, V., Koulocheris, D., Vrazopoulos, H. and Kanarachos, A. (2003), “Time-series parametric modeling using evolution strategy with deterministic mutation operators”, *Proceedings of the IFAC International Conference on Intelligent Control Systems and Signal Processing*, Faro, Portugal, April.
- Dimou, C.K. and Koumoussis, V.K. (2003), “Competitive genetic algorithms with application to reliability optimal design”, *Adv. Eng. Softw.*, **34**(11), 773-785.
- Fleming, P. and Purshouse, R. (2002), “Evolutionary algorithms in control systems engineering: a survey”, *Control Eng. Pract.*, **10**(11), 1223-1241.
- Fletcher, R. (2000), *Practical Methods of Optimization*, (2nd Ed.), John Wiley & Sons Ltd., New York, USA.
- Friswell, M. (2007), “Damage identification using inverse methods”, *Philos. T. Roy. Soc. A.*, **365**(1851), 393-410.
- Fritzen, C.P., Klinkov, M. and Kraemer, P. (2013), “Vibration-based damage diagnosis and monitoring of external loads”, (Eds., Ostachowicz, W. and Güemes, J.A.), *New Trends in Structural Health Monitoring*, Springer, Vienna, Austria.
- Gibson, S. and Ninness, B. (2005), “Robust maximum-likelihood estimation of multivariable dynamic systems”, *Automatica*, **41**(10), 1667-1682.
- Huang, J.N. and Pappa, R.S. (1985), “An Eigensystem Realization Algorithm (ERA) for modal parameter identification and model reduction”, *J. Guid. Control Dynam.*, **8**, 620-627.
- Katayama, T. (2005), *Subspace Methods for System Identification*, Springer-Verlag, Berlin, Germany.
- Kim, J. and Lynch, J.P. (2012), “Subspace system identification of support-excited structures-part I: theory and black-box system identification”, *Earthq. Eng. Struct. D.*, **41**(15), 2235-2251.
- Koulocheris, D., Dertimanis, V. and Spentzas, C. (2008), “Parametric identification of vehicle structural characteristics”, *Forschung im Ingenieurwesen*, **72**(1), 39-51.
- Koulocheris, D., Vrazopoulos, H. and Dertimanis, V. (2003), “Vehicle suspension system identification using evolutionary algorithms”, *Proceedings of the EUROGEN*, Barcelona, Spain, July.
- Kristinsson, K. and Dumont, G. (1992), “System identification and control using genetic algorithms”, *IEEE T. Syst. Man. Cybernetics*, **22**(5), 1033-1046.
- Lau, J., Lanslots, J., Peeters, B. and Van der Auweraer, H. (2007), “Automatic modal analysis: reality or myth?”, *Proceedings of the 25th International Modal Analysis Conference, IMAC-XXV*, Orlando, USA.
- Lin, J.W. and Betti, R. (2004), “On-line identification and damage detection in non-linear structural systems using a variable forgetting factor approach”, *Earthq. Eng. Struct. D.*, **33**(4), 419-444.
- Ljung, L. (1999), *System Identification: Theory for the User*, (2nd Ed.), Prentice-Hall Inc., Englewood Cliffs, NJ, USA.
- McKelvey, T., Helmersson, A. and Ribarits, T. (2004), “Data driven local coordinates for multivariable linear systems and their application to system identification”, *Automatica*, **40**(9), 1629-1635.
- McKelvey, T. and Helmersson, A. (1997), “System identification using an over-parameterized model class-improving the optimization algorithm”, *Proceedings of the 36th IEEE Conference on Decision and Control*, San Diego, CA.
- Moré, J.J. (1978), “The Levenberg-Marquardt algorithm: implementation and theory”, (Ed. Watson, G.A.), *Lecture Notes in Mathematics 630*, Springer, Berlin, Germany.
- Nagarajaiah, S. and Basu, B. (2009), “Output only modal identification and structural damage detection using time frequency & wavelet techniques”, *Earthq. Eng. Eng. Vib.*, **8**(4), 583-605.
- Papadimitriou, C., De Roeck, G. and Lombaert, G. (2012). “Predictions of fatigue damage accumulation in the entire body of metallic bridges by analysing operational vibrations”, *Proceedings of the 3rd International Symposium on Life-Cycle Civil Engineering*, Vienna, Austria.
- Reynders, E. (2012), “System identification methods for (operational) modal analysis: review and comparison”, *Earthq. Eng. Struct. D.*, **19**(1), 51-124.
- Schwefel, H.P. (1995), *Evolution and Optimum Seeking*, Wiley, New York, NY.
- Tang, H.S., Xue S.T. and Fan. C. (2008), “Differential evolution strategy for structural system

- identification”, *Comput. Struct.*, **86**(21-22), 2004-2012.
- Verhaegen, M. and Verdult, V. (2007), *Filtering and System Identification: a Least Squares Approach*, Cambridge University Press, Cambridge, UK.
- Wu, T. and Kareem, A. (2013), “Aerodynamics and aeroelasticity of cable-supported bridges: identification of nonlinear features”, *J. Eng. Mech. - ASCE*, **139**(12), 1886-1893.
- Yu, X. and Gen, M. (2010), *Introduction to Evolutionary Algorithms*, Springer-Verlag, London, UK.

Appendix A: the normalized modal dispersion metric

Given a state-space description of a structural system with n DOF and provided a zero-mean Gaussian white excitation, a modal decomposition of covariance matrix of the output can be expressed as

$$\mathbf{\Gamma}_{yy}[h] = \sum_{k=1}^{2n} \mathbf{Q}_k \lambda_k^h \quad (\text{A.1})$$

with the involved quantities defined in Dertimanis (2013). At zero lag, the covariance matrix becomes

$$\mathbf{\Gamma}_{yy}[0] = \sum_{k=1}^{2n} \mathbf{Q}_k = \sum_{k=1}^{2n} \mathbf{Q}_k + \mathbf{Q}_k^* \quad (\text{A.2})$$

where the asterisk denotes complex conjugate. Eq. (A.2) can be used to the evaluation of the total vibration energy associated with the output of the involved state-space realization. Define the k th modal dispersion matrix

$$\mathbf{E}_k = \mathbf{Q}_k + \mathbf{Q}_k^* \quad (\text{A.3})$$

and the k th normalized modal dispersion matrix as the one with elements

$$[\mathbf{\Delta}_k]_{ij} = \frac{\mathbf{E}_k}{[\mathbf{\Gamma}_{yy}[0]]_{ij}} \quad (\text{A.4})$$

Then, any norm can be implemented towards the derivation of a metric that quantifies the significance of the k th mode. The L_2 modal dispersion metric is thus defined as

$$\delta_k = \|\mathbf{\Delta}_k\|_2 \quad (\text{A.5})$$

while its normalized version as

$$\bar{\delta}_k = \frac{\delta_k}{\max_k(\delta_k)} \quad (\text{A.6})$$