On Complex Spectra and Metastability of Markov Models

Sean Meyn, Gregory Hagen, George Mathew, Andrzej Banasuk

Abstract—The purpose of this paper is to develop methods for model reduction for diffusion processes that exhibit cyclic behavior. For this purpose we extend techniques based on the spectral theory of Markov processes to the case of complex spectra. The main idea is to augment the state process for the diffusion with a clock process. For each complex eigenvalue for the original diffusion there exists a real eigenvalue for the augmented process. Results concerning metastability (or quasi-stationarity) are then applied to the augmented process. For the special case of a linear diffusion in two dimensions, this is analogous to analyzing the process in a rotating coordinate frame. The results are illustrated through a linear diffusion, and an empirical model of combustion dynamics.

I. Introduction

Extensions of the classical Wentzell–Freidlin theory for model reduction have appeared in numerous papers over the past decade. Much of this work has concerned Markov processes that are reversible [14], [5], [2], [9], [3], [4]. The goal in these papers is to understand the statistics of exit times from a given subset of the state space.

Some results for non-reversible Markov chains are available. Fill's paper [8] extends the convergence-rate bound of Diaconnis and Stroock [6] to non-reversible Markov chains. For this purpose the transition matrix is replaced by its symmetrization, and the rate of convergence is bounded by the eigenvalues of the resulting self-adjoint matrix. These ideas are the basis of [10] that establishes exit time statistics from a set for a discrete-time non-reversible Markov chain.

Extensions of Wentzell–Freidlin theory to non-reversible processes appeared for the first time in [10]. The foundation of this paper is the theory of quasi-stationarity, building on the work of [7]. The main idea of [10] can be summarized as follows: Suppose that $\boldsymbol{X} = \{X(t) : t \in \mathbb{T}\}$ is a diffusion process evolving on $X = \mathbb{R}^d$, with transition semigroup denoted $\{P^t: P^t\}$

 $t \in \mathbb{T}$. We say that Λ is an eigenvalue with (non-zero) eigenfunction h if for each t,

$$P^t h = e^{\Lambda t} h$$

Suppose that Λ is real and negative. In this case we can assume that h is also real-valued, and we also assume that it is continuous. We would like to consider Doob's h-transform, $\check{P}^t := e^{-\Lambda t}I_h^{-1}P^tI_h$, where I_g is the multiplication operator: For each $x \in X$ and $A \subset X$ we have $I_g(x,A) = g(x)\mathbb{I}\{x \in A\}$. The h-transform, like importance-sampling, is intended to lead to a new Markov model whose properties provide insight into the problem of interest. Unfortunately $\{\check{P}^t\}$ is not a valid Markov semigroup since h may take on negative values. Instead we consider the following restricted definition.

Let M denote a connected component of the set $\{x: h(x) > 0\}$. We let $T_{\bullet} = \inf(t > 0: Y(t) \in \mathsf{M}^c)$, and for $t \in \mathbb{T}$ denote $t_{\bullet} = t \wedge T_{\bullet}$. The twisted semigroup is defined for each $t \in \mathbb{T}$, $x \in \mathsf{X}$, and $A \in \mathcal{B}$ (i.e. A Borel measurable) via,

$$\check{P}^t(x,A) := \frac{1}{h(x)} \mathsf{E}_x \big[e^{-\Lambda t_{\bullet}} h(X(t_{\bullet})) \mathbb{I}\{X(t_{\bullet}) \in A\} \big]$$
(1)

Under general conditions, it is shown in [10] that the twisted semi-group corresponds to a diffusion process on M that is exponentially ergodic. Exponential ergodicity of the twisted process then implies a form of quasi-stationarity, and from this it follows that the exit time from M is approximately exponentially distributed with parameter $|\Lambda|$.

The inspiration for consideration of the twisted process was the work of [7], and techniques from the large deviations analysis contained in [1], [11].

The main result of this paper is the extension of the results of [10] to the case in which $\Lambda \in \mathbb{C}$ is complex. The main idea is to augment the state process for the diffusion with a clock process. For each complex eigenvalue for the original diffusion there exists a real eigenvalue for the augmented process. Results concerning metastability contained in [10] are then applied to the augmented process.

This paper is organized as follows. In section II we present the problem setup and augment the state space with the clock process. Metastability of the twisted

S. Meyn is with the Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign. meyn@control.csl.uiuc.edu

G. Hagen, G. Mathew and A. Banaszuk are with United Technologies Research Center, 411 Silver Lane, East Hartford, CT. 06108 HagenGS@utrc.utc.com, MathewGA@utrc.utc.com, BanaszA@utrc.utc.com

process is established in section III. In section IV we present some examples.

II. SETUP

The setting described informally in the introduction is adopted throughout the paper: It is assumed that $\mathbf{X} = \{X(t) : t \in \mathbb{T}\}$ is a diffusion process evolving on $\mathsf{X} = \mathbb{R}^d$, with transition semigroup denoted $\{P^t : t \in \mathbb{T}\}$. Letting u denote the drift, and Σ the covariance matrix, the differential generator (see e.g. [12]) is defined for C^1 functions $h : \mathsf{X} \to \mathbb{C}$ by $\mathcal{D}h(x) :=$

$$\sum_{i} u_i(x) \frac{d}{dx_i} h(x) + \frac{1}{2} \sum_{ij} \Sigma_{ij}(x) \frac{d^2}{dx_i dx_j} h(x) \quad (2)$$

or, in more compact notation,

$$\mathcal{D} = u \cdot \nabla + \frac{1}{2} \operatorname{trace} \left(\Sigma \Delta \right)$$

For each $\beta > 0$ the resolvent kernel is given as the Laplace transform,

$$R_{\beta} := \int_0^\infty e^{-\beta t} P^t dt. \tag{3}$$

We write $R := R_{\beta}$ when $\beta = 1$.

It is assumed as in [10] that the diffusion is *V*-uniformly ergodic: For a probability measure ν on \mathcal{B} , some constants $b < \infty$ and $\overline{\Gamma} > 0$, a function $s : \mathsf{X} \to [0,\infty)$, and a $V : \mathsf{X} \to [1,\infty)$:

$$\begin{array}{ccc} \mathcal{D}V & \leq & -\overline{\Gamma}V + bs \\ R & \geq & s \otimes \nu. \end{array} \tag{V4}$$

The second inequality in (V4) means that the function s and probability measure ν are *small*. This terminology and the outer product notation are taken from [13]. This 'smallness assumption' is equivalently expressed,

$$R(x, A) \ge s(x)\nu(A), \qquad x \in X, \ A \in \mathcal{B}.$$

Suppose that \mathcal{D} has a complex eigenvalue Λ , which we write as

$$\Lambda = -\Gamma + i\vartheta$$

with $\Gamma > 0$, and $\vartheta \neq 0$, with associated eigenvector h. Consider the clock process defined by,

$$\Phi(t) = \Phi(0)e^{i\vartheta t}, \qquad t > 0, \tag{4}$$

with initial condition restricted to the unit circle in \mathbb{C} , which is denoted U. The clock process is Markov, as is the bivariate process,

$$Y(t) = \begin{pmatrix} X(t) \\ \Phi(t) \end{pmatrix}, \quad t \ge 0.$$

In fact Y is a diffusion on $Y = X \times U$ whose covariance matrix for y is given by,

$$\Sigma_Y(y) := \operatorname{diag}(\Sigma(x), 0). \tag{5}$$

Throughout the paper we adopt the notation $y = (x, \phi)$ for $y \in Y$, with $x \in X$, $\phi \in U$.

We define for each real $\beta \in \mathbb{R}$ the real-valued function,

$$g_{\beta}(y) = \operatorname{Re}\left(\left(e^{i\beta}/\phi\right)h(x)\right), \quad y \in Y.$$
 (6)

Proposition 2.1: For each $\beta \in \mathbb{R}$ the function g_{β} is an eigenfunction for the process Y, with eigenvalue $\Lambda_Y = -\Gamma$.

Proof: The differential generator for X can be extended in the obvious way to Y. Given the simple dynamics of Φ we have for any function $f: U \to \mathbb{C}$,

$$\mathcal{D}f\left(\phi\right) = i\vartheta\phi f'(\phi)$$

With $f(\phi) = 1/\phi$ the eigenfunction equation holds,

$$\mathcal{D}f(\phi) = -i\vartheta\phi/(\phi)^2 = -i\vartheta f(\phi), \qquad \phi \in \mathsf{U}.$$

Hence the generator applied to g_{β} gives,

$$\mathcal{D}g_{\beta}(y) = \operatorname{Re}\left((e^{i\beta}/\phi)\mathcal{D}h(x)\right)$$

$$+ \operatorname{Re}\left(-i\vartheta(e^{i\beta}/\phi)h(x)\right)$$

$$= \operatorname{Re}\left((e^{i\beta}/\phi)\Lambda h(x)\right)$$

$$+ -i\vartheta(e^{i\beta}/\phi)h(x))$$

$$= -\Gamma g_{\beta}(y), \qquad y \in Y.$$

III. THE TWISTED PROCESS

To define the twisted process we fix $\beta=0$ in the definition (6), and let M denote a connected component of $\{y:g_0(y)>0\}$. It is assumed that this set has nice topological properties: M is equal to the closure of its interior. Following [10], we define $T_{\bullet}=\inf(t>0:Y(t)\in \mathsf{M}^c)$, and the associated twisted process as follows:

The twisted process is the Markov process \check{Y} with state space M whose semigroup is defined using (1) based on the eigenfunction g_0 . Equivalently, for each $f \in L_{\infty}(M)$, and any $x \in M$, $\check{P}^s f(y) := \check{\mathsf{E}}_y[f(\check{Y}(s))] =$

$$\frac{1}{g_0(y)} \mathsf{E}_y[g_0(Y(s \wedge T_{\bullet})) f(Y(s \wedge T_{\bullet})) \exp((s \wedge T_{\bullet})\Gamma)]$$

The twisted process has a generator defined for C^2 functions $f: Y \to \mathbb{C}$ by,

$$\check{\mathcal{D}}f = g_0^{-1}\mathcal{D}(g_0f) + \Gamma f. \tag{7}$$

Two key assumptions are imposed in [10]: First, that the diffusion is hypoelliptic (which is used to conclude that the resolvent possesses a density with respect to Lebesgue measure). Second, it is assumed that the gradient of the eigenfunction does not vanish on the boundary of M. The gradient assumption is maintained here. To ensure that Y is hypoelliptic we assume that X is elliptic, meaning that its covariance is strictly positive. These assumptions are collected together as follows:

$$\Sigma(x) > 0$$
,

$$\nabla_{x} g_{0}(y) = \operatorname{Re} \left(\phi^{-1} \nabla h \left(x \right) \right) \neq 0, \forall y \in \partial M.$$
 (8)

It is not hard to see that the assumption (8) always fails when X is a diffusion in one-dimension. We see in the next section that it does hold in many examples, such as the linear diffusion in two or more dimensions.

The following result is a consequence of Theorem 3.7 of [10]. The reader is referred to this paper for a precise definition of metastability — Its main conceptual conclusion is that the exit time T_{\bullet} is approximately exponentially distributed, and that the process 'almost' reaches a 'local' steady-state prior to exiting M.

Theorem 3.1: Assume that (V4) is also satisfied for a continuous function $V: \mathsf{X} \to [1,\infty)$. Suppose that h is an eigenfunction with complex eigenvalue $\Lambda = -\Gamma + i\vartheta$ satisfying the following conditions:

- (a) $0 < \Gamma < \overline{\Gamma}$.
- (b) $g_0(y) > 0$ for all $x \in M$, and $g_0(x) = 0$ for $x \in \partial M := \overline{M} \setminus M$.
- (c) Condition (8) holds. Consequently, for $y \in \partial M$, $(\nabla g_0(x))^T \Sigma_Y(y) (\nabla g_0(y)) > 0$.
- (d) $K_n := \{x \in X : V(x) \le ng_0(x)\}$ is a compact subset of X for each n > 1.

Then

- (i) The escape-time from M for the twisted process is infinite a.s. for $\check{Y}(0) = y \in M$;
- (ii) The twisted process is \check{V}_1 -uniformly ergodic with $\check{V}_1(y) = V(x)/g_0(y), \ y \in \mathsf{Y}.$
- (iii) The set M is both metastable and V-metastable, with exit rate $\Gamma(M) = \Gamma_V(M) = \Gamma$. In particular,

$$\mathsf{E}[e^{\varepsilon T_{\bullet}}] \begin{cases} = \infty & \text{if } \varepsilon \geq \Gamma \\ < \infty & \text{otherwise.} \end{cases}$$

The proof of Theorem 3.1 amounts to establishing a version of (V4) for the twisted process. We can follow the same steps as in [10] to construct the required Lyapunov function.

For a given $0 < \alpha < 1$ write

$$\check{V}_1 := g_0^{-1} V, \ \check{V}_2 := g_0^{-1} g_0^{\alpha}, \quad \text{and } \check{V} := \check{V}_1 + \check{V}_2.$$

We denote $G_0 = \log(g_0)$, where g_0 is the eigenfunction for Y. From (V4) and the eigenvector equation we have,

$$\check{\mathcal{D}}\check{V}_{1} = [I_{g_{0}}^{-1}\mathcal{D}I_{g_{0}} + \Gamma I]g_{0}^{-1}V
= g_{0}^{-1}[\mathcal{D}V + \Gamma V]
\leq -(\overline{\Gamma} - \Gamma)\check{V}_{1} + bg_{0}^{-1}s$$

$$\check{\mathcal{D}}\check{V}_{2} = [I_{g_{0}}^{-1}\mathcal{D}I_{g_{0}} + \Gamma I]g_{0}^{-1+\alpha}
= g_{0}^{-1}[\mathcal{D}g_{0}^{\alpha} + \Gamma g_{0}^{\alpha}]
= \alpha g_{0}^{\alpha-1}[\Gamma - \frac{1}{2}(1-\alpha)\nabla G_{0}^{T}\Sigma_{Y}\nabla G_{0}].$$

Following arguments in [10], we obtain a version of (V4) for the twisted process: For a finite constant b_0 , and a compact set $S \subset M$,

$$\check{\mathcal{D}}\check{V} \le -\frac{1}{2}(\overline{\Gamma} - \Gamma)\check{V} + b_0 \mathbb{I}_S.$$

IV. EXAMPLES

We discuss an analytic example as well as an example motivated by an empirical model of limit-cycling combustion dynamics.

A. Ornstein-Uhlenbeck process

Consider the Ornstein-Uhlenbeck process,

$$dX(t) = AX(t) dt + dW(t), (9)$$

where W is a full-rank Gaussian process. Suppose that Λ is a complex eigenvalue, and v a (non-zero) left-eigenvector for A, satisfying

$$A^{T}v = \Lambda v$$
.

The generator for X shares this eigenvalue, and the function $h(x) = v^T x$ is an eigenfunction:

$$\begin{split} \mathcal{D}h\left(x\right) &= (Ax)^{\mathsf{\scriptscriptstyle T}} \nabla h\left(x\right) + \frac{1}{2} \mathrm{trace}\left(\Sigma \Delta h\left(x\right)\right) \\ &= x^{\mathsf{\scriptscriptstyle T}} A^{\mathsf{\scriptscriptstyle T}} v = \Lambda h(x). \end{split}$$

We now check to see if (8) is satisfied. We have,

$$\nabla_x q_0(y) = \operatorname{Re}(\phi^{-1}v), \quad y = (x, \phi) \in Y.$$

This is zero if and only if $\operatorname{Re}(\phi^{-1}v_k)=0$ for each $k=1,\ldots,n$. If this holds for some $\phi\in U$, it then follows that $v^*=i\phi^{-1}v$ is a purely real eigenvector for A, which is impossible since Λ is complex. We conclude that (8) is satisfied.

Consider the two-dimensional model with

$$A = \begin{bmatrix} -a & 1\\ -1 & -a \end{bmatrix}$$

where a>0. The matrix A possesses a pair of complex eigenvalues in the left-hand complex plane, satisfying $\Gamma=a$:

$$eig(A) = -a \pm i$$
.

A left eigenvector for A is given by $v^{T} = [-1, i]$, which gives

Re
$$(e^{-jt}v^{T}X(t)) = \cos(t)X_{1}(t) + \sin(t)X_{2}(t)$$
.

If X(0) satisfies $\operatorname{Re}(v^TX(0))>0$, we can expect that $\operatorname{Re}(e^{-jt}v^TX(t))>0$ for a period of time approximately exponentially distributed, with mean 1/a. Applying Theorem 3.1 we conclude that the first exit time $T_{\bullet}=\inf(t>0:\operatorname{Re}(e^{-jt}v^TX(t))=0)$ shares the following property with the exponential distribution:

$$\mathsf{E}[e^{\varepsilon T_{\bullet}}] \begin{cases} = \infty & \text{if } \varepsilon \geq a \\ < \infty & \text{otherwise.} \end{cases}$$

B. Empirical Model of Limit-Cycling Combustion Dynamics

We apply the analysis to a Markov model describing the nonlinear dynamics of limit-cycling combustion oscillations. The data was obtained from an experimental combustion rig described in [16]. The two-dimensional phase space was obtained as in [17] as follows. A POD analysis was done on the temporal flame images and the data was projected on to the first two dominant POD modes. The dynamics of the flame data projected on to this two-dimensional space is shown in Figure 1. The phase portrait shows a noisy limit-cycle where the direction of oscillation is in the clockwise direction.

A discrete time Markov model was constructed for the dynamics on this two-dimensional space. The eigenvalues are shown in Figure 2. The complex eigenvalues suggest cyclic behavior and a metastability analysis can be done using the corresponding eigenfunctions as described in the previous sections.

We describe the metastable sets associated with the eigenvalues shown on the right in Figure 2. In particular, the eigenvalue at $\lambda := |\lambda| e^{i\psi} = 0.98 + j0.995$ is associated with an eigenfunction that varies in the tangential direction and has no radial variation. The associated eigenvector h(x) is complex as shown in Figure 3. We take the clock process to be the discrete time equivalent of (4),

$$\phi_k = e^{i\psi k}\phi_0, \qquad k = 1, 2, \dots,$$

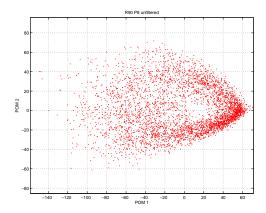


Fig. 1: Phase space showing a noisy limit-cycle of combustion dynamics. The oscillations move in a clockwise direction.

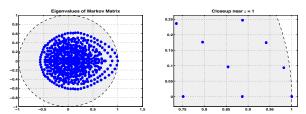


Fig. 2: Eigenvalues of the Markov matrix associated with the combustion dynamics data shown in Figure 1.

where ψ is the angle of the eigenvalue λ . Setting $\phi_0 = 1$, the eigenfunction of the associated twisted process is

$$g_0(y) = \operatorname{Re} \left(e^{-i\psi k} h(x) \right), \qquad y = (x, \phi) \in \mathsf{Y}.$$

Fig. 3: The complex eigenvector h(x) (magnitude on right, phase angle on left) associated with the complex eigenvalue $\lambda=0.98+j0.995$ shown in Figure 2.

The plot in Figure 4 shows the sign of $g_0(y)$ for different phase-shifts (i.e, after multiplication by $e^{-i\psi k}$ for different values of k.). Note how the sets with positive support and negative support rotate around the phase space and the exit time marks the point when the system exits one of these rotating sets (i.e., exhibits a phase-shift in its oscillations).

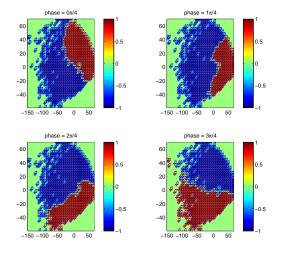


Fig. 4: The sign of the eigenfunction with a complex eigenvalue close to the unit circle, rotating with incremental phase-shifts of $\frac{\pi}{4}$ between 0 and $\frac{3\pi}{4}$.

The eigenvalue $\lambda=0.89$ is purely real and hence has a purely real eigenvector with no tangential variation, but variation in the radial direction. The sign of the eigenvector is shown in Figure 5. Since the eigenvalue is real, this eigenvector is not associated to any cyclic behavior. This eigenvector and the related exit time simply indicates when the system moves from a state of low amplitude oscillation to high amplitude oscillation, and vice-versa.

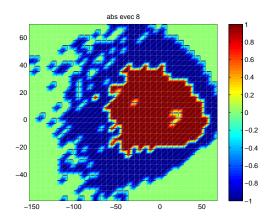


Fig. 5: The sign of the radial eigenvector of the Markov matrix.

Finally, the complex eigenvalue $\lambda = 0.851 + j0.99$ has an eigenvector exhibiting both tangential and radial components. Note again how the metastable set rotates around the phase-space, as indicated by the phase-shifted sign of the eigenvector shown in Figure

6.

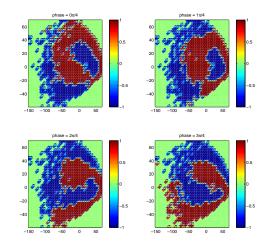


Fig. 6: The sign of the eigenvector with tangential and radial variation, shown rotating with incremental phase-shifts of $\frac{\pi}{4}$ between 0 and $\frac{3\pi}{4}$.

By examination of the magnitudes of the eigenvalues, the eigenvectors associated with these three metastable sets have decreasing mean exit times. This is intuitively confirmed by the fact that the sets become increasingly complicated. A hierarchy of such sets along with the spectral properties of the Markov matrix can be used to construct a reduced order model of the measured process through techniques described in [15].

V. CONCLUSION

We have presented a framework for analysing Markov models with semi-rotational dynamics by considering the complex spectra, and illustrated the approach using an application involving limit-cycling combustion oscillations. The ultimate goal of this research is to construct low order models that capture essential structure, such as the hidden Markov models proposed in [10]. The most interesting open problems are application specific. For example, can we justify the consideration of a two-dimensional model obtained from POD coefficients? If not, what are alternative approaches to treat the full-order Markov model?

ACKNOWLEDGEMENTS

This work was supported in part by AFOSR contract FA9550-07-C-0045, which is gratefully acknowledged.

REFERENCES

- S. Balaji and S. P. Meyn. Multiplicative ergodicity and large deviations for an irreducible Markov chain. Stoch. Proc. Applns., 90(1):123–144, 2000.
- [2] A. Bovier, M. Eckhoff, V. Gayrard, and M. Klein. Metastability in stochastic dynamics of disordered mean-field models. *Probab. Theory Related Fields*, 119(1):99–161, 2001.
- [3] Anton Bovier, Michael Eckhoff, Véronique Gayrard, and Markus Klein. Metastability in reversible diffusion processes. I. Sharp asymptotics for capacities and exit times. J. Eur. Math. Soc. (JEMS), 6(4):399–424, 2004.
- [4] Anton Bovier, Véronique Gayrard, and Markus Klein. Metastability in reversible diffusion processes. II. Precise asymptotics for small eigenvalues. J. Eur. Math. Soc. (JEMS), 7(1):69–99, 2005.
- [5] P. Deuflhard, W. Huisinga, A. Fischer, and Ch. Schütte. Identification of almost invariant aggregates in reversible nearly uncoupled Markov chains. *Linear Algebra Appl.*, 315(1-3):39–59, 2000.
- [6] Persi Diaconis and Daniel Stroock. Geometric bounds for eigenvalues of Markov chains. Ann. Appl. Probab., 1:36–61, 1991.
- [7] P. A. Ferrari, H. Kesten, and S. Martínez. R-positivity, quasistationary distributions and ratio limit theorems for a class of probabilistic automata. Ann. Appl. Probab., 6:577–616, 1996.
- [8] J. A. Fill. Eigenvalue bounds on convergence to stationarity for nonreversible Markov chains, with an application to the exclusion process. *Ann. Appl. Probab.*, 1(1):62–87, 1991.
- [9] W. Huisinga. Metastability of Markovian Systems: A transfer Operator approach to molecular dynamics. PhD thesis, Free University Berlin, 2001.
- [10] W. Huisinga, S. Meyn, and C. Schutte. Phase transitions and metastability in Markovian and molecular systems. *The Annals of Applied Probability*, 14(1):419–458, 2004.
- [11] I. Kontoyiannis and S. P. Meyn. Spectral theory and limit theorems for geometrically ergodic Markov processes. *Ann. Appl. Probab.*, 13:304–362, 2003. Presented at the INFORMS Applied Probability Conference, NYC, July, 2001.
- [12] A. Lasota and M. C. Mackey. Chaos, Fractals, and Noise: Stochastic Aspects of Dynamics. Springer-Verlag, New York, 1994.
- [13] E. Nummerlin. General Irreducible Markov Chains and Nonnegative Operators. Cambridge University Press, Cambridge, 1984
- [14] L. Rey-Bellet and L. E. Thomas. Asymptotic behavior of thermal nonequilibrium steady states for a driven chain of anharmonic oscillators. *Comm. Math. Phys.*, 215(1):1–24, 2000.
- [15] T. Runolfsson and Y. Ma. Model reduction of nonreversible Markov chains. Proceedings of the 46th IEEE Conference on Decision and Control, 2007.
- [16] K. A. Teerlinck, G. Hagen, A. Banaszuk, J. A. Lovett, J. M. Cohen, A. Ateshkadi, and B. V. Kiel. Demonstration of thermoacoustic instability control in a three-flameholder combustor. 53rd JANNAF (Joint Army-Navy-NASA-Air Force) Propulsion Meeting, Monterey, CA. Dec 5-8, 2005.
- [17] U. Vaidya and G. Hagen. Empirical model extraction from time resolved image data. Proceedings of the American Control Conference, 2008.