NORTHWESTERN UNIVERSITY

Reduced Models for Noise-driven Limit-cycle Oscillators

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Applied Mathematics

By

Avinash Jagdish Karamchandani

EVANSTON, ILLINOIS

March 2023

ABSTRACT

Reduced Models for Noise-driven Limit-cycle Oscillators

Avinash Jagdish Karamchandani

Motivated by rhythms in the brain, we investigate the synchronization of noisy and all-to-all pulse-coupled oscillators. We consider a case where the oscillatory excursions are of varying amplitude and where only sufficiently large excursions result in the output pulses that drive the interactions between the oscillators. In the regime of weak noise and weak interaction strength, we use a standard reduction to a Kuramoto-type phase-description with continuous, non-sinusoidal coupling and the associated Fokker-Planck equation for the population density of oscillator phases. Its linear stability analysis identifies the appearance of various emergent cluster states. Direct simulations of the oscillator equations reveal that, in order to achieve quantitative agreement with the phase reduction theory, the coupling strength and the noise have to be extremely small. Even moderate noise leads to significant variation in the timing of the large oscillations, which can enhance the diffusion coefficient in the Fokker-Planck equation by orders of magnitude. Introducing an effective diffusion coefficient extends the range of agreement significantly. We find that the effective diffusion coefficient, which can be computed efficiently via simulations of a single, noise-driven oscillator, is highly nonlinear as of a function of the input noise strength.

In a broader setting, for any limit-cycle oscillator that produces output conditional on the amplitude of its oscillations, we also treat the effective diffusion coefficient theoretically. In a novel framework, we model the outputs by "events" that correspond to distinguished crossings of a Poincare section. Using a linearization of the noisy Poincare map and its description under phase-isostable coordinates, we derive the effective diffusion coefficient for the occurrence and timing of the events using Markov renewal theory. We show that for many oscillator models the corresponding point process can exhibit

"unruly" diffusion: with increasing input noise strength, the diffusion coefficient vastly increases compared to the standard reduction analysis, and, strikingly, it also *decreases* when the input noise strength is increased further. The appearance of "unruliness" thus reflects a break down of the standard phase reduction and, in the context of coupled oscillators, of the Fokker-Planck theory. We provide a thorough analysis in the case of planar oscillators, which exhibit unruliness in a finite region of the natural parameter space.

Acknowledgements

I gratefully acknowledge support by Grant No. NSF-CMMI1435358 and RTG Grant Nos. NSF-DMS-0636574, NSF-DMS-1547394, and NSF-DMS-1937229.

Table of Contents

List of Tables

List of Figures

CHAPTER 1

Introduction

Many natural and engineering systems feature oscillatory dynamics. A key, quantitative feature of such dynamics is *phase*, which identifies the state of the oscillator's "internal clock". Systems with oscillatory dynamics can often be captured in a mathematical model that features an attracting limit cycle, and in that case the phase variable parameterizes the cycle. Notably, after any transients have subsided and the system has relaxed to the cycle, phase fully describes the state of the system. Indeed, two identical oscillators with different initial states can only differ after long times by their location on the cycle. After a long times, their time evolutions will be identical except for a fixed advance or delay relative to each other: a phase shift. Similarly, in situations where the oscillator is subject to exogenous forcing for only a finite duration of time, the only "memory" that the oscillator has of the forcing in the long time limit is a net shift in the phase.

Phase can be extended beyond the limit cycle and even transient dynamics can be characterized in terms of phase. Art Winfree [**85**] introduces *isochrons,* codimension-1 surfaces that foliate the basin of attraction of the limit cycle and each intersect the cycle at one point. And, under the oscillator's intrinsic dynamics, they evolve steadily into each other. Importantly, all trajectories that initially lie on the same isochron converge to the trajectory that starts at the point of intersection with the limit cycle. That is, all of the points on a given isochron have the same "asymptotic phase". Thus, the isochrons extend the definition of phase from the limit cycle to the full basin of attraction; phase is a well-defined coordinate on the basin.

In situations where the timing of certain dynamical features are of primary interest, a description of the dynamics by phase alone is desirable. Examples of such situations include questions of how a person and their sleep-wake cycle responds to rapid travel

across time zones, i.e. "jet lag", and the how the timing of action potentials in neurons are affected by synaptic currents. Phase reduction, a standard dimension reduction technique for limit cycle oscillators, eliminates all of the *other* coordinates that e.g. characterize the deviations from the limit cycle. In the absence of forcing, the equation for phase under phase reduction is closed (in the sense that it does not depend on the discarded variables), exact (in the sense that it is not an approximation), and trivial. In that situation, phase evolves steadily:

$$\dot{\phi} = f_0.$$

But, when the oscillator is subject to forcing, a single phase equation can only approximately capture the non-trivial dynamics that result. Indeed, in the presence of additive forcing $u(t)$, the phase evolution becomes

$$\dot{\phi} = f_0 + Z\left(\phi, \text{other d.o.f.s}\right) u(t),$$

where the Z quantifies the response of the phase to the forcing and typically depends on other degrees of freedom besides phase. In that case, phase reduction is only valid in the limit of strong scale separation: the attraction to the limit cycle must be strong in comparison with the strength of forcing. Geometrically, in the phase space, this means the oscillator's trajectory remains close to the limit cycle and therefore its state can be well-approximated as a position on the limit cycle. In this limit, the phase responds approximately via the *phase response curve* $Z_0\left(\phi\right)$, which is the value of Z evaluated at the point on the limit cycle with phase ϕ. Z_0 is a function of phase alone and approximately

closes the phase equation,

$$\dot{\phi} \sim f_0 + Z_0(\phi)\, u(t). \tag{1.1}$$

When there is not adequate scale separation, however, (1.1) will fail to capture transient or sustained dynamics that appear as a result of perturbative forcing, such as exogenous noise or interactions from other, coupled oscillators.

This thesis investigates a break-down of the standard phase reduction technique in settings where noisy oscillators interact with each other. Neural tissue is a prime example: the membrane voltage oscillations of a single neuron affect the electrical activity of other neurons, most commonly through synapses or gap junctions. Neural systems and other systems of coupled oscillators are therefore complex systems: the oscillators individually show nontrivial dynamics and population-level dynamics emerge from the interactions. Phase reduction has been used in an effort to understand the emergent behavior, like synchronization, in those systems; it simplifies the dynamics of the individual oscillators, and therefore focuses study on the nature of the interactions and the effect it has on the emergent dynamics. In Chapter 2 of this thesis, for example, we use phase reduction to investigate oscillators with delayed pulsatile interactions, and we highlight how the delay and pulse width affect the emergent states.

Complex systems have often been understood via phase transitions between different types of emergent states, and coupled oscillator systems are no exception. Consider a situation in which the oscillators not only interact with each other but are also subject to independent white noise. There are then two main parameters: the noise strength and the interaction strength. For sufficiently strong interactions and sufficiently weak

noise, one expects such a system to support coherent, emergent states, like synchrony. But, as the noise strength is increased or interaction strength decreased, coherence will be lost. Locating and understanding the transition that occurs as a function of those two parameters organizes our understanding of the emergent states. The analysis of the transition can be made tractable by taking the oscillators to be identical, all-to-all coupled, and infinite in number. Taken together with phase reduction, that means that the population of oscillators can be characterized by a density ρ of their phases, which is governed by a nonlinear "Fokker-Planck" equation,

$$\frac{\partial \rho}{\partial t} = D_{phase}\frac{\partial^2 \rho}{\partial \phi^2} - \frac{\partial}{\partial \phi}\left[f_0 \rho + K\left\{\text{nonlinear, nonlocal interaction term}\right\}\right]. \tag{1.2}$$

In (1.2), D_{phase} quantifies diffusion of phase density due to the noise, and its value is given by the phase reduction. K is the interaction strength. (1.2) makes predictions for which coherent states emerge and for what values of D_{phase} and K. An analysis of the linear stability of the incoherent state within the Fokker-Planck theory reveals a simple prediction for when the interaction strength is large enough to overcome the noise: it must be greater than a critical value which is proportional to the noise strength, $K > K_{crit} \propto D_{phase}$. We expect this prediction to be accurate in the limit of weak noise and weak coupling.

In this thesis, we identify a system of coupled oscillators for which the phase reduction fails dramatically to accurately predict the phase transition between the emergent coherent states and incoherence. The natural questions that follow, and the ones central to this thesis, are then: (1) "When does the Fokker-Planck phase transition analysis fail in coupled oscillator systems with additive noise?", (2) "Why?", and (3) "What can we

do about it?". We address question (2) first: we find that the phase reduction does not always accurately account for the response of the oscillators to the noisy drive. In a coupled system, it is the response of the oscillators' *outputs* to the noise that is important, where the outputs are those dynamical features that drive the oscillator interactions. In pulse-coupled systems, certain discrete dynamical "events" may trigger the pulses, and thus those events are the relevant outputs. The action potentials in neuronal oscillators, seen as "spikes" in the membrane voltage, are an example of such events, since they drive the synaptic interactions with other neurons. Thus, of interest in this example is the perturbation of the spike times due to the noise, which may not be well captured by the diffusion coefficient D_{phase} that is given by phase reduction. Answering questions (1) and (3) begins with a quantification of the actual noise response of the event timings, which we call the "effective phase diffusion coefficient" D_{eff}. As it turns out, the effective diffusion is often highly nonlinear as a function of the input noise strength, and its graph often takes on a characteristic, "unruly" shape. This is in stark contrast with the phase diffusion coefficient D_{phase}, which, as predicted by the standard phase reduction, is linear. It is when the effective diffusion deviates from the phase diffusion predicted by the standard phase reduction that the phase transition analysis under the Fokker-Planck framework fails. The effective diffusion coefficient is also the remedy: we find that the accuracy of (1.2) improves dramatically when we replace D_{phase} with D_{eff}.

The effective diffusion coefficient is thus the answer to all three of the above questions, and is the common thread throughout this thesis. The thesis is organized in two chapters. Chapter 2 is reproduced from a publication in Chaos, [**36**]. There we introduce the initial motivation for our work: rhythm formation in the mammalian olfactory bulb. We describe

rhythms within the Fokker-Planck framework as emergent states where the phases of the neuronal oscillators are clustered. We show, however, that the standard approach, where D_{phase} is defined via phase reduction, fails to accurately predict the onset of those states even for weak noise. We introduce a computational definition of the effective diffusion coefficient D_{eff} based on simulations of a single oscillator, and demonstrate that replacing D_{phase} with D_{eff} extends the accuracy of the theory to much larger noise and interaction strengths. In Chapter 3, we complement the computational treatment with a theoretical accounting of the effective diffusion coefficient. There, we move from a phase-oriented framework to an event-oriented framework, defining the events as arising from distinguished crossings of a Poincare section. Treating the events as a point process, we recast the effective diffusion coefficient as a point process statistic, and introduce a formula for it using Markov renewal theory. Using that formula, we investigate the origin of the unruly quality of D_{eff} as function of the input noise strength and demonstrate that "unruliness" is not uncommon amongst planar oscillators.

CHAPTER 2

Pulse-Coupled Mixed-Mode Oscillators: Cluster States and Extreme Noise Sensitivity

2.1. Introduction

The synchronization of oscillators coupled in a network is of great interest in many areas of science and engineering. Examples range from arrays of microwave oscillators [**86**] and lasers [**8**], networks of Josephson junctions [**81**] and optomechanical oscillators [**88**] to biological systems like pacemaker cells in the heart [**52**] and in the suprachiasmatic nucleus of the brain, which controls the circadian rhythm [**49**]. In the brain coherent rhythmic activity of large ensembles of neurons manifests itself in macroscopically observable rhythms, which have been found in many brain regions [**78**]. Among them the widely observed γ-rhythm (30-100 Hz) has been studied particularly extensively [**10, 9**].

For weak coupling one can generically describe coupled oscillators in terms of their phase. In this regime a substantial body of work has addressed the role of the network structure in the synchronization of phase oscillators with simple sinusoidal coupling, as is the case in the Kuramoto model [**37, 16, 64**]. The effect of modular network structure has also been addressed to some extent for strongly coupled, simple oscillators [**51**].

Even in the weak-coupling limit the coupling function is, however, in general not sinusoidal. Instead, higher harmonics are expected to come into play, which can affect the dynamics substantially; even just including a phase-shifted second harmonic can introduce qualitatively new dynamics [**29**]. The Fourier content of the interaction function depends on the type of coupling as well as on the properties of the periodic orbit of the uncoupled individual oscillators via their phase-resetting curve [**66**]. Thus, for oscillators undergoing complex periodic orbits like mixed-mode oscillations the phase-interaction function can exhibit a similarly complex phase dependence. This has been studied in detail for oscillators that are coupled diffusively, which in neuronal systems arises in the presence of gap

junctions; it was shown that such phase-interaction functions can stabilize a variety of states in which the oscillator phases form different numbers of clusters [**21**].

Most neurons in the brain interact, however, instead via chemical synapses, which provide output pulses that are driven by the spikes of the neurons. The shape of these pulses contributes to the Fourier content of the interaction function. Moreover, in contrast to diffusive coupling, pulse-coupling involves delays reflecting the finite rise time of the pulse and the propagation of the signal between neurons, which will be particularly significant if the transmission involves communication with other brain areas. In the weak-coupling limit the delays associated with pulse-coupling induce shifts in the phase dependence of the interaction function, which strongly affect the synchronization properties of the oscillators [**19, 23**].

A key requirement for the applicability of weak-coupling theory is that the periodic orbits of the oscillators are only weakly perturbed by their interaction. Since the subthreshold oscillations (STOs) that are part of mixed-mode oscillations typically have very small amplitude, even very weak perturbations may be able to trigger skipping or inserting of a STO cycle in these oscillations, amounting to large deviations from the periodic orbit. Among those perturbations is also noise that may be present in the circuit. This raises the question to what extent weak-coupling theory is suited to study the synchronization of mixed-mode oscillators.

An example motivating the present work is the mammalian olfactory system. There the inhibitory interaction between the principal neurons of the olfactory bulb can on the one hand be mediated by a interneuron population within the bulb itself, but on the other hand it can also involve cells in a second, cortical area. Intriguingly, the olfactory

bulb exhibits two different coherent rhythms: a faster one which is formed in the olfactory bulb alone, and a slower one which requires the input from the second brain area [**56, 57**]. Moreover, the principal cells that drive these rhythms exhibit various types of mixed-mode oscillations in which different numbers of small sub-threshold oscillations are interspersed between the large voltage spikes that are responsible for the output of the neurons [**14, 3**]. Due to the interplay between these mixed-mode oscillations and the delays of the pulsatile synaptic coupling the synchronization properties of a neuronal network like that of the olfactory bulb are expected to be complex.

In this chapter we investigate the synchronization properties of all-to-all coupled networks of identical mixed-mode oscillators in which the inhibitory pulse coupling, which drives synchronization, is counteracted by uncorrelated noise. We elucidate, in particular, the role that the delay in the interaction plays in the stabilization of different phase-locked states and the limitations that arise for the weak-coupling theory from the small size of the sub-threshold oscillations. In addition to synchronous solutions in which all oscillators fire at the same time, we find cluster solutions in which subsets of oscillators are phase-locked with each other. While one might have expected that the number of stabilized clusters would be directly related to the number of sub-threshold oscillations that the mixed-mode oscillations exhibit [**21**], this is not the case. Even for mixed-mode oscillations that include only a single sub-threshold oscillation the pulsatile coupling can - depending on the delay and the pulse width - stabilize a variety of solutions exhibiting up to 5 or more different clusters. Further, bistability between different cluster states as well as oscillations in the different order parameters can appear.

With regard to the validity of the weak-coupling theory we find that it captures the dynamics of the full system adequately only for extremely weak coupling and correspondingly extremely weak noise. This is the case even when the phase-resetting curve, which underlies the coupling term in a Fokker-Planck equation describing the synchronization, is adjusted for the finite coupling strength. However, our numerical simulations reveal that quantitative agreement for the onset of synchronization can be achieved over a much larger range of the diffusion coefficient if an effective diffusion coefficient is introduced in the Fokker-Planck equation. For the parameters investigated in this work this effective diffusion can be up to a factor of 500 larger than what is expected within the weak-coupling theory. This large boost in the diffusion is due to noise-driven skipping or adding of sub-threshold oscillations.

This chapter is organized as follows. In Section 2.2 we introduce the neuronal model that we use to investigate the interaction between mixed-mode oscillators and discuss some essential aspects of the mixed-mode oscillations. In Section 2.3 we use weak-coupling theory to determine the onset of synchronization and cluster formation and investigate the role of the shape and delay of the inhibitory pulse coupling. We compare the results from the weak-coupling theory with direct simulations of the spiking neurons in Section 2.4, where we introduce an effective diffusion coefficient. In Section 2.5 we use weakly nonlinear analysis of the Fokker-Planck equation to identify super- and sub-critical formation of cluster states. Within the Fokker-Planck framework we also find secondary bifurcations arising from the interactions between different cluster states and a tertiary Hopf bifurcation leading to oscillations of the order parameters that characterize those states. We confirm these results in direct simulations of the full spiking neuron model.

Currents:

Capacitive	Leak	Potassium	Sodium	Synaptic
$C_m = 1\frac{\mu\text{F}}{\text{cm}^2}$	$E_{leak} = -65\text{mV}$	$E_K = -72\text{mV}$	$E_{Na} = 55\text{mV}$	$E_{syn} = -70\text{mV}$
	$g_{leak} = 0.01\frac{\text{mS}}{\text{cm}^2}$	$\bar{g}_K = 50\frac{\text{mS}}{\text{cm}^2}$ $\bar{g}_{Ks} = 310\frac{\text{mS}}{\text{cm}^2}$ $\bar{g}_{Ka} = 10\frac{\text{mS}}{\text{cm}^2}$	$\bar{g}_{Na} = 500\frac{\text{mS}}{\text{cm}^2}$ $\bar{g}_{Nap} = 0.7\frac{\text{mS}}{\text{cm}^2}$	

Ion Channel Dynamics:

$$\alpha_\chi(V)\,[\text{ms}^{-1}],\ \beta_\chi(V)\,[\text{ms}^{-1}] = \frac{a+bV}{c+d\text{e}^{-\left(V-V_{half,2}\right)/k_{V,2}}}$$
$$\chi_\infty(V)\,[1],\ \tau_\chi(V)\,[\text{ms}] = y_0 + \frac{a+b\text{e}^{-\left(V-V_{half,1}\right)/k_{V,1}}}{c+d\text{e}^{-\left(V-V_{half,2}\right)/k_{V,2}}}$$

		y_0	a	b	$V_{half,1}$	$k_{V,1}$	c	d	$V_{half,2}$	$k_{V,2}$
m_K	α	-	1	0	-	-	0	1	11	19
	β	-	1	0	-	-	0	1	-113	-61
m_{Na}	α	-	0.32×50	0.32	-	-	1	-1	-50	4
	β	-	0.28×23	0.28	-	-	-1	1	-23	-5
h_{Na}	α	-	0.128	0	-	-	0	1	-46	-18
	β	-	4	0	-	-	1	1	-23	5
h_K	χ_∞	0.135	0.865	0	-	-	1	1	-14	-6.53
	τ	50	0	0	-	-	-	-	-	-
m_{Ks}	χ_∞	0	1	0	-	-	1	1	-34	6.5
	τ	7	0	0	-	-	-	-	-	-
h_{Ks}	χ_∞	0	1	0	-	-	1	1	-65	-6.6
	τ	200	220	0	-	-	1	1	-71.6	6.85
m_{Ka}	χ_∞	0	1	0	-	-	1	1	70	14
	τ	0	0	25	-45	-13.3	1	1	-45	-10
h_{Ka}	χ_∞	0	1	0	-	-	1	1	-47.4	-6
	τ	0	0	55.5	-70	-5.1	1	1	-70	-5
m_{Nap}	χ_∞	0	1	0	-	-	1	1	-51	5

Table 2.1. Mitral cell model parameters as used throughout this chapter. The dynamics of each m or h variable follow one of two functional forms, involving either α and β or χ_∞ and τ that are unique to that m or h.

2.2. A Neuronal Mixed-Mode Oscillator

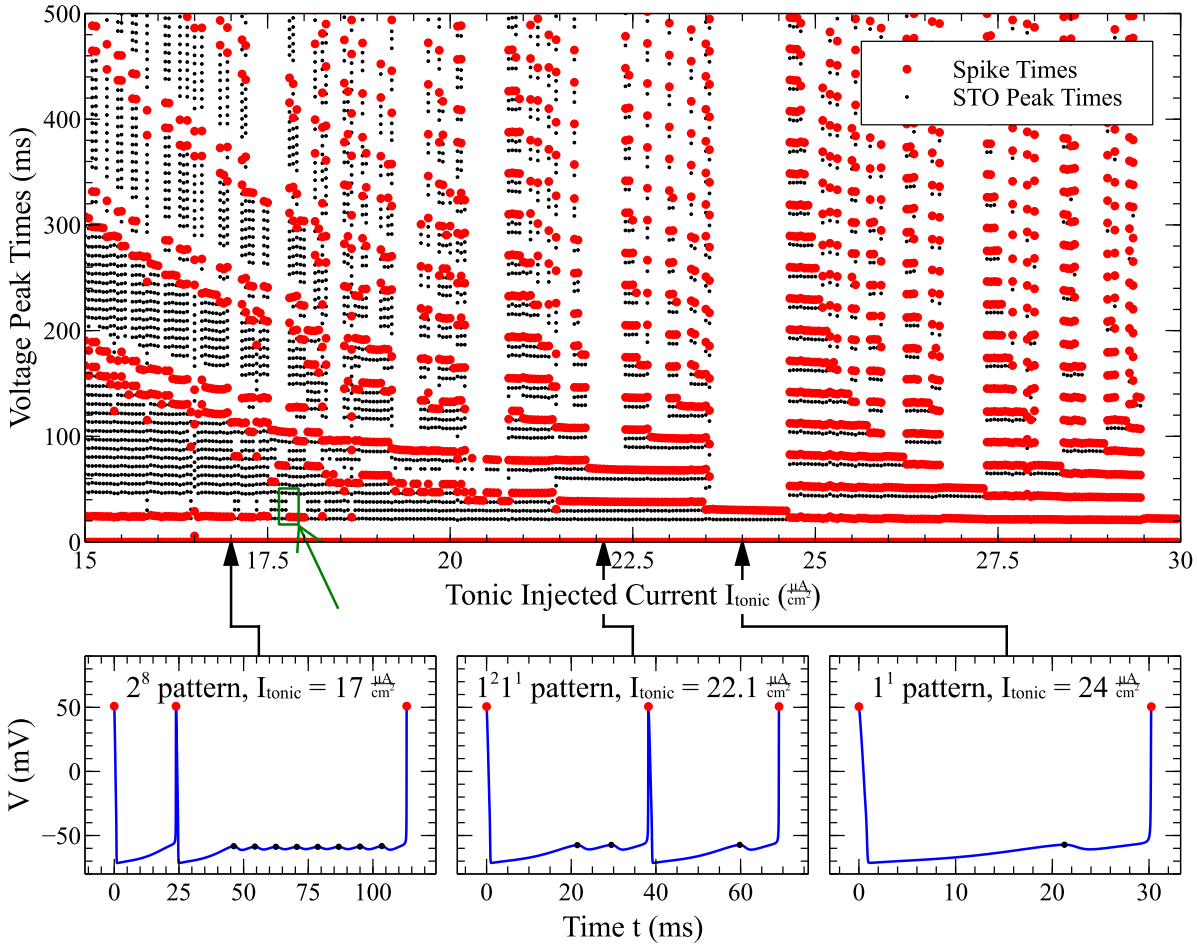

Figure 2.1. **A variety of limit cycle MMO patterns appear as the injected current I_{tonic} is varied.** The time between a spike and the following STO (or spike) is fairly consistent over a wide range of I_{tonic} and about three times the time between STOs (as seen e.g. in the green box).

As an example of a network of mixed-mode oscillators (MMOs), we consider a caricature of the mammalian olfactory system. The principal cells of the olfactory bulb, the mitral cells, exhibit mixed-mode oscillations and indirectly inhibit each other via two different synaptic pathways, one involving an interneuron population within the bulb and the other in addition a third population in a separate brain area. The two pathways are

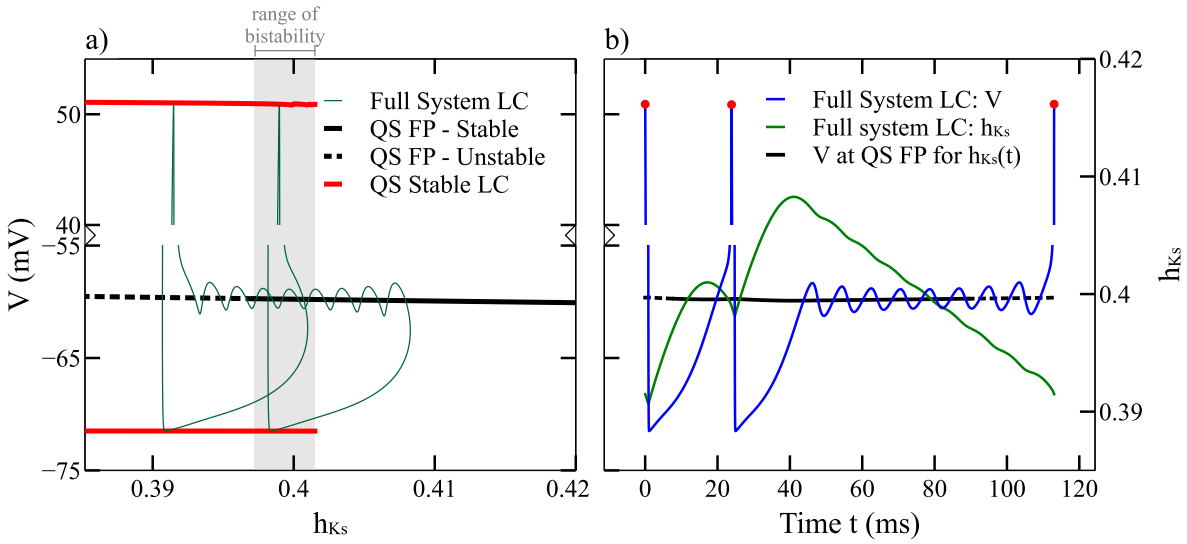

Figure 2.2. **A quasi-static bifurcation structure underlies the mixed-mode oscillations.** ($I_{tonic} = 17 \frac{\mu A}{cm^2}$) (a): The quasi-static bifurcation structure for slow h_{Ks} shows bistability between a fixed point ("QS FP") and a limit cycle (maximum and minimum voltage values depicted, "QS LC"). The trajectory of the full 2^8-limit cycle (LC) is super-imposed. (b): The temporal evolution of V and h_{Ks} along the full limit cycle.

expected to differ in particular in their effective propagation delays, and we focus in this work on the role of delay in the pulsatile interaction between the MMOs. In terms of the connectivity structure we consider here all-to-all coupling. While not biologically realistic, this allows a description of the onset of cluster formation in terms of a Fokker-Planck equation, which provides significant insight that is likely to carry over to more realistic circumstances.

We use a single-compartment conductance-based model that features a fast, inactivating sodium current, a persistent sodium current, a delayed rectifier potassium current,

an A-type potassium current, and a slow inactivating potassium current[**3, 7**],

$$
\begin{aligned}
C_m \frac{dV}{dt} &= g_{leak}\left(E_{leak} - V\right) + g_K\left(E_K - V\right) \\
&\quad + g_{Na}\left(E_{Na} - V\right) + g_{syn}\left(E_{syn} - V\right) \\
&\quad + I_{inj} \qquad (2.1) \\
g_K &= \bar{g}_K m_K^2 h_K + \bar{g}_{Ka} m_{Ka} h_{Ka} + \bar{g}_{Ks} m_{Ks} h_{Ks} \\
g_{Na} &= \bar{g}_{Na} m_{Na}^3 h_{Na} + \bar{g}_{Nap} m_{Nap,\infty}\left(V\right)
\end{aligned}
$$

$$
\frac{d\chi}{dt} = \begin{cases} \alpha_\chi\left(V\right)\left(1-\chi\right) + \beta_\chi\left(V\right)\chi, \\ \quad \chi \in \{m_K, m_{Na}, h_{Na}\} \\ \frac{\chi_\infty(V)-\chi}{\tau_\chi(V)}, \\ \quad \chi \in \{h_K, m_{Ka}, h_{Ka}, m_{Ks}, h_{Ks}\} \end{cases} . \qquad (2.2)
$$

Here the m and h are the activation and inactivation variables of the currents through various ion gates. The dynamics of each follow one of two equivalent functional forms, (2.2), involving functions of voltage that describe either the rate at which ion gates open and close (α and β) or the variable's steady-state value and it's timescale (χ_∞ and τ). The specific functions for each variable, as well as the constants that appear in (2.1) are given in Table 2.1. Compared to the mitral-cell models of Bathellier et al. [**3**] and Brea et al. [**7**] we modified some parameters, particularly in cases where their models do not agree with each other, so as to qualitatively match the experimental results in Bathellier et al. [**3**].

We model excitatory feed-forward input as a tonic current plus white Gaussian uncorrelated noise, i.e. for cells j and j' we have

$$
\begin{aligned}
I_{inj,j} &= I_{tonic} + \sqrt{2D_{in}}\,\eta_j(t) \qquad (2.3)\\
\langle \eta_j(t)\,\eta_{j'}(t')\rangle &= \delta_{j,j'}\delta(t-t').
\end{aligned}
$$

The effective inhibitory interactions between these cells are modeled by delayed synaptic pulses given by the difference of two decaying exponentials,

$$
\begin{aligned}
g_{syn}(t) &= \bar{g}_{syn}\sum_{j,n} s_{DE}(t - t_{j,n} - t_{delay}; t_w) \qquad (2.4)\\
&= \bar{g}_{syn}\sum_{j,n} s(t - t_{j,n}),
\end{aligned}
$$

with

$$
s_{DE}(t; t_w) = \begin{cases} \mathrm{e}^{-\beta_1 t/t_w} - \mathrm{e}^{-\beta_2 t/t_w}, & t \geq 0 \\ 0 & t < 0 \end{cases}.
$$

Here $t_{j,n}$ is the time of the n^{th} spike of the j^{th} cell and t_{delay} denotes the pulse delay. In this work we choose $\beta_1$and β_2 so that their ratio is fixed at 1.1151 and t_w is the half-height width of the pulses, which we vary as a parameter. We refer to $s(t)$ as the abstract pulse that incorporates both the delay and pulse width.

The dependence of rhythmogenesis on the noise strength D_{in}, the strength of interaction $\bar{g}_{syn}$, and the delay and shape of the synaptic pulses are of primary interest in this chapter. We investigate it for a few fixed values of I_{tonic} with the remaining parameters fixed as given in Table 2.1. Note that the functions τ_{h_K}, $\tau_{h_{Ka}}$, and $\tau_{h_{Ks}}$, which describe

the time scales of the inactivation variables h_K, h_{Ka}, and h_{Ks}, take on large values; these variables are slow.

As a function of the tonic current, the noiseless model exhibits an extensive variety of periodic MMO patterns characterized by large amplitude oscillations - the spikes (red dots, Figure 2.1) - and small amplitude sub-threshold oscillations (STOs) (black dots). To identify these patterns we use a customary notation [**15**] where $n_{spike,1}^{n_{STO,1}} n_{spike,2}^{n_{STO,2}} \ldots$ indicates a sequence of $n_{spike,1}$ spikes, $n_{STO,1}$ STOs, $n_{spike,2}$ spikes, $n_{STO,1}$ STOs, and so on. Voltage traces of some exemplary patterns that persist for a moderate range of injected current values are shown in Figure 2.1 (bottom). They all have relatively short periods suggesting that simple patterns tend to be more resilient to changes in input current. In this work we therefore focus on the simplest and perhaps most relevant patterns, e.g. the 1^1 pattern found at $I_{tonic} = 24 \frac{\mu\text{A}}{\text{cm}^2}$ and the 2^8 pattern at $I_{tonic} = 17 \frac{\mu\text{A}}{\text{cm}^2}$.

Through the use of geometric singular perturbation theory a great deal of progress has been made in understanding the mechanisms underlying MMOs in fast-slow systems [**15**]. We find here, though, that a less sophisticated and qualitative approach is sufficient to gain intuition about this model in the limited context of spike-driven interactions. In particular, we find it useful to consider the quasi-static system obtained by freezing the value of h_{Ks}. In the case of $I_{tonic} = 17 \frac{\mu\text{A}}{\text{cm}^2}$, the quasi-static system shows a relatively simple bifurcation structure as a function of the fixed value of h_{Ks}, exhibiting bistability between a limit cycle (spiking) and a spiral fixed-point (quiescence) (Figure 2.2). In the full system, the inactivating variable h_{Ks} shows a net increase along the quasi-static limit cycle, but decreases monotonically when the voltage is near the quasi-static fixed-point. This sets up a bursting cycle that switches between the spikes and the STOs, which form

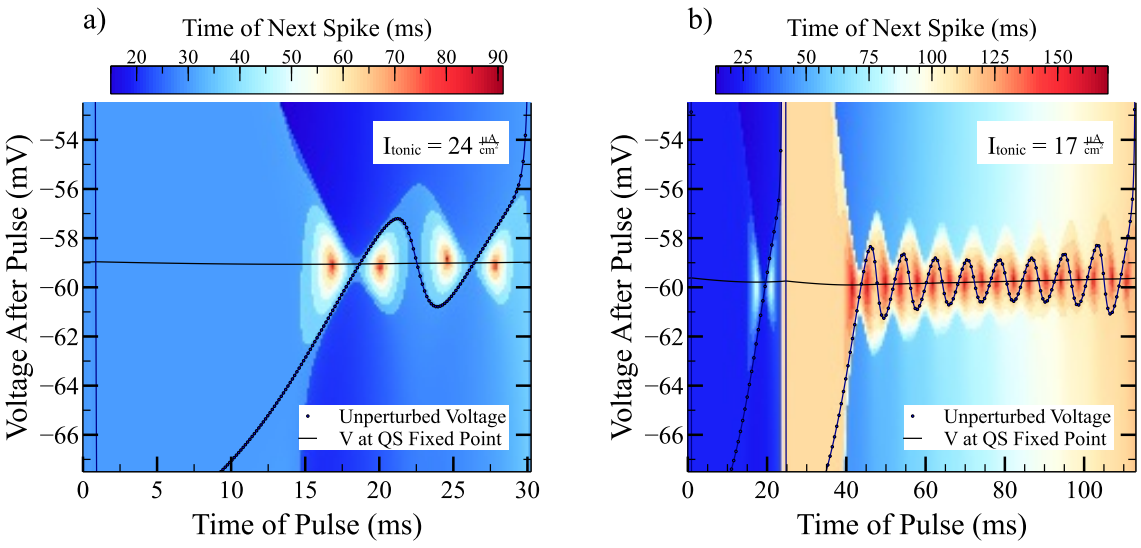

Figure 2.3. **The MMOs are sensitive to the timing and size of perturbative inputs.** δ-pulse inputs that instantaneously shift the voltage while leaving the other variables unchanged can significantly advance or delay subsequent spikes. Their impact depends on their timing relative to the first spike of the periodic MMO (horizontal axis) and the new value to which the voltage is shifted (vertical axis). The colors indicate the time between the first spike of the MMO and the first spike following the δ-pulse. In (a), $I_{tonic} = 24 \frac{\mu A}{cm^2}$, and in (b), $I_{tonic} = 17 \frac{\mu A}{cm^2}$.

as transient oscillations around the fixed point. Note that when the h_{Ks} is not frozen the small oscillations can extend to values of h_{Ks} for which the quasi-static fixed point is unstable by a delayed-bifurcation effect. The characteristic decreasing and then increasing magnitude of the STOs in MMO patterns like the 2^8 comes from the slow attraction to and then repulsion from the fixed-point. We expect that this mechanism operates at other values of the tonic injected current as well, resulting in some consistency in the model. In particular, the time between two voltage maxima is fairly consistent across I_{tonic} values, and, interestingly, the time following a spike before another spike or STO peak is about 3 times as large as the time following an STO peak (green box, Figure 2.1).

The quasi-static picture also provides some insight into how the presence of the STOs affects the system's response to external perturbations. Figure 2.3 shows two examples of the response of the cell to δ-pulse inputs that produce instantaneous shifts of the voltage away from the limit cycle. The color indicates the new, perturbed time of the next spike relative to the beginning of the MMO ($t = 0$). Notably the next spike can be delayed significantly when the perturbation drives the voltage close to its value at the quasi-static fixed point, producing through the delayed bifurcation effect far more STOs than on the limit cycle. This strong effect is not seen at all points along the STOs since the instantaneous perturbations only involve the voltage. Instead, as further investigation reveals, the effect is maximized close to the points where the activating variable m_{Ks} is near its quasi-static fixed point value. We therefore expect that the temporal form of STOs strongly affects the sensitivity of the system to injected or synaptic inputs.

2.3. Phase Resetting and Cluster Formation in the Weak-Coupling Theory

The response of limit cycle oscillations, like the MMOs of the mitral cell, to perturbations can be characterized in part by a phase resetting curve (PRC). States resulting from perturbations that push the system off the limit cycle but remain within its basin of attraction will converge back to the limit cycle. The only effect that persists after a long time is a shift along the limit cycle, i.e. a lag or lead, as compared to the unperturbed system. The PRC quantifies this effect, giving the change in phase, i.e. the position along the limit cycle, for some input as a function of the phase before the perturbation. PRCs therefore have been used extensively to capture the lowest order temporal features of neural oscillators under the effect of inputs.

Under the assumption of weak perturbative inputs (here weak synaptic pulses between cells and weak injected noise), limit cycle oscillators can formally be reduced to phase oscillators whose interaction is governed by the PRC [**22**]. In this limit, each cell evolves nearly periodically along its limit cycle, and deviations from periodic spiking impact other cells only at higher order. Spiking can therefore be considered periodic in the interaction terms s. For a given time t, the time $t_{j,n}$ of each past spike (cf. (2.4)) can then be parameterized as $t_{j,k,l}$ by the cell number j and the number of full cycles k since that spike. In the case that the periodic orbit has multiple spikes per cycle, l denotes the index of that spike within the cycle. Further, the time elapsed since the spike can be written in terms of the integer number of cycles elapsed k, the current phase $\phi_j(t)$ of the cell, and that spike's phase ϕ_l^{LC} within the cycle,

$$t - t_{j,k,l} \xrightarrow[\text{coupling}]{\text{weak}} \frac{k + \phi_j(t) - \phi_l^{LC}}{f_0}.$$

Here the phase ϕ_j runs between 0 and 1, f_0 is the natural firing rate of each cell, and ϕ_l^{LC} has fixed values for a given MMO.

Then by (2.4))

$$\begin{aligned} \frac{g_{syn}(t)}{\bar{g}_{syn}} &= \sum_j \sum_{k=0}^{\infty} \sum_{l=1}^{L} s\left(\frac{k + \phi_j(t) - \phi_l^{LC}}{f_0}\right) + h.o.t. \\ &\equiv \sum_j \sum_{l=1}^{L} \hat{s}\left(\phi_j(t) - \phi_l^{LC}\right) + h.o.t. \\ &\equiv \sum_j \hat{s}_{all}\left(\phi_j(t)\right) + h.o.t., \end{aligned} \tag{2.5}$$

where the infinite sum over k remains accurate to first order for exponentially decaying pulses s [**30**]. We define $\hat{s}$ as the 1-periodic sum of contributions from the infinite history of the pulses received as a results of each of the j^{th} cell's limit cycle spikes and $\hat{s}_{all}$ as the total pulsatile input from all of those spikes. Each oscillator's phase then evolves according to

$$\dot{\phi}_j = f_0 + \left[\bar{g}_{syn} \sum_{j'=1}^{N} \hat{s}_{all}\left(\phi_{j'}\right)\left(E_{syn} - V_{LC}\left(\phi_j\right)\right) + \sqrt{2D_{in}}\eta_j\left(t\right)\right] Z\left(\phi_j\right), \tag{2.6}$$

where N is the number of oscillators and the PRC Z characterizes the response of the phase to both the pulsatile and noise inputs. We define Z as giving the phase response to injected inputs; the response to synaptic inputs is then given by $\left(E_{syn} - V_{LC}\left(\phi\right)\right) Z\left(\phi\right)$. We note that we have omitted in (2.6) terms arising from Ito's Lemma [**87**]. Ultimately, the only effect of those terms is to slightly perturb the frequency f_0, the exact value of which is inconsequential in this analysis. Taking $\phi_j = f_0 t + \delta\phi_j$, (2.6) becomes

$$\dot{\delta\phi}_j = \left[\bar{g}_{syn} \sum_{j'=1}^{N} \hat{s}_{all}\left(f_0 t + \delta\phi_{j'}\right)\left(E_{syn} - V_{LC}\left(f_0 t + \delta\phi_j\right)\right) + \sqrt{2D_{in}}\eta_j\left(t\right)\right] Z\left(f_0 t + \delta\phi_j\right), \tag{2.7}$$

For small $\bar{g}_{syn}$ and D_{in}, the evolution of the phase perturbation $\delta\phi$ is slow compared to $1/f_0$ and we can apply the method of averaging. Averaging the coupling terms over one

period results in a continuous interaction function,

$$
\begin{aligned}
H\left(\delta\phi_j - \delta\phi_{j'}\right) &= \int_0^1 \hat{s}_{all}\left(\tau + \delta\phi_{j'}\right)\left(E_{syn} - V_{LC}\left(\tau + \delta\phi_j\right)\right) \\
&\quad \times Z\left(\tau + \delta\phi_j\right) d\tau \\
&= \int_0^1 \hat{s}_{all}\left(\phi - \left(\delta\phi_j - \delta\phi_{j'}\right)\right) \\
&\quad \times \left(E_{syn} - V_{LC}\left(\phi\right)\right) Z\left(\phi\right) d\phi.
\end{aligned} \tag{2.8}
$$

Averaging the noise term gives additive rather than multiplicative noise [**70**], where the new diffusion coefficient describes the diffusion of the phase variable:

$$
D_{phase} = D_{in} \int_0^1 Z^2\left(\phi\right) d\phi. \tag{2.9}
$$

Reintroducing the natural frequency f_0, one arrives at a generalized Kuramoto model,

$$
\dot{\phi}_j = f_0 + \bar{g}_{syn} \sum_{j'=1}^{N} H\left(\phi_j - \phi_{j'}\right) + \sqrt{2D_{phase}}\,\eta_j. \tag{2.10}
$$

For a large all-to-all coupled network ($N \to \infty$) of identical cells with uncorrelated noisy inputs, the population-level dynamics can be captured by a non-linear Fokker-Planck equation [**39**] for the probability density $\rho(\phi, t)$ of oscillators with phase ϕ at time t,

$$
\rho_t = D_{phase}\rho_{\phi\phi} - \frac{\partial}{\partial\phi}\left[f_0\rho + K\rho \int_0^1 H\left(\phi - \phi'\right)\rho\left(\phi'\right) d\phi'\right], \tag{2.11}
$$

where we take $K = N\bar{g}_{syn}$ to scale the synaptic coupling appropriately as $N \to \infty$. The emergence of rhythmic solutions can be studied by analyzing the destabilization of the incoherent state ($\rho = 1$) by Fourier modes, where the growth of a perturbation of a mode

$e^{2\pi\imath m\phi}$ with wavenumber m indicates the onset of coherent activity by the formation of m clusters in phase. The eigenvalues for these modes are [**39**]

$$\lambda_m = -(2\pi m)^2 D - (2\pi\imath m)[f_0 + KH_0 + KH_m], \tag{2.12}$$

where

$$H_m = \int_0^1 H(\phi)\, e^{-2\pi\imath m\phi} d\phi \tag{2.13}$$

and, as we will do in the remainder of this section, we have written D_{phase} as D for brevity. We will consider the location of the onset of cluster states ($\mathrm{Re}\lambda_m = 0$) as a function of the strengths of noise and coupling and of the pulse delay and shape. Note that in this limit of weak coupling K and weak noise D the stability only depends on the ratio $\frac{D}{K}$ with the incoherent state unstable to the m-cluster state for $D/K < (D/K)_{crit,m}$, where

$$\left(\frac{D}{K}\right)_{crit,m} = \frac{\mathrm{Im}H_m}{2\pi m}. \tag{2.14}$$

The linear stability analysis predicts that as noise is decreased or coupling is increased, the incoherent states is first destabilized by a coherent state that has $m = \mathrm{argmax}_{m'} \left(\frac{D}{K}\right)_{crit,m'}$ clusters.

Further information, such as bistability between the incoherent and cluster states can be obtained via weakly-nonlinear analysis. Taking $\rho = 1 + \mathrm{Re}\left\{Ae^{2\pi\imath m(\phi-(f_0+KH_0)t)}\right\}$, the evolution of the magnitude of A, $|A| = r_m$, is given by

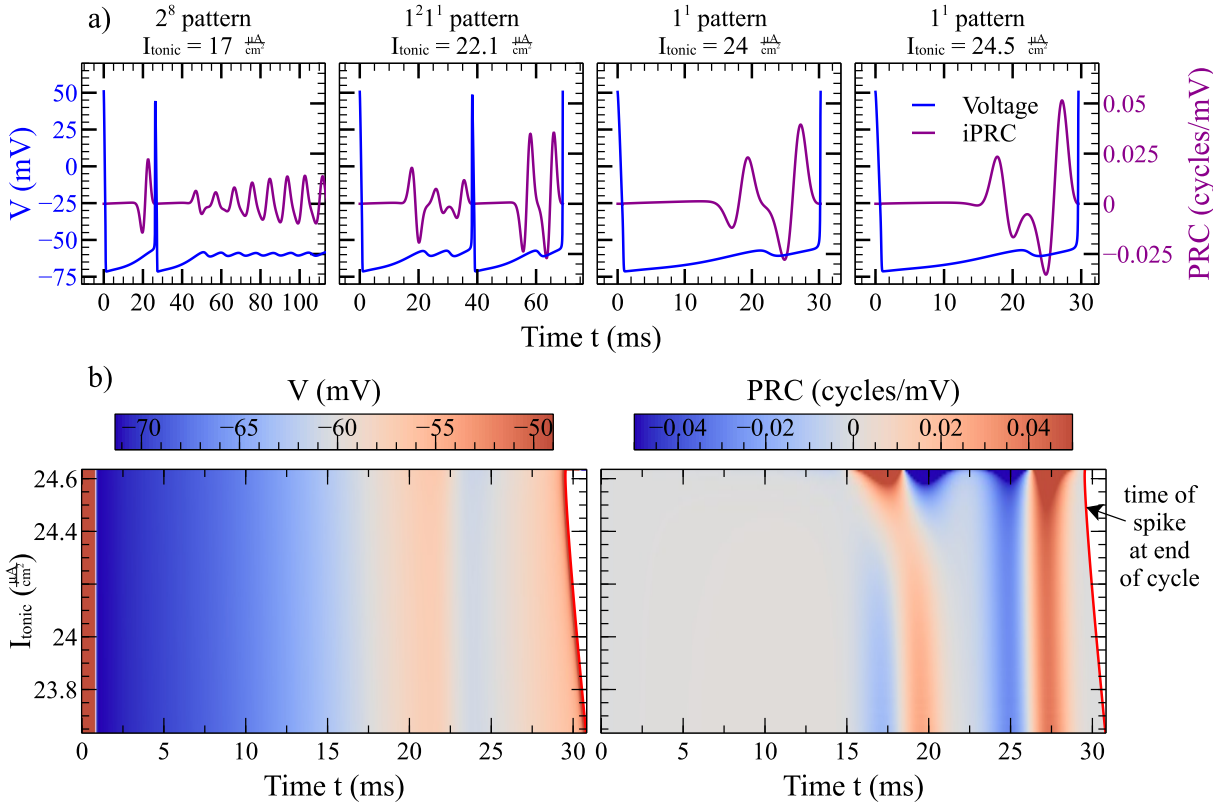

Figure 2.4. **The weak-input phase resetting of the mixed-mode oscillator is complex.** (a): Each spike and STO in the voltage trace (blue) results in an oscillatory "feature" in the iPRC (magenta). (b): The iPRC (right panel) can vary significantly across tonic injected current values, even when the voltage trace (left panel) remains approximately fixed.

$$\frac{1}{2\pi m K}\frac{dr_m}{dt} = \left(\gamma_0 - 2\pi m\frac{D}{K}\right) r_m + \gamma_2 r_m^3 + \gamma_4 r_m^5 + \mathcal{O}\left(r_m^7\right), \tag{2.15}$$

where

$$\gamma_0 = \mathrm{Im}\left\{H_m\right\}$$

$$\gamma_2 = \mathrm{Im}\left\{H_m \mu_2\right\}$$

$$\gamma_4 = \mathrm{Im}\left\{H_m \mu_2^2 \left[-\frac{3}{2} + \mu_4 g_{201} + \mu_2^* g_{210} + \frac{\mu_2^*}{\mu_2}\mu_4 g_{111}\right]\right\}$$

with

$$\begin{aligned}
\mu_2 &= \frac{2\left(H_m^* + H_{2m}\right)}{3H_m - H_m^* - 2H_{2m}} \\
\mu_4 &= \frac{H_m + H_{2m}}{2H_m - H_m^* + H_{3m}} \\
g_{201} &= -\frac{H_m^* + H_{3m}}{2\left(H_m^* + H_{2m}\right)^2} \\
g_{210} &= \frac{H^*}{2\left(H_m^* + H_{2m}\right)} \\
g_{111} &= \frac{H_m^*\left(H_{2m}^* + H_{3m}\right)}{\left|H_m^* + H_{2m}\right|^2},
\end{aligned}$$

and the star denotes complex conjugate. The amplitude equation predicts bistability of the incoherent state ($r_m = 0$) and an m-cluster state ($r_m > 0$) if the bifurcation is subcritical, $\gamma_2 > 0$, and the quintic term saturating, $\gamma_4 < 0$. In that case the bistability ranges between the primary bifurcation point $\frac{D}{K} = \left(\frac{D}{K}\right)_{crit,m}$ and the saddle-node bifurcation at $\frac{D}{K} = \left(\frac{D}{K}\right)_{SN,m} \sim \frac{1}{2\pi m}\left(\gamma_0 - \frac{\gamma_2^2}{4\gamma_4}\right)$. Thus, in the limit of weak coupling and weak noise, we can predict from the PRC and the stereotypical interaction pulses the nature of the emergent, population-level dynamics (cf. Figures 2.12,2.13 below).

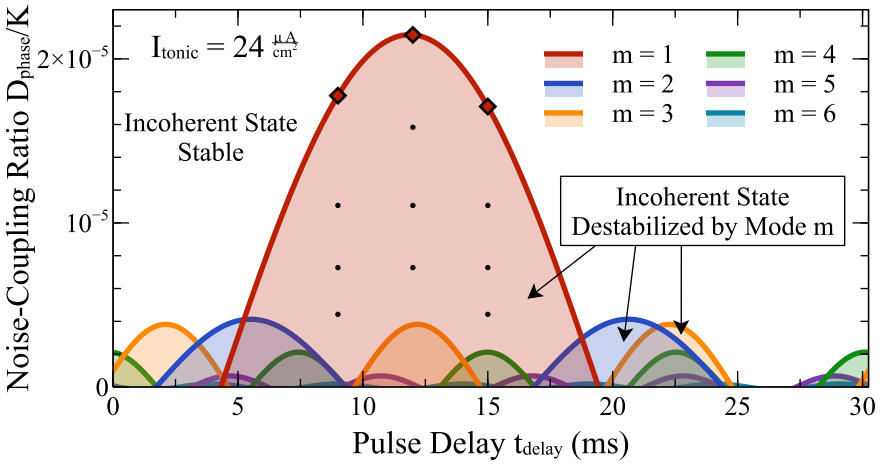

Figure 2.5. **A sample phase diagram over pulse delay and noise-coupling ratio.** Phase diagram for $I_{tonic} = 24\frac{\mu A}{cm^2}$ and a pulse half-height width t_w of about 12.5 ms. The incoherent state is linearly stable above the shaded areas. The critical noise-coupling ratio for mode m has a $\frac{T}{m}$-perioidic sinusoidal dependence on delay. The points marked with diamonds and dots are probed in simulations of the spiking model (Figure 2.9).

For the mitral cell model we use the adjoint method [**22**] to compute infinitesimal PRCs (iPRCs), i.e. the coefficient in the linearization of the PRC in the input strength. The timing of the perturbation relative to the STOs substantially modulates its impact (Figure 2.3). In the iPRC this manifests itself in substantial oscillations (Figure 2.4a), indicating rapid changes from a phase lead to a phase lag or vice-versa. We note, however, that while there seems to be an oscillatory feature in the iPRCs for every oscillation in voltage (spike or STO), a general pattern for the magnitude of those responses and the timing relative to the voltage oscillations is not apparent. Specifically, we do not expect that the iPRC can be well predicted from the voltage trace alone. Indeed, iPRCs can change dramatically across I_{tonic} even when the voltage trace remains largely unchanged (Figure 2.4b), reflecting the relevance of the other dynamical variables.

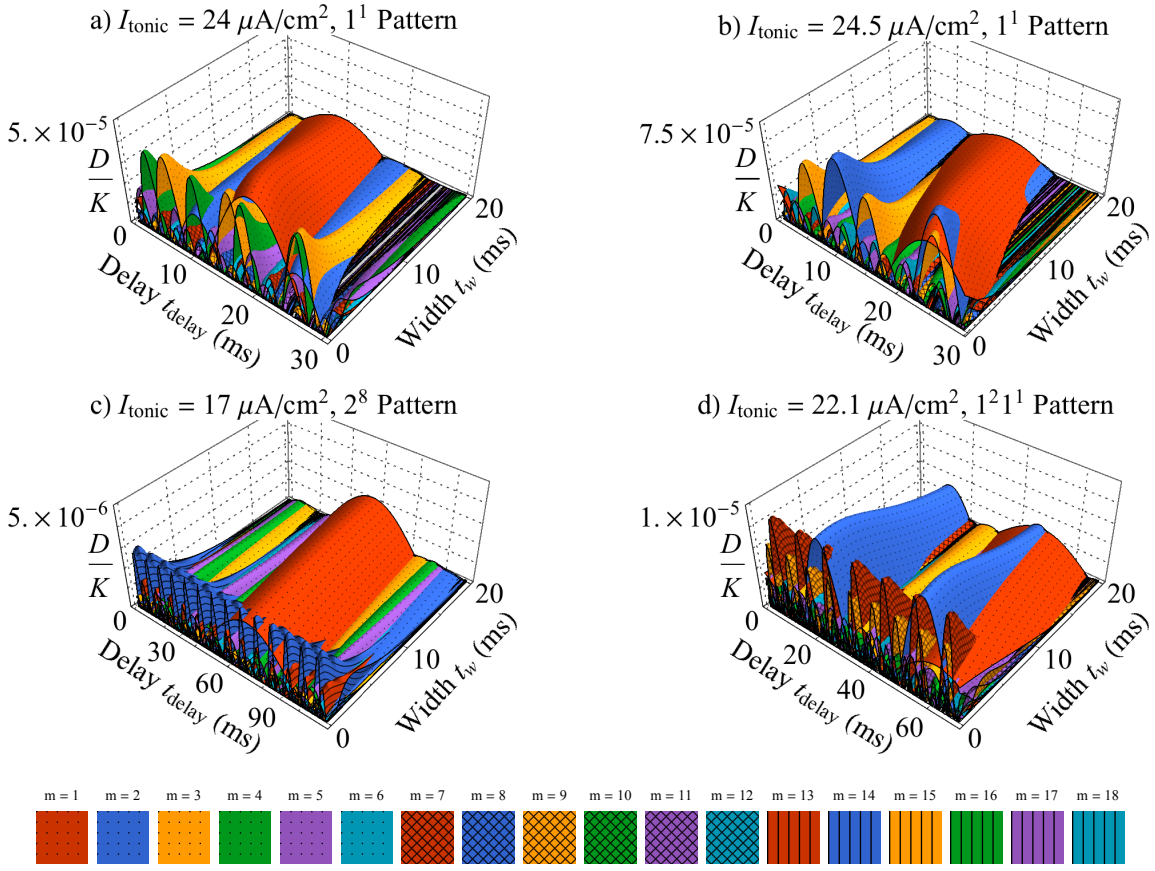

Figure 2.6. **Phase diagrams for different tonic injected currents show a variety of cluster states.** For smaller pulse width t_w the cluster states with high m become more relevant, in particular for those MMOs with many STOs.

Since the timing of the synaptic pulses strongly affects the interaction between the oscillators, the delay t_{delay} plays a key role in the formation of cluster states. This is in contrast to the diffusive coupling formed by gap junctions [**21**]. For MMOs with multiple spikes per period ($L > 1$) the timing of these spikes, characterized by their phases $\phi^{LC}_{l=1\ldots L}$ within the oscillation period, also affects the interaction significantly. To make the role of t_{delay} and ϕ^{LC}_l explicit, we write the synaptic input $\hat{s}_{all}(\phi)$ as a sum over the spikes within

a period,

$$\hat{s}_{all}(\phi) = \sum_{l=1}^{L} \hat{s}_{DE}\left(\phi - \phi_l^{LC} - f_0 t_{delay}\right), \tag{2.16}$$

where, like $\hat{s}$, $\hat{s}_{DE}(\phi)$ is defined as the periodic sum of the infinite history of single pulses $s_{DE}(t)$,

$$\begin{aligned}\hat{s}_{DE}(\phi) &= \sum_{k=0}^{\infty} s_{DE}\left(\frac{\phi(t)+k}{f_0}\right) \\ &= \frac{\mathrm{e}^{-\frac{\phi}{f_0 t_{syn,1}}}}{1-\mathrm{e}^{-\frac{1}{f_0 t_{syn,1}}}} - \frac{\mathrm{e}^{-\frac{\phi}{f_0 t_{syn,2}}}}{1-\mathrm{e}^{-\frac{1}{f_0 t_{syn,2}}}}.\end{aligned}$$

Similarly, we decompose H into the phase-shifted contributions from the different spikes within a period,

$$H(\phi) = \sum_{l=1}^{L} h\left(\phi - \phi_l^{LC} - f_0 t_{delay}\right), \tag{2.17}$$

where

$$h(\phi) = \int_0^1 \hat{s}_{DE}(\phi' - \phi)\left(E_{syn} - V_{LC}(\phi')\right) Z(\phi')\, d\phi'. \tag{2.18}$$

In the equation for the onset of cluster mode m, (2.14), the dependence on t_{delay} and ϕ_i^{LC} can be expressed explicitly using the Fourier coefficients

$$h_m \equiv \int_0^1 h(\phi)\, \mathrm{e}^{-2\pi\imath m\phi} d\phi$$

of h, resulting in

$$\left(\frac{D}{K}\right)_{crit,m} = \frac{1}{2\pi m}\mathrm{Im}\left\{h_m \sum_{l=1}^{L} \mathrm{e}^{2\pi\imath m\phi_l^{LC}}\right\} \times \sin\left(2\pi m f t_{delay}\right). \tag{2.19}$$

We summarize this relationship with phase diagrams showing the critical noise-coupling ratio for each Fourier mode as a function of delay, pulse shape, and/or tonic injected current. Figure 2.5 shows the critical ratio plotted as a function of the pulse delay for the 1^1 mixed-mode pattern at $I_{tonic} = 24$ for different values of m. The shading marks the regions in parameter space for which the linear stability calculation predicts the growth of a Fourier mode from the incoherent state. Depending on the coefficients γ_2 and γ_4 of the weakly nonlinear analysis (2.15), this is expected to lead to the formation of m-clusters, as discussed in Section 2.5 below. Notably, the regions corresponding to mode m are sinusoidal and $\frac{T}{m}$-periodic in the delay, where $T \equiv \frac{1}{f_0}$ is the period of the MMO (cf. (2.19)).

Comparison of the phase diagrams for different values of I_{tonic} elucidates which aspects of the different patterns' iPRCs ultimately affect the linear stability predictions. For example, $I_{tonic} = 24$ and $I_{tonic} = 24.5$ both produce very similar 1^1 voltage traces (Figure 2.4a), and both show similar sets of prominent modes ($m = 1, 2, 3, 4$) (Figure 2.6a,b). However, their iPRCs differ significantly from each other (Figure 2.4a), resulting in quite different linear stability surfaces (Figure 2.6a,b).

Depending on the MMO patterns that are obtained as I_{tonic} is varied, a wide variety of modes emerge readily, i.e. already for relatively small coupling or large noise (Figure 2.6). The most prominent m-cluster states originate from large Fourier modes m in the iPRC

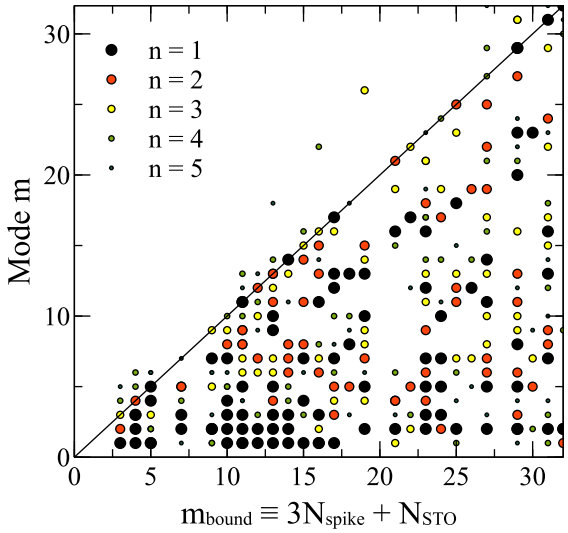

Figure 2.7. **Cluster number is roughly limited by the MMO pattern.** Across the MMO patterns shown in Figure 2.1, the $m = m_{bound}$ state is an approximate upper limit to each pattern's n^{th} most prominent cluster state for $n = 1, 2, \ldots, 5$.

(see (2.18) and (2.19)). Since the iPRCs generally have an oscillatory feature for each spike and for each STO that appears in the MMO pattern (Figure 2.4), longer patterns have iPRCs with higher spectral content and therefore support cluster states over larger ranges of m (Figures 2.6c,d). In fact, looking across tonic injected currents and focusing on the 5 most prominent modes, which have the largest onset values $(D/K)_{crit,m}$ when maximized over t_{delay} and t_w, we find that they seem to be approximately limited by an upper bound m_{bound} (Figure 2.7). We find that the bound can be written in terms of the number of spikes N_{spike} and the number of STOs N_{STO} in one period of the MMO,

$$m_{bound} \equiv 3N_{spike} + N_{STO} \approx \frac{T}{T_{STO}}. \tag{2.20}$$

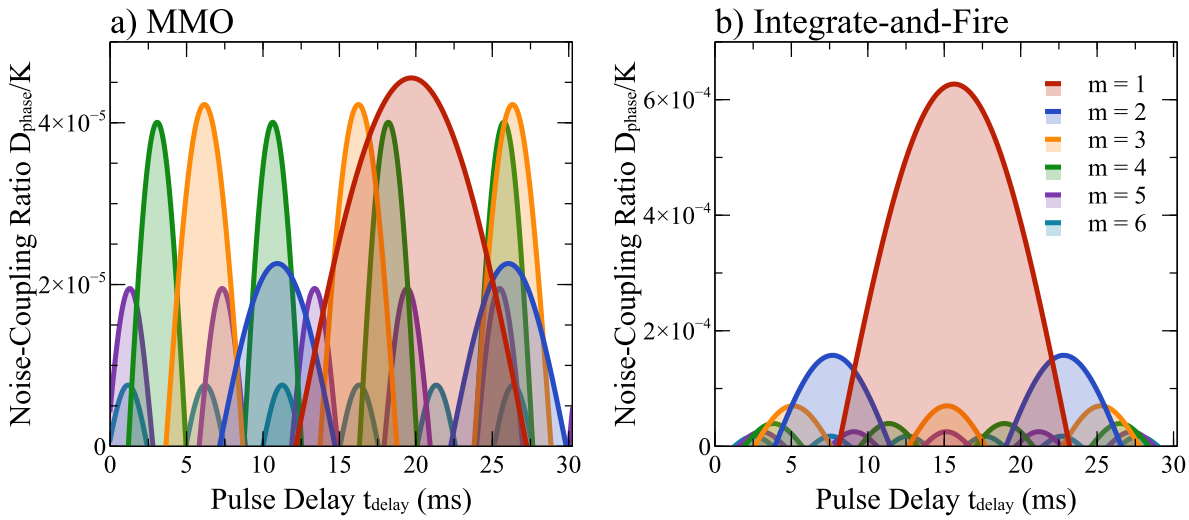

Figure 2.8. **A comparison of the MMO with an integrate-and-fire model:** phase diagrams in the limit of vanishing pulse width, $t_w = 0$. (a): For the MMO the higher modes appear as readily as the 1-cluster state (red region). (b): In contrast, in the integrate-and-fire model, which does not exhibit MMOs, the critical coupling strengths decay quickly with the mode number.

This reflects the fact that over a wide range of tonic injected currents the duration of spikes is approximately $3T_{STO}$ (Figure 2.1), where T_{STO} is the duration of one STO ($T_{STO} \approx$ 7.5ms). Since strong oscillatory features in the iPRCs are at a minimum separated by about T_{STO} (Figure 2.4), it is not unreasonable that its significant spectral content and therefore the resulting cluster states are approximately limited by m_{bound}.

The width of the synaptic pulses filters all modes that appear in the iPRC (cf. (2.18)), affecting those modes with the largest values of $\frac{m}{T}$ the most. For a given tonic injected current value, abundance of the states with large m relative to those with small m thusly decreases as pulse width is increased (Figure 2.6). This has the potential to produce

transitions from one cluster state to another as pulse width is increased, e.g. $m = 4$ to $m = 2$ for $t_{delay} \approx 10$ms and $I_{tonic} = 24 \frac{\mu A}{cm^2}$ (Figure 2.6a).

On the other hand in the extreme limit of a delta pulse the full effect of the MMOs in the mitral cell model is revealed with cluster states emerging roughly as readily as or more readily than the synchronous state (Figure 2.8a). In contrast, in an integrate-and-fire model (see Appendix) without mixed-mode oscillations, the synchronous state always persists for far stronger noise than cluster states, even for narrow pulses (Figure 2.8b).

We next test the predictive power of this theory and the resulting diagrams by comparison with numerical simulation of a network of the spiking conductance-based model (the "spiking simulations").

2.4. Weak-Coupling Theory Applies only for Extremely Weak Noise

To test the weak-coupling theory we simulate an all-to-all coupled network of $N = 1024$ spiking cells, focusing on the linear stability of the incoherent state. To assess the onset of cluster formation quantitatively we use the order parameters

$$r_m = \left| \frac{1}{N} \sum_{n=1}^{N} e^{2\pi \imath m \phi_n} \right|, \tag{2.21}$$

where the phase values ϕ_m are approximated from spike times by assuming ϕ varies linearly from 0 to 1 in each inter-spike interval. For $m = 1$, this is Kuramoto's order parameter for synchrony, and the r_m are appropriate extensions for cluster solutions. Note that in the limit $N \to \infty$, these are the magnitudes of the Fourier modes that destabilize the incoherent state in the Fokker-Planck Equation, (2.11).

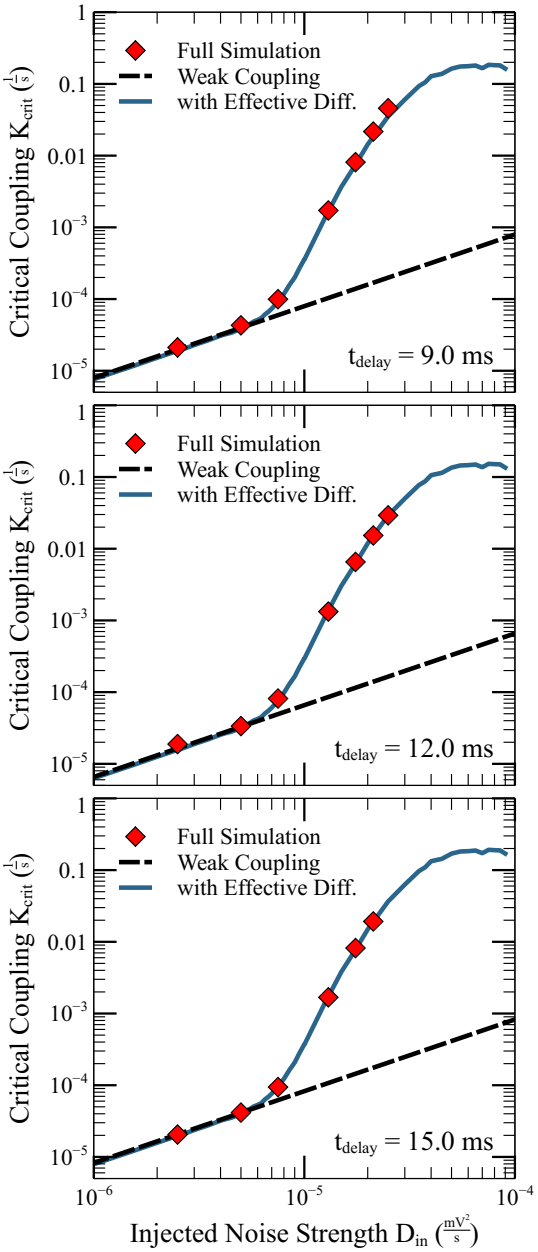

Figure 2.9. **Critical coupling values for the destabilization of the incoherent state as measured by simulation deviate greatly from the weak-coupling predictions.** Replacing the phase diffusion by an effective diffusion in the predictions improves agreement with simulation. Parameters: $I_{inj} = 24\frac{\mu\text{A}}{\text{cm}^2}$, $t_w \approx 12.5$ms (cf. Figure 2.5).

The order parameters from simulation are very noisy due to finite size effects, and we compensate by running many realizations. In an effort to compare with the weak-coupling results for the linear stability of the incoherent state, we measure the order parameter growth rates rather than amplitudes. For a limited number of delays and noise strengths, we estimate the critical coupling strength K_{crit} (Figure 2.5, diamonds) via simulations for a few coupling strengths greater than the critical value (dots). We use a linear fit of the growth rates as a function of the coupling coefficient, extrapolate them to find the $K = K_{crit}$ that corresponds to 0 growth, and estimate the error via boot strapping. As a function of noise strength, the measured K_{crit} agrees well with the theoretical, linear prediction for noise levels $D_{in} \lesssim 10^{-5}s^{-1}$. However, for even slightly larger values of the noise, the coupling levels needed to synchronize the spiking neurons are orders of magnitude larger than those predicted by the weak-coupling theory (Figure 2.9). The noise levels and the corresponding coupling strengths in the linear regime are extremely weak, as evidenced by the slow pace of the dynamics in those simulations. For example, for $D_{in} = 5 \times 10^{-6}\mathrm{s}^{-1}$, simulation runtimes of up to 2.7×10^{5}s (about 9 million limit cycle oscillations!) were required to measure the growth rate of the order parameter.

The weak-coupling theory assumes that the limit cycle is only slightly perturbed by the inputs from other neurons. This may not be the case for the coupling strengths used in Figure 2.9. However, for a large all-to-all coupled network in the incoherent state, the global synaptic current is nearly constant in time and close to the mean of the periodic synaptic pulse $\hat{s}$. To assess the impact of the finite coupling we include this mean value as a fixed synaptic current added to each cell, much like the tonic injected current. This modifies the limit cycle and its iPRC, but for $K \lesssim 0.1s^{-1}$ the resulting change in the

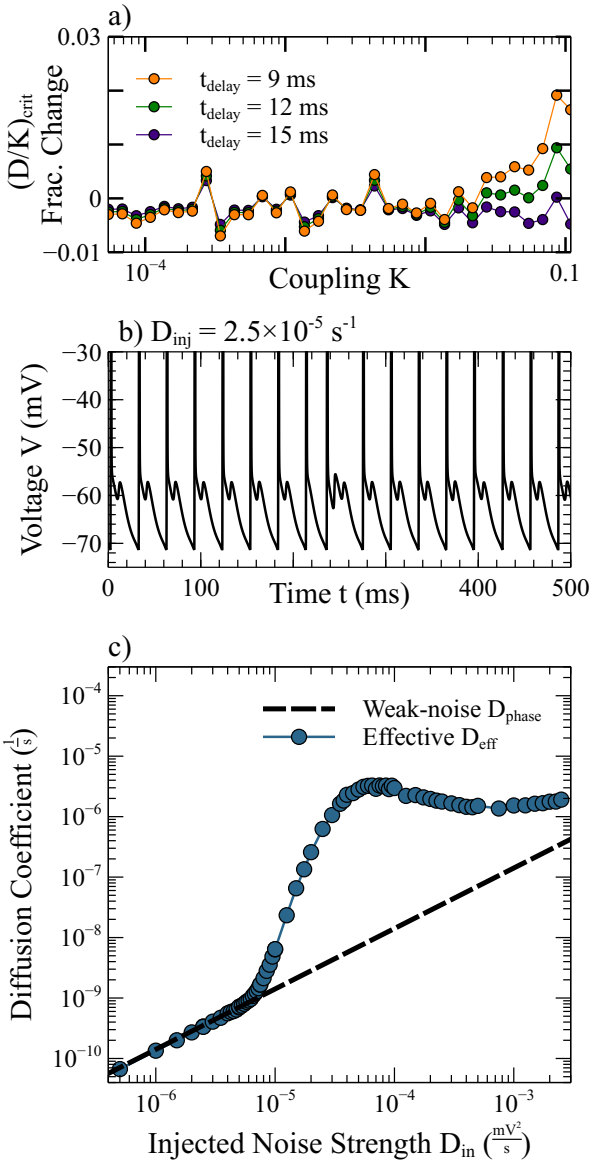

Figure 2.10. **The effect of finite noise and coupling.** (a): Finite coupling has only a small affect on linear stability. In the incoherent state, the finite coupling induces a mean synaptic current, shifting the limit cycle and its iPRC. This does not significantly affect the weak-coupling prediction in the regime in question (cf. Figure 2.9). (b): Even small noise can induce skipping and adding of STOs ($D_{in} = 2.5 \times 10^{-5} s^{-1}$). (c): Effective phase diffusion is greatly amplified compared to the linear weak-coupling prediction.

predicted critical diffusion-coupling ratio is less than 3% (Figure 2.10a). We expect that in our large but finite system, the impact of the finite coupling strengths on the critical coupling values is just as insignificant.

Finite noise also could be a culprit. Noise does in fact cause significant deviations from the limit cycle in the form of added or skipped STOs (Figure 2.10b). Even at low noise, for which adding events are rare but typically followed by skipping events and vice versa, the variability of the spike times over long time scales is greatly enhanced [**35**]. We find that we can measure the effect of these deviations via an approximate effective phase diffusion. In the absence on coupling, the Fokker-Planck equation gives an exponential decay of the order parameters with the decay rate proportional to the diffusion coefficient D_{phase},

$$\frac{dr_m}{dt} = -(2\pi m)^2 D_{phase} r_m. \tag{2.22}$$

Clearly, in making use of the weak-noise approximation for D_{phase}, this equation is inaccurate. But from numerical simulation over many oscillations of uncoupled spiking cells, we can still measure the decay rate in the first order parameter, r_1, and empirically define an effective diffusion coefficient, D_{eff}. While the theoretical weak-noise phase diffusion coefficient D_{phase} is proportional to the injected noise level (cf. (2.9)), the measured effective diffusion D_{eff} is strongly nonlinear (Figure 2.10c), increasing by a factor of 1000 when the injected noise strength D_{in} is increased from $10^{-5}s^{-1}$ by only a factor of 5. Replacing the phase diffusion coefficient D_{phase} in (2.11) by this effective diffusion coefficient D_{eff}, the theory agrees very well with the direct simulations over quite some range of the coupling and diffusion coefficient (Figure 2.9, blue traces).

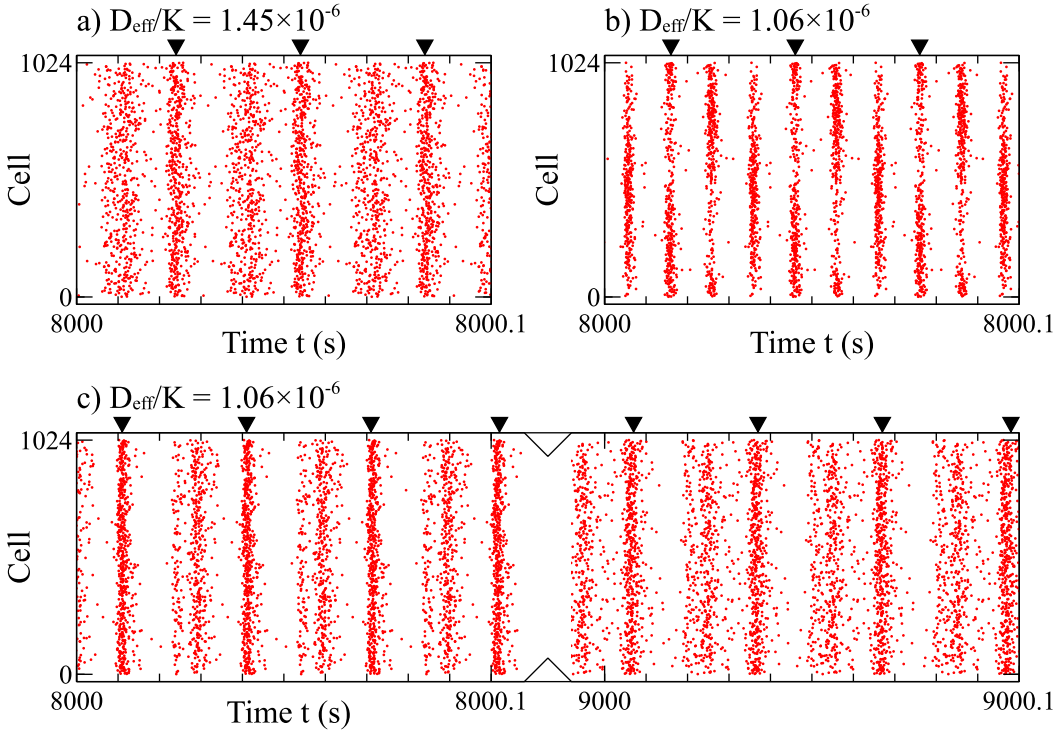

Figure 2.11. **The coupled MMOs exhibit complex emergent states:** cluster states in the spiking simulation for $I_{inj} = 24\frac{\mu A}{cm^2}$, $t_{delay} = 3.62$ms and $K = 0.012s^{-1}$. The triangles indicating the spike times of a single cell. (a) shows a moderately strong 2-cluster with $r_2 \approx 0.6$ ($D_{in} = 2.37 \times 10^{-s}s^{-1}$), (b) a strong 3-cluster with $r_3 \approx 0.9$ ($D_{in} = 2.25 \times 10^{-s}s^{-1}$), and (c) an oscillatory multi-cluster state that shows at times a 2-cluster and at others a state with 2-, 3-, and 5-cluster qualities ($D_{in} = 2.25 \times 10^{-s}s^{-1}$, cf. Figure 2.14b.i,ii).

2.5. Complex Dynamics of Cluster States

Having established the predictive power of the weak-coupling theory when using the effective diffusion, we now turn to exploring dynamics that appear in the Fokker-Planck equation (2.11) and how they reflect the dynamics of the full spiking system.

The presence of high Fourier modes in the interaction function allows states that are more complex than found in Kuramoto's model with sinusoidal coupling. The most

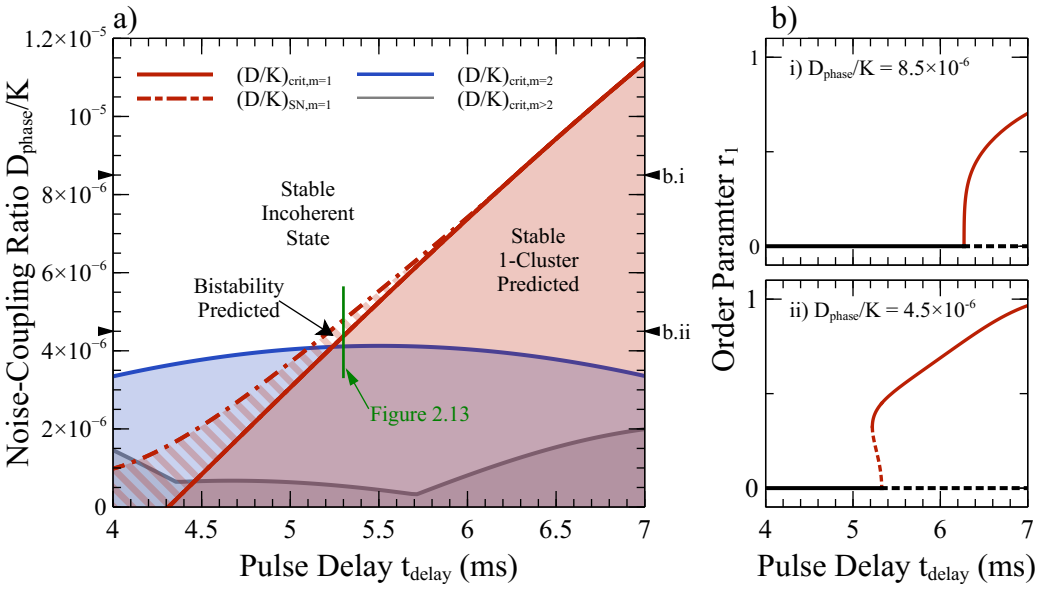

Figure 2.12. **Weakly-nonlinear analysis predicts bistability of the 1-cluster and the incoherent state.** (a): Phase diagram as in Figure 2.5 with predicted saddle-node bifurcation of the 1-cluster included (cf. (2.15) with $m = 1$). Bistability is predicted in the diagonally hashed area. The vertical line at $t_{delay} = 5.3$ms indicates the parameter range covered in the bifurcation diagram shown in Figure 2.13 below. (b): Bifurcation diagrams obtained by varying the delay with noise and coupling fixed at the values indicated by black triangles in (a). (b.i) shows a supercritical bifurcation and (b.ii) a subcritical primary bifurcation with a stable upper branch arising in a saddle-node bifurcation. Stable (unstable) branches are indicated by solid (dashed) lines.

prominent features that result are the cluster states predicted by the linear stability analysis. States with more than one cluster do in fact appear robustly in the spiking simulations, as seen in raster plots of the cell number against spike time (Figure 2.11). To illustrate that in these states the frequency of the population spike volleys is greater than the spiking frequency of the individual neurons, the spike times of an individual cell are marked by black triangles. Particularly noteworthy is the bistability between a 3-cluster

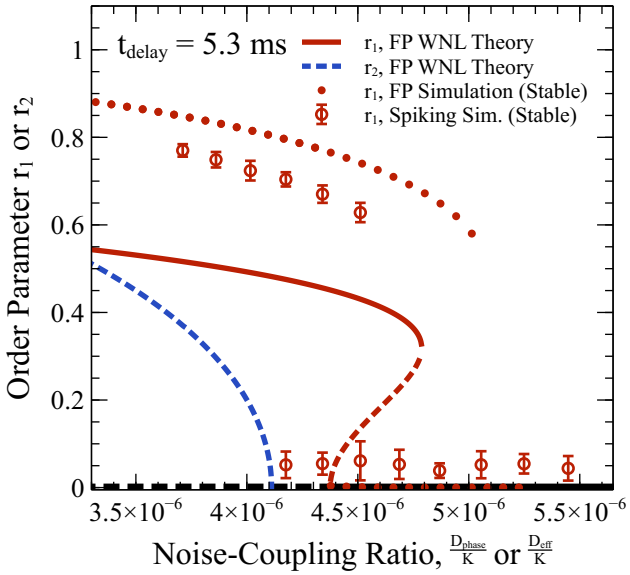

Figure 2.13. **Simulations confirm the bistability.** Simulations of the Fokker-Planck equation (solid circles) and the spiking system (open circles) both confirm the bistability predicted by the weakly nonlinear theory for the Fokker-Planck equation (red lines) for $t_{delay} = 5.3$ms (cf. vertical green line in Figure 2.12a and (2.15) with $m = 1$). The weakly nonlinear result for the 2-cluster (blue line, $m = 2$) is included for reference.

state and an oscillatory state that is predominantly a 2-cluster state but has strongly oscillating 3- and 5-cluster components (cf. Figure 2.14 below).

The weakly-nonlinear analysis for the order parameter r_m, (2.15), predicts parameter regions of bistability between the incoherent and the cluster states (Figure 2.12). Specifically, the bifurcation of the 1-cluster state is found to be subcritical for delays less than 6.2ms. The analysis to quintic order predicts, in addition, a saddle-node bifurcation (dash-dotted line in Figure 2.12a) and an upper, stable branch (Figure 2.12b.ii).

In fact, for a delay of $t_{delay} = 5.3$ms, we do find bistability in the spiking simulations as predicted by the weakly-nonlinear analysis (Figure 2.13). While the location of the saddle-node bifurcation and the amplitude of the upper branches differ quantitatively

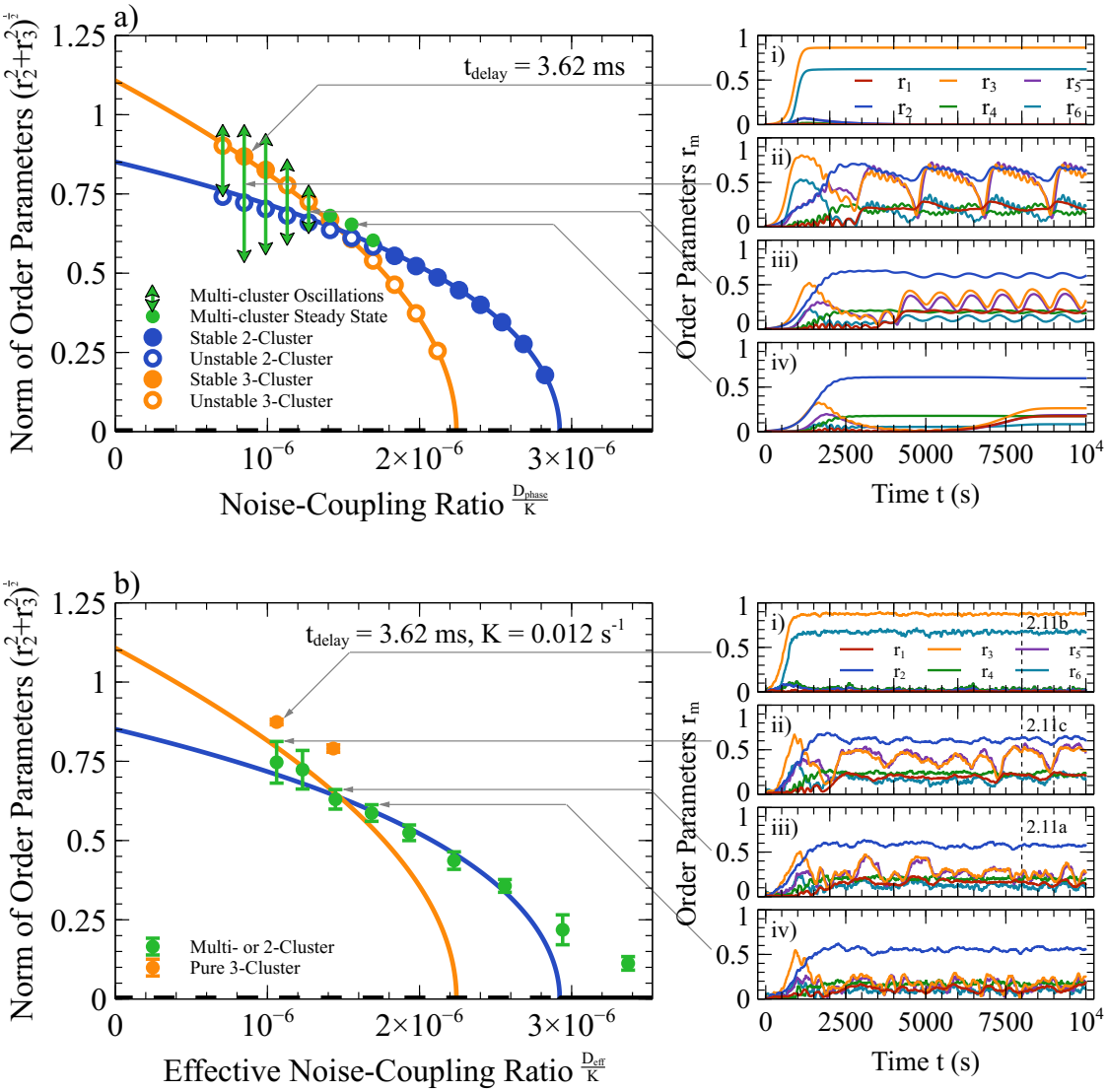

Figure 2.14. **The Fokker-Planck theory compares well with the full spiking simulation once D_{phase} is replaced with D_{eff}.** Simulations of (a) the Fokker-Planck equation and (b) the spiking model show bistability between a 3-cluster steady state (right panels, i) and a multi-cluster oscillatory state (ii-iii), as well as a steady multi-cluster state (iv). In the bifurcation diagrams in the left panels, arrows (a) and error bars (b) show the full range of the stable oscillations and the standard deviation across the last 4000 s of the simulation, respectively. The vertical dashed lines (b.i-iii) indicate the origin of the raster plots in Figure 2.11.

from the values predicted by the weakly-nonlinear analysis, the qualitative agreement is good. It should be noted that the noise inherent in the spiking simulations limits the extent to which the saddle-node bifurcation can be reached. We also show the results of simulations of the Fokker-Planck equation (solid circles), which agree much better with the spiking simulations than the weakly nonlinear theory. The discrepancy between the spiking simulations and the weakly nonlinear theory reflects therefore more a limitation of the weakly-nonlinear analysis than of the weak-coupling theory that it is based on.

The appearance of codimension-2 points, where the critical $\frac{D}{K}$ curves for two modes intersect, suggests potentially interesting mode-mode interactions. In fact, for $t_{delay} =$ 3.62ms, where the supercritical onset of the 2-cluster and 3-cluster states approximately coincide (Figure 2.5), numerical simulations of the Fokker-Planck equation reveal both steady and oscillatory mixed-mode states (Figure 2.14a). For decreasing noise-to-coupling ratio, the 2-cluster state is initially stable until a mixed-mode branch forms. That mixed-mode state undergoes a Hopf bifurcation, producing stable oscillations in the order parameter (Figure 2.14a.iii). Decreasing $\frac{D}{K}$ further, faster oscillations form on top of the existing ones. Interestingly, the pure 3-cluster branch also becomes linearly stable for smaller $\frac{D}{K}$, leading to the stable coexistence of the 3-cluster state and the oscillations for the same noise-coupling ratio (Figure 2.14a.i,ii). States strikingly similar to these appear in the spiking simulation (Figure 2.14b) when D_{in} is adjusted so that the effective diffusion coefficient D_{eff} matches the value of D_{phase} used in the Fokker-Planck simulations. While the order parameters are noisy due to finite-system-size effects, the mixed states, oscillations, and their coexistence with a 3-cluster state are clearly seen. Variations on top of the oscillations for the lowest noise-coupling ratio may reflect the small, fast oscillations

seen in the Fokker-Planck simulations, but they cannot easily be distinguished from noise. Overall, however, the correspondence between the weak-coupling Fokker-Planck results and the direct spiking simulations is quite good. In this case the weak-coupling theory is not only predictive, but it also provides insight into the behavior seen in the spiking simulation. Indeed, without the theoretical picture for comparison, one might conclude that the noisy oscillations obtained in the spiking simulations are just the result of noise.

2.6. Discussion

Motivated by the observation of different rhythmic activities in networks of neurons in the olfactory system [**56**], which are associated with the coherent activity of large ensembles of neurons, we have investigated the synchronization behavior of networks of mixed-mode oscillators. The role of STOs in neural rhythms has been previously studied computationally [**44, 53**]. It was found that STOs can enhance the rhythms, in their strength and in particular in robustness of their frequency. Via a semi-analytical approach, we aimed to elucidate the origin of the rhythmic cluster states arising from the phase resetting interactions of spikes and STOs. In contrast to previous work that investigated diffusive coupling of the oscillators [**21**], we considered pulse coupling and focused particularly on the impact of delays and pulse shape. In the neuronal contexts such delays can arise from the dynamics of the chemical synapses, from the propagation of the signal between neurons, and from the dynamics of intermediate neurons that may relay the signal. In the olfactory system that motivated our analysis most of the intermediate neurons are in the same brain area, but one of the observed rhythms is known to involve also feedback from a different brain area, which introduces a longer delay [**38**].

To gain semi-analytical insight into the system we considered weak coupling between the neurons, which allows a reduction of the high-dimensional equations for each neuronal oscillator to equations for the phases of the oscillators. Assuming all-to-all coupling and large network size allows a further reduction to a Fokker-Planck equation for the phase distribution of the oscillators. The mixed-mode character of the oscillations - together with the shape of the pulses mediating the coupling - manifests itself in the high Fourier content of the phase-interaction function. It is the large strength of the higher Fourier modes that leads to the rich behavior of this system.

The interaction we considered is inhibitory. A classic macroscopic rhythm that can arise from inhibitory coupling is the interneuron network gamma rhythm (ING), in which the period of the population rhythm is determined by the time it takes the inhibition to decay sufficiently for the neurons to resume spiking [**80, 10**]. Thus, in the ING-rhythm the period is always monotonic in the delay of the inhibition. By contrast, for the mixed-mode oscillators with weak coupling that we considered here the delay defines the relative phase of the Fourier modes in the phase-interaction function and determines the period of the population rhythm by selecting which cluster solution is stable. As a result, the period of the rhythm is not directly related to the delay; in fact, an increase in the delay often reduces the period. For instance, in Figure 2.6a increasing the delay from 10ms to 15ms increases the population frequency since the 2-cluster state is replaced by a state with 3 clusters.

We focused on a parameter regime in which the mixed-mode oscillation is relatively simple and contains only a single sub-threshold oscillation. Nevertheless, for sufficiently narrow interaction pulses and depending on the delay we find that the solutions that

emerge readily from incoherence can have even up to 4 or 5 clusters, reflecting population frequencies approximately up to the sub-threshold oscillation frequency. Generally, though, the shape of the voltage trace does not directly translate into specific cluster states. The phase-resetting curve, which underlies the phase-interaction function, cannot simply be deduced from the voltage trace, reflecting its dependence on the other variables describing the oscillator.

The relatively large amplitude of the higher Fourier modes of the phase-interaction function also introduces interesting nonlinear behavior of the order parameter. For instance, we find that the bifurcation off the incoherent state to the cluster states is often subcritical and our weakly nonlinear analysis also predicts hysteretic transitions between cluster states and the incoherent state. Moreover, numerical simulations of the Fokker-Planck equation reveal a tertiary Hopf bifurcation off a mixed-mode solution involving two different cluster states. Direct numerical simulations of the spiking-neuron model confirm the hysteresis as well as the oscillations in the order parameters.

The high Fourier content of the phase-interaction function discussed here reflects largely the complexity of the mixed-mode oscillations arising from the sub-threshold oscillations. The amplitude of high Fourier modes can also be enhanced by strongly skewing the phase-resetting curve. For quadratic integrate-fire neurons this has been shown to arise if there is strong spike-frequency adaptation [**39, 43**]. It also leads to the formation of clusters and to population rhythms with frequencies that are much higher than the spiking frequencies of the individual neurons.

Our quantitative comparison between the weak-coupling theory and direct numerical simulations of the spiking neurons showed agreement only for extremely small coupling

and correspondingly weak noise. Demonstrating agreement in this weak limit posed computational challenges as the simulation time required to identify the growth of cluster states in some cases exceeded 3×10^5s. This necessitated the use of high-performance GPU computation and limiting the networks to a population size N that could reasonably be simulated. A single simulation for 1024 oscillators and times on the order of 10^5 seconds still took over 100 hours of computation time on 1024 cores.

We found that the discrepancy for slightly larger noise is mostly due to skipping and inserting of sub-threshold oscillations, which arises already for quite small values of the noise. It can lead to an increase in the diffusion of the phase by three orders of magnitude even when the injected noise is only increased by a factor of 5 to a noise level at which only about 1% of the oscillation periods are affected by skipping or adding. After introducing an effective diffusion coefficient in the Fokker-Planck equation quite good agreement with the weak-coupling theory is achieved.

The fact that the small sub-threshold oscillations allow even weak noise to induce cycle skipping raises new interesting questions. How can this enhanced diffusion be described theoretically? It is not sufficient to consider a randomly selected sequence of step sizes drawn from a multi-modal distribution since it turns out that the skipping and adding events are strongly correlated over a long time scale[**35**].

CHAPTER 3

Events in Noise-Driven Oscillators: Markov Renewal Processes and the "Unruly" Breakdown of Phase-Reduction Theory

3.1. Introduction

3.1.1. Background: Phase reduction, Population Dynamics of Coupled and Noise-driven Oscillators, and Limitations

Phase reduction is a well-established dimension-reduction technique for dynamical systems that exhibit limit cycle oscillations [**85, 41, 20, 67**]. In appropriate limits, such as weak perturbative forcing or strong contraction to the limit cycle, it replaces the possibly high-dimensional oscillator state by a single phase variable ϕ, which represents the oscillator's "internal clock", and allows one to describe the long-time dynamics of the oscillator by a single differential equation for ϕ. Since this achieves a substantial reduction in the complexity of the description, phase reduction is a powerful tool for the analysis of individual and coupled oscillators [**41**] and their applications in many areas, including neuroscience [**74**] and circadian rhythms [**28**].

In the presence of noise and in the limit of infinitely many oscillators, globally coupled oscillators can be described by a nonlinear, non-local Fokker-Planck equation (1.2) for the population distribution $\rho(\phi, t)$ of the phase [**72, 75, 40, 36**]. In this framework, a key role in determining the tendency of the oscillators to synchronize is played by the diffusion coefficient of the phase, D_{phase}. It represents the leading impact of the noise on the oscillators, which tends to distribute their phases uniformly. (1.2) quantifies the intuition that, as long as the diffusion is small in comparison with the strength K of the interactions between the oscillators, coherent population-level dynamics, like synchronous states, will emerge. In particular, (1.2) predicts that coherent states arise for coupling strengths K above a critical value K_{crit} that is proportional to D_{phase}: $K > K_{crit} \propto D_{phase}$ (Figure

3.1c, dashed line). This prediction should be accurate in the limit of weak interactions and weak noise.

In Chapter 2, we applied the Fokker-Planck framework to mixed-mode oscillations comprised dichotomously of large- and small-amplitude oscillations (Figure 3.1a). This model arose in a neuronal context where only the large oscillations, which represent action potentials, lead to any output to other neurons via chemical synapses. Therefore, from the perspective of the other neurons in the network only these large-amplitude oscillations are relevant *events*, and it is their timing that determines the population-level dynamics. Numerical simulations reveal transitions from asynchrony to synchrony and more exotic coherent states, which depend on the strength of the noise D_{in}. Standard phase reduction provides a direct linear connection between D_{in} and D_{phase}, and thus the Fokker-Planck theory (1.2) predicts the onset K_{crit} for these coherent states as a function of the input noise strength D_{in}. For the mixed-mode oscillations investigated in Chapter 2, that prediction fails spectacularly when compared to simulation of the full system, even when the noise and the interactions are relatively weak (compare the dashed line and blue dots in Figure 3.1c). To wit, it vastly underestimates the strength of the oscillator interactions needed to overcome the independent noise and produce coherent dynamics states. In Chapter 2, we found that the amount of diffusion is substantially underestimated by the phase reduction theory, since D_{phase} does not take into account the fact that the oscillators interact only via the amplitude-dependent events. Standard phase reduction, it turns out, fails because it discards key amplitude information.

3.1.2. Going Beyond Standard Phase Reduction: Events and Effective Phase

The need to go beyond the standard phase reduction has been recognized previously in various other contexts, e.g. in shear-induced chaos [**46**] and systems involving multiple time scales [**54, 45**]. Recent works have considered augmenting the phase variable with one or more amplitude-like variables to increase the fidelity of the reduction (see e.g. [**79, 11**]). Of particular interest is the phase-isostable reduction [**84**], offering a coordinate system in which the unperturbed oscillator has especially simple dynamics.

Here we offer an *event-centric* extension to the standard phase reduction. In particular, we address the influence of noise on oscillators in which the oscillatory excursions can be divided into two classes: excursions that qualify as events and excursions that do not. More concretely, describing each oscillatory excursion as the crossing of a Poincare section, events correspond to a crossing within a limited domain of the Poincare section; crossings outside that domain constitute non-events. In the neuronal mixed-mode oscillators, for example, only the large-amplitude peaks in the voltage are events (action potentials), whereas small-amplitude extrema are non-events.

In noise-driven oscillators, the events occur in an irregular fashion and define a stochastic point process. We are particularly interested in the long-term statistics of the events that reflect the degree of diffusion in the stochastic process. The quantity that is of central interest for the description of the event irregularity is the growth rate of the variance of the total number of events with time, which we propose to call the "temporal variance growth rate" (TVGR)[1] and represent by the symbol $\mathcal{V}_E^{(t)}$. As it turns out, this quantity has a direct relationship to the phase diffusion coefficient D_{phase} in the limit of

[1]This quantity has also been referred to as "the slope of the variance-time curve" [**12**].

weak noise. Namely, $\mathcal{V}_E^{(t)} \sim 2D_{phase}$ as $D_{in} \to 0$. This connection between the event-based TVGR and the phase diffusion motivates us to introduce an *effective* phase that increases by 1 between events, whose diffusion coefficient is given by an *effective* phase diffusion $D_{\mathit{eff}} \equiv \frac{1}{2}\mathcal{V}_E^{(t)}$.

In Chapter 2, we suggested that D_{phase} be replaced by D_{eff} in (1.2) as a way to remedy the deficiencies in the Fokker-Planck theory. We determined D_{eff} computationally via simulations of a single, uncoupled oscillator and found that it indeed compensates for the discrepancy in the predictions and recovers the actual onset of coherent states (Figure 3.1b,c). The effective phase thus not only provides an intuitive picture of the long-time dynamics of the events for an isolated oscillator, but also captures quantitatively the synchronization behavior of a large ensemble of such oscillators when their interaction is contingent on the events. Comparing the effective diffusion with the phase diffusion shows just how poorly the standard phase reduction performs: D_{eff}, which accounts for the fact that the oscillators interact only via the amplitude-dependent events, is orders of magnitude larger than D_{phase} (Figure 3.1b).

Strikingly, we find that the effective diffusion coefficient, exhibits "unruly" behavior for a wide range of oscillators: while in the limit of weak noise the effective diffusion coefficient D_{eff} converges to the diffusion coefficient D_{phase} obtained from the standard phase reduction, already for surprisingly small values of the noise it can become orders of magnitude larger than D_{phase}, only to *decrease* when noise is increased further to larger values (see Figure 3.1b for an example). For coupled such oscillators, this implies strongly enhanced sensitivity to desynchronization for intermediate noise values and reduced sensitivity for strong noise (see Chapter 2).

In the present chapter, we address the source and abundance of the unruliness in the TVGR and thus the effective diffusion coefficient. In contrast with our prior, purely computational approach, we obtain explicit expressions for the TVGR by considering the point process that arises from a linearized Poincare map. The paradigm of point processes has already previously been used in the analysis of oscillatory and excitable systems. In neuroscience, for example, the times T_n at which action potentials occur in a neuron are often considered to arise from a point process. Indeed, the distribution of so-called inter-spike-intervals, $T_{n+1} - T_n$, is a subject of great interest (see e.g. [**26**]). More broadly, point-process theory offers various measures of the temporal variability in oscillatory and excitable dynamics (Table 3.1). The Fano factor, for instance, has been used to measure the similarity of a given point process to a Poisson process, which has a Fano factor of 1. For noise-driven, nonlinear oscillatory and excitable systems, the Fano factor and effective diffusion coefficient have each been used to identify a variety of "resonances" that appear as a function of input noise strength: coherence resonance and incoherence maximization [**58, 48**].

The analysis of point processes is often limited to renewal processes, i.e. to point processes in which the intervals between events are independent and identically distributed. Noise-driven limit-cycle oscillators will, however, in general maintain correlations from one cycle to the next and therefore do not fit the framework of renewal processes. Indeed, non-renewal dynamics are seen in real world oscillatory systems [**24, 26, 2**]. These correlations do not preclude a description of the oscillators within the framework of point processes. In this work, we therefore include correlations in the time intervals between

events by considering the Markov renewal process associated with the stochastic Poincare map.

3.1.3. Organization of the Chapter

The chapter is organized as follows. In Section 3.2 we introduce the effective phase via the events and their variance statistics and outline in more detail our approach for its analysis. In the somewhat technical Section 3.3 and the corresponding Appendxi B.1, we review the theory for variance statistics of Markov and Markov renewal reward processes, offering a novel formula for the TVGR. We use that formula in Section 3.4 to demonstrate in a simple toy model how unruliness can arise in $\mathcal{V}_E^{(t)}$, offering an simple explanation of the eventual decrease of the TVGR with increasing noise strength. We then apply the general theory to the linearization of Poincare map dynamics for limit cycle oscillators in the Section 3.5 and the corresponding Appendix B.3. In Section 3.6 we narrow our focus to two-dimensional oscillators and make an argument that unruliness in $\mathcal{V}_E^{(t)}$ is a common occurrence. We conclude with plausible extensions to this work in Section 3.7 and a broader discussion in Section 3.8.

3.2. Phase Diffusion and The Temporal Variance Growth Rate

3.2.1. Phase Reduction and Its Breakdown

The specific subjects of this chapter are noise-driven oscillatory systems and the statistical characteristics of the events they produce. Motivated by systems in which output is only generated by events, we will take the event timings to be the relevant read-out of the

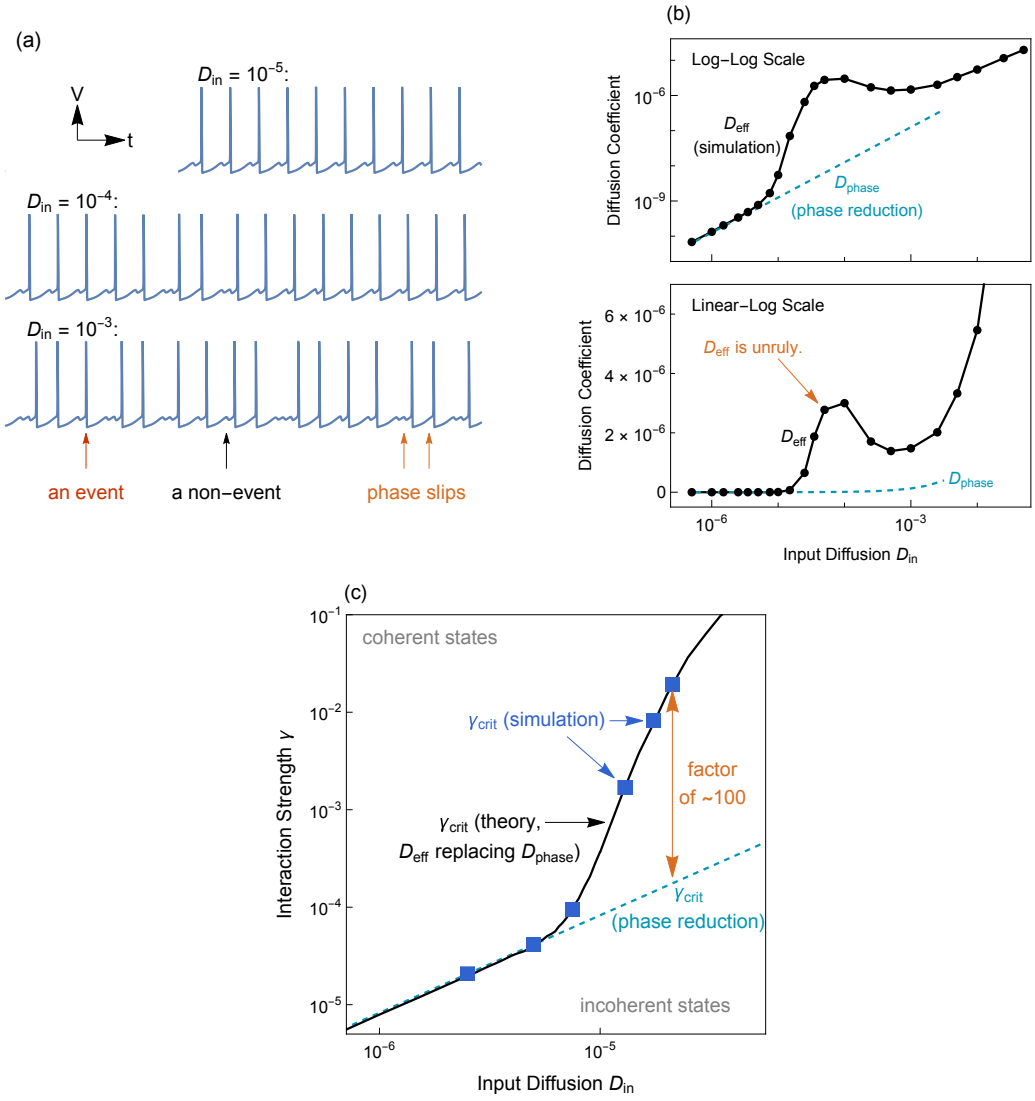

Figure 3.1. **Effective Diffusion in a Mixed-Mode Oscillator.** (See also Figures 2.9 and 2.10.) (a): Sample voltage traces of the neural MMO driven by noise. Only the large-amplitude excursions are events in which the oscillator produces "output"; the small-amplitude excursions are non-events. Phase slips occur for non-zero noise D_{in}. (b): The phase diffusion D_{phase} from phase reduction theory and the "unruly"effective diffusion D_{eff} (same data shown on two different scales). (c): A phase diagram for population states in globally coupled oscillators. Standard phase reduction with the Fokker-Planck theory, (1.2), predicts that the boundary between stable coherent states and stable incoherence is linear. The theory only agrees with the full, coupled-oscillator simulation once D_{phase} is replaced with D_{eff}.

system's dynamics, and we will be particularly interested in quantifying the long-time-scale, noise-driven dispersion of events. In this section we connect the statistics of the events with the diffusive dynamics of the phase and introduce an effective phase, which becomes relevant when the noise induces "phase slips" during which events are skipped or extra ones are produced (as in, for example, Figure 3.1a).

We investigate systems in $d+1$ dimensions of the form

$$d\vec{y} = \vec{F}(\vec{y})\,dt + \sqrt{2D_{in}}\mathbf{G}\left(\vec{y}\right)\,d\vec{W} \tag{3.1}$$

that have a stable limit cycle when $D_{in} = 0$, where $\vec{W}$ is the standard Wiener process of dimension $d+1$ and the matrix $\mathbf{G}(\vec{y})$ characterizes the noise correlations. To define events we choose a set E in phase space, such that an event occurs when $\vec{y}\,(t)$ intersects E. For instance, in neuroscience applications the event often corresponds to an action potential, where the voltage variable has a large "spike".

For weakly perturbed limit cycles, the system description can be substantially simplified by a phase reduction, in which the state $\vec{y} \in \mathbb{R}^{d+1}$ is approximately represented by a single phase variable ϕ parameterizing the limit cycle $\vec{y}_{LC}\left(\phi\right)$. The dynamics of (3.1) are thus captured via a single equation [**41, 1**],

$$d\phi = f_0\,dt + \sqrt{2D_{in}}\,Z^T\left(\phi\right)\mathbf{G}\left(\vec{y}_{LC}\left(\phi\right)\right)\,d\vec{W} + \mathcal{O}\left(D_{in}\right),$$

in which $\phi \in [0,1)$; upon reaching 1, ϕ is reset to 0. Subsequently applying the method of averaging (see e.g. Section 6.2 of [**71**]) yields

$$d\phi \sim f_0\,dt + \sqrt{2D_{phase}}\,dW, \tag{3.2}$$

where

$$D_{phase} = \left[f_0 \int_0^{1/f_0} Z^T(\phi)\, \mathbf{G}(\vec{y}_{LC}(\phi))\, \mathbf{G}^T(\vec{y}_{LC}(\phi))\, Z(\phi)\, d\phi \right] D_{in}. \tag{3.3}$$

The reduction thus offers a prediction for the amount of diffusion D_{phase} in the phase variable. In particular, from (3.3), it is proportional to the input diffusion coefficient. The dynamics of the reduced phase model are localized to the limit cycle, and so the events produced by (3.2) are governed by the intersection of the limit cycle with E. In this work, unless otherwise indicated, we will consider the simplest scenario, in which the noiseless limit cycle intersects E once and does so transversely. Without loss of generality, we take the phase on the limit cycle at the intersection to be $\phi = 0$.

We can compare the full dynamics (3.1) with that of the phase oscillator (3.2) via the corresponding point processes for the events. Define the point process T_n for the full system (3.1) as the time t at which the n^{th} event occurs. In contrast with the point process for the phase oscillator, given by the times at which $\phi = 0$, the process T_n will in general be non-renewal. Indeed, a noise-driven limit-cycle oscillator will support correlations in its dynamics from one intersection of E to the next. As the noise strength is increased from infinitesimal values, $\vec{y}$ will deviate from $\vec{y}_{LC}$ and the time interval, $\Delta T_n = T_{n+1} - T_n$, between events will deviate from the period of the unperturbed limit cycle. More drastic changes occur with increasing noise strength if the event surface E extends only a finite distance away from the limit cycle. In that case non-infinitesimal noise may cause the trajectory to miss E in one cycle and return to it only in a subsequent cycle, inserting an additional non-event crossing. The phase ϕ will then not capture the event point process and - in terms of describing the events - the phase equation (3.2) breaks down.

3.2.2. Effective Phase Diffusion for Events

In order to obtain a description that is able to deal with phase slips we introduce an alternative, *effective* phase oscillator, one whose *event* statistics by definition match that of the full oscillator in the *long-time limit*. It will be useful to characterize the point process by the function N_t, the number of events that have occurred by time t. Take

$$d\phi_{\mathit{eff}} = f_{\mathit{eff}}\, dt + \sqrt{2D_{\mathit{eff}}}\, dW, \tag{3.4}$$

where each reset from $\phi_{\mathit{eff}} = 1$ to 0 constitutes an event and the effective oscillation frequency,

$$f_{\mathit{eff}} \equiv \frac{1}{\mu} = \lim_{t\to\infty} \frac{1}{t} \mathrm{E}\left\{N_t\right\}, \tag{3.5}$$

and the effective phase diffusion coefficient,

$$D_{\mathit{eff}} \equiv \frac{1}{2}\mathcal{V}_E^{(t)} = \frac{1}{2} \lim_{t\to\infty} \frac{1}{t} \mathrm{var}\left\{N_t\right\}, \tag{3.6}$$

are defined by the event point process N_t of the full oscillator. We expect that in the weak-noise limit, where phase reduction applies, $f_{\mathit{eff}} \to f_0$, $D_{\mathit{eff}} \to D_{\mathit{phase}}$, and so $\phi_{\mathit{eff}} \to \phi$. In general, and in contrast with ϕ, however, we do not think of a given value of ϕ_{eff} as referring to a particular point or set of points in phase space. Instead, ϕ_{eff} is analogous to an averaged dynamical variable whose dynamics only reflect those of the original system over long time scales.

Note that because each passage from $\phi_{\mathit{eff}} = 0$ to $\phi_{\mathit{eff}} = 1$ governed by (3.4) is independent from all others, the corresponding point process N_t^{eff} (with events defined by

	definition	... in the long-time limit	... in the effective phase oscillator or effective renewal process
Event Rate	$\frac{1}{t}\mathrm{E}\{N_t\}$	$1/\mu$	$f_{\mathit{eff}} = \frac{1}{\mathrm{E}\{\Delta\tau\}}$
Mean Inter-event Interval	$\frac{1}{n}\mathrm{E}\{T_n\}$	μ	mean first passage time $= \mathrm{E}\{\Delta\tau\}$
Temporal Variance Growth Rate	$\frac{1}{t}\mathrm{var}\{N_t\}$	$\mathcal{V}_E^{(t)}$	$2D_{\mathit{eff}} = \frac{\mathrm{var}\{\Delta\tau\}}{\mathrm{E}\{\Delta\tau\}^3}$
Event Dispersion Rate	$\frac{1}{n}\mathrm{var}\{T_n\}$	$\mu^3\mathcal{V}_E^{(t)}$	variance in the first passage time $= \mathrm{var}\{\Delta\tau\}$
Fano Factor (or Index of Dispersion)	$\frac{\mathrm{var}\{N_t\}}{\mathrm{E}\{N_t\}}$	$\mu\mathcal{V}_E^{(t)}$	$\frac{\mathrm{var}\{\Delta\tau\}}{\mathrm{E}\{\Delta\tau\}^2}$

Table 3.1. Typical point process statistics, their inter-relationships in the long-time limit ($n \to \infty$ or $t \to \infty$), and their connection to the effective phase oscillator (3.4). Note that as long as these long-time-scale statistics exist for the general (non-renewal) point process, they can be reproduced by a renewal process: choose the random renewal intervals $\Delta\tau$ such that $\mu = \mathrm{E}\{\Delta\tau\}$ and $\mathcal{V}_E^{(t)} = \frac{\mathrm{var}\{\Delta\tau\}}{\mathrm{E}\{\Delta\tau\}^3}$. (3.4) is just one example of such a renewal process.

$\phi_{\mathit{eff}} = 1$) is a renewal process, and its statistics are fully determined by the first passage time distribution. Ignoring the reset, (3.4) produces Brownian motion with a drift, which has a first passage time density given by an inverse Gaussian distribution with mean μ and variance $\mu^3\mathcal{V}_E^{(t)}$ (see e.g. [**25, 1**]). We note that because (3.4) incorporates each a mean- and variance- type statistic, N_t^{eff} will correctly reproduce all of the long-time statistics of N_t that appear in Table 3.1, which are inter-related [**12**]. To facilitate the comparison of the standard phase reduction (3.2) with the effective phase oscillator (3.4), here we specifically focus on the effective diffusion $D_{\mathit{eff}} = \frac{1}{2}\mathcal{V}_E^{(t)}$. In this context, a "break-down" in the phase reduction theory occurs when D_{eff} deviates greatly from the linear prediction, D_{phase}.

The break-down can be quite dramatic. Indeed, for the neuronal mixed-mode oscillators (see Chapter 2), as a function of the input noise strength D_{in}, the effective diffusion coefficient D_{eff} grows orders of magnitude above the linear phase diffusion before eventually decreasing to comparable levels for large input noise (Figure 3.1b). We call this behavior "unruly" because its graph is highly nonlinear and non-monotonic in stark contrast with the linear prediction from phase reduction, (3.3). Similarly dramatic effects have been noted in the literature: non-monotonicity in the Fano factor has been called incoherence maximization [**48**] when a maximum appears and coherence resonance [**58**] when there is a minimum.

3.2.3. Theoretical Approach

The main contribution of this work is an asymptotic expression for the effective diffusion coefficient (3.6) that applies to a wide class of limit-cycle oscillators and that is readily analyzed and interpreted. In deriving that formula, we make the following over-arching choices (see also Figure 3.2):

(1) We define an event as occurring when the trajectory $\vec{y}$ of the system (3.1) crosses a subset E of a codimension-1 surface S that is transverse to the limit cycle.

(2) Of the variance-like statistics listed in Table 3.1, we compute in particular the asymptotic temporal variance growth rate, $\mathcal{V}_E^{(t)} = 2D_{\mathit{eff}}$.

(3) We approximate the dynamics in the vicinity of the limit cycle to linear order, but - in contrast with the phase reduction (3.2) - we explicitly include dynamics off of the limit cycle.

Choice 1 defines a point process, which arises from crossings of a Poincare section S. It has the structure of a Markov renewal process, where the "Markov" aspect governs the positions on the Poincare section and the "renewal" aspect reflects the randomly varying times between crossings. As compared with the renewal process generated by (3.2), including the state space of the Poincare section enriches the dynamics. The dynamics can be non-renewal, since the position on the section and the time intervals between crossings will in general be correlated. We make the further choice to analyze the system by including all crossings of S as steps in the process but distinguishing those in E as events. There are therefore two related point processes that we consider, given by T_n^S (or N_t^S) for the process on S and T_n^E (or N_t^E) for the one limited to E. While in principle it is also an option to simply limit the point process to E in the first place by taking $S = E$, our choice makes calculation feasible. Unless otherwise indicated, we restrict our consideration to sections that intersect the noiseless limit cycle once with the intersection occurring in E. Then, in the absence of noise, $T_n^S = T_n^E$, which is also the point process produced by the phase-reduced model, (3.2).

The inclusion of E informs our Choice 2, since the asymptotic temporal variance growth rate can be calculated as

$$\mathcal{V}_E^{(t)} = \lim_{t\to\infty} \frac{1}{t} \mathrm{var}\left\{N_t^E\right\} = \lim_{t\to\infty} \frac{1}{t} \mathrm{var}\left\{\sum_{k=1}^{N_t^S} 1_E\left(x_k\right)\right\}, \tag{3.7}$$

where 1_E is the indicator function for E on S and $x_k \in S$ is the position of the k^{th} crossing on the section - the k^{th} step in the Markov process, which occurs at time T_k^S. We note that with this specification of the function 1_E on the state space, this process could be considered a Markov renewal reward process in which each event ($x_k \in E$) produces a

reward of 1 and each non-event crossing ($x_k \in S$, but $x_k \notin E$) a reward of 0. In our non-rigorous analysis, we will assume that the Markov processes are sufficiently "nice"; at the least, we require that the processes have an invariant probability density $\pi(x)$ which is nonzero for all $x \in S$ and has the necessary recurrence properties to render $\mathcal{V}_E^{(t)}$ independent of initial conditions.

Finally, Choice 3 means that, while our calculation of the effective diffusion coefficient $D_{eff} = \frac{1}{2}\mathcal{V}_E^{(t)}$ will be an approximation of (3.6), it will be a correction to D_{phase}. In contrast with the work presented in Chapter 2, where we found D_{eff} computationally, we here offer an analytical treatment that is made possible by the linearization. Our analysis combines the ideas presented by Hitczenko and Medvedev [**32**] and Wilson and Moehlis [**84**]. The former make use of an orthonormal moving coordinate frame to describe the Poincare map for the position x_k on the Poincare section, approximating the dynamics to linear order in x_k. We expect this approximation to be accurate when the Poincare map is roughly linear in and near the subset E, e.g. if E is sufficiently small. We perform this analysis using phase-isostable coordinates [**84**] under which the linearized dynamics are particularly simple.

We note that Wilson and Ermentrout [**83**] have previously introduced a notion of phase that also addresses features present in output signals, like neuronal voltage spikes. In light of the work we present here, they, too, consider output features that correspond to the crossing of a Poincare section S, and, under their "operational" definition of phase, the oscillator completes one cycle with every crossing of S. Using phase-isostable coordinates, they derive a representation of the transient response of the operational phase to forcing. But, since their analysis is perturbative, it is not clear that the operational phase can be

immediately generalized to account for amplitude-dependent events (and thus phase slips) defined via an event subsection E. The amplitude-dependent events, however, underlie the unruly noise response of oscillators that is of interest here. In our approach, in (3.7) in particular, we capture the distinguished crossings of the Poincare section within the Markov renewal framework via an additional element: the reward function $1_E(x)$.

We also note that Schwalger and Lindner [**69**] have previously considered Markov renewal processes on finite state spaces where, likewise, only some transitions are distinguished as events, and they provide a broad analysis of such models. As an application, they consider the negative correlations between subsequent inter-spike-intervals that are seen in some neurons, and they offer ad-hoc, minimal models with a small number of discrete states that qualitatively reproduce the correlation. In this work, we have a similar goal: capturing a statistical property of a non-renewal point process, here the unruliness in the temporal variance growth rate for events, in a reduced model. Rather than using an ad-hoc model, however, we make use of a systematic reduction that connects the full system (3.1) directly to a linear model on the continuous-state-space Poincare section.

3.3. Asymptotic Variance Growth Rates for Markov and Markov Renewal Processes

Our immediate goal is an understanding of the (asymptotic)[2] temporal variance growth rate (3.7). In this section and the corresponding Appendix B.1, we derive an expression for it, (3.15), that is amiable to the analysis that appears in the following sections. We find (3.15) in two steps: we first write the temporal variance growth rate in terms of a

[2]"Asymptotic" here reflects the fact that we are taking the limit $t \to \infty$ or $n \to \infty$. Since, throughout this chapter, we are only considering long-time statistics - i.e. this type of limit, we will drop this modifier in future references to this and related quantities.

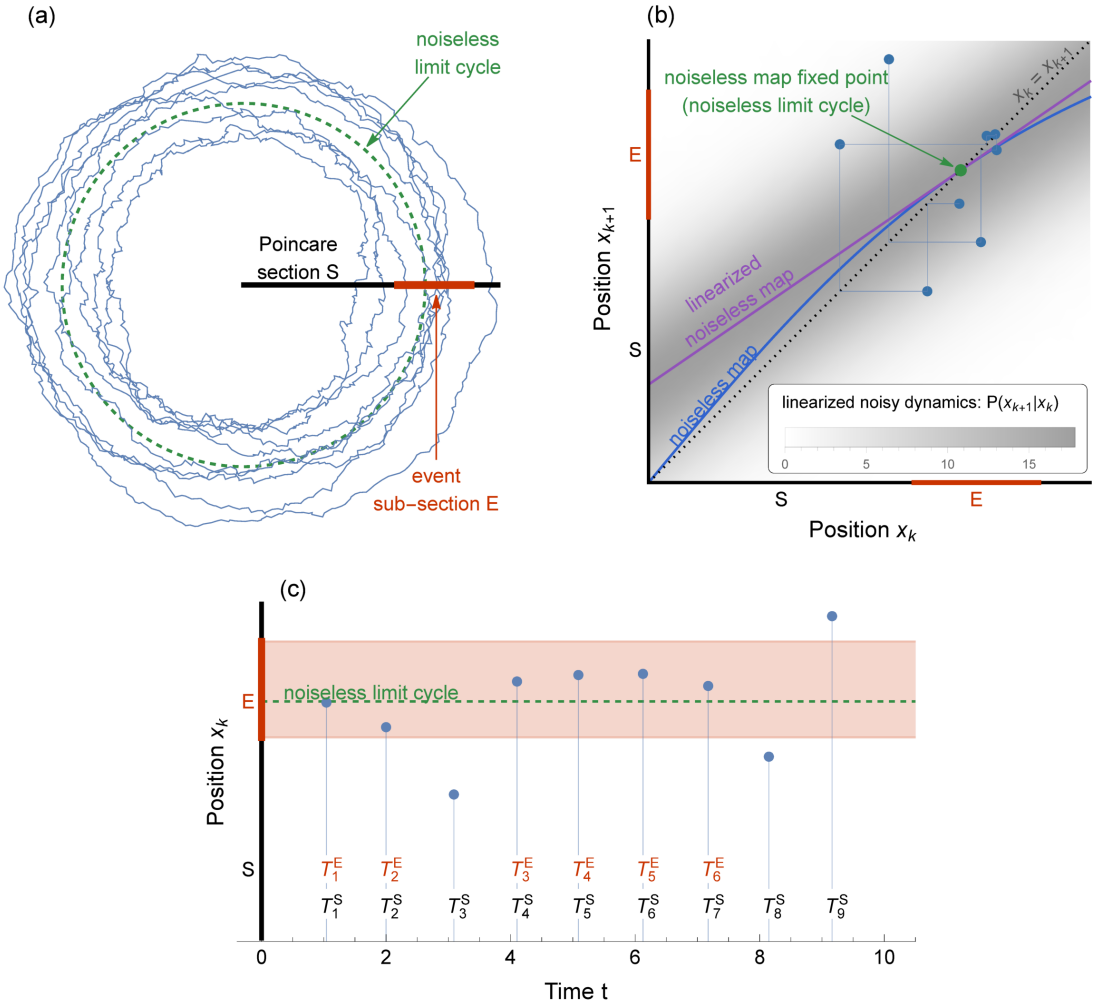

Figure 3.2. **Events in a Noise-driven Limit-cycle Oscillator.** (a): Noisy trajectory crossing a Poincare section S and an event sub-section E. (b): The linearized, stochastic, discrete-time dynamics for x_k on the Poincare section S. The shading gives the probability of x_{k+1} conditioned on x_k. (c): The designation $1_E(x_k)$ of each x_k as an event or non-event and the two associated point processes, T_k^S and T_k^E.

sequential variance growth rate (Section 3.3.2 and Appendices B.1.1 and B.1.2), and then we decompose the resulting formula into components that will show qualitatively different behaviors when applied to the Markov renewal processes arising from oscillatory systems (Section 3.3.3).

3.3.1. Sequential Variance Growth Rates

The TVGR $\mathcal{V}_{\mathrm{E}}^{(t)}$, (3.7), is a statistical property of the Markov renewal process T_k^S on the Poincare section S. Recall that a Markov renewal process is a Markov process in which each step k occurs at a particular time T_k. The time interval $\Delta T_k \equiv T_{k+1} - T_k$ between steps is a random variable that may depend on the preceding and subsequent states, x_k and x_{k+1}, and the ΔT_k are independent of each other once conditioned on the states x_k and x_{k+1} (see e.g. Chapter 6 of [**34**] for a pedagogical introduction). Since $\mathcal{V}_{\mathrm{E}}^{(t)}$ is defined as a growth rate over time, it incorporates that temporal information. By contrast, we define the sequential variance growth rate $\mathcal{V}_{x\mapsto f(x)}^{(n)}$ (SVGR) of a function f on a Markov process with states $x \in S$ as

$$\mathcal{V}_{x\mapsto f(x)}^{(n)} \equiv \mathcal{V}^{(n)}\left[x \mapsto f\left(x\right)\right] \equiv \lim_{n\to\infty} \frac{1}{n} \mathrm{var}\left\{\sum_{k=1}^{n} f\left(x_k\right)\right\}. \tag{3.8}$$

In our notation, the subscript $x \mapsto f\left(x\right)$ of $\mathcal{V}$ identifies how one maps the Markov states x to the value $f\left(x\right)$ that is then accumulated from one step to the next. In some instances, for readability, we will alternatively write the expression $x \mapsto f\left(x\right)$ in square brackets. As in the TVGR, the variance in the SVGR is taken over sample sequences given by x_k. However, unlike the TVGR, the SVGR $\mathcal{V}_{x\mapsto f(x)}^{(n)}$ does not take into account any information about the time between steps. In the case that f is an indicator function for those x that correspond to events, the function variance is a simplified version of the TVGR wherein the randomness of the time intervals in the Markov renewal process is ignored and all intervals are set to 1.

Note that, like x_k, $f_k = f(x_k)$ is itself a sequence of correlated random variables. When the variance of the normalized sum $\mathcal{V}^{(n)}_{x\mapsto f(x)}$ exists, it follows the standard formula for series of random variables,

$$\mathcal{V}^{(n)}_{x\mapsto f(x)} = \mathrm{var}\left\{f(x_0)\right\} + \sum_{k=1}^{\infty} \mathrm{cov}\left\{f(x_0), f(x_k)\right\} + \sum_{k=1}^{\infty} \mathrm{cov}\left\{f(x_k), f(x_0)\right\}. \tag{3.9}$$

Here, we imagine choosing the initial state x_0 randomly according to the invariant density of the Markov process. In that case, by the stationarity of the Markov process, x_0 is as good as any arbitrary element of the sequence. We write the two sums in (3.9) separately to allow f to be vector-valued, in which case $\mathrm{cov}(\cdot,\cdot)$ is in general a non-symmetric cross-covariance matrix. Note that the definition of the SVGR (3.8) and the formula (3.9) also allow for situations where the function $f(x)$ itself is not deterministic, but takes on random values that may depend on x. This is a feature we will make use of in the following formulae.

We will also make use of a generalization of the SVGR, allowing for functions f that depend on adjacent points x_k and x_{k+1} of the Markov process:

$$\mathcal{V}^{(n)}_{(x\to x')\mapsto f(x,x')} \equiv \lim_{n\to\infty} \frac{1}{n} \mathrm{var}\left\{\sum_{k=1}^{n} f(x_k, x_{k+1})\right\}. \tag{3.10}$$

For that we find it useful to define the Markov process on the "edges" of the original process. We notate the sequence of states of the new process as $z_k \equiv (x_k \to x_{k+1})$. The process's dynamics are wholly induced by that of the original process: the probability to transition from $(a \to b)$ to $(c \to d)$ under the new process is the probability to transition from b to d under the original process if $b = c$ and is 0 otherwise. The generalized SVGR

on the original process is then just a standard SVGR on the new process, $\mathcal{V}^{(n)}_{(x\to x')\mapsto f(x,x')} = \mathcal{V}^{(n)}_{z\mapsto f(z)}$, and so it can be computed via (3.9).

3.3.2. The Event Temporal Variance Growth Rate as a Sequential Variance Growth Rate

In Appendix B.1, we use a result from reward-renewal theory [**73**] to make the non-rigorous argument that the TVGR for events $\mathcal{V}^{(t)}_E$ can be written in terms of the generalized SVGR of an (ordinary, non-renewal) Markov process. We find

$$\mathcal{V}^{(t)}_E = \frac{1}{\mathrm{E}\{\Delta T\}}\mathcal{V}^{(n)}\left[(x\to x')\mapsto\left(1_E(x) - \mathrm{E}\{1_E(x)\}\frac{\Delta t(x,x')}{\mathrm{E}\{\Delta T\}}\right)\right], \tag{3.11}$$

where $\Delta t(x,x')$ is a random-valued function with the property that $\Delta t(x_k, x_{k+1})$ has the same distribution as ΔT_k when conditioned on x_k and x_{k+1}. Below and in the following sections, we will rescale time so that the expected time interval is 1, and, because we will use it often, introduce $\mathcal{E}$ as the event probability; thus

$$\mathrm{E}\{\Delta T\} = 1 \tag{3.12}$$

$$\mathrm{E}\{1_E(x)\} = \mathcal{E}. \tag{3.13}$$

With these simplifying choices, (3.11) becomes

$$\mathcal{V}^{(t)}_E = \mathcal{V}^{(n)}_{(x\to x')\mapsto(1_E(x)-\mathcal{E}\Delta t(x,x'))}, \tag{3.14}$$

which can in principle be computed via (3.9).

Before moving on, we make the following disclaimer. In Appendix B.1.1, we only give a plausibility argument that the TVGR can by computed via (3.14). We offer a more detailed derivation in Appendix B.1.2 resulting in the same formula, but that derivation is only valid for Markov chains with finitely-many states. Even so, (3.14) has a clear interpretation for more general state spaces. We assume that the result holds for the Markov renewal processes we consider in this work, which, as we discuss in Section 3.5.2, are innocuous Gaussian processes. We provide some empirical validation of (3.14) in Section 3.6.6.

3.3.3. Decomposition of the Temporal Variance Growth Rate

The TVGR (3.14) is a (type of) variance of the sum of the correlated quantities $1_E(x')$ and $-\mathcal{E}\Delta t(x,x')$, and it would be reasonable to guess that it can be decomposed as the variance of each term plus a type of covariance $\mathcal{CV}$,

$$\mathcal{V}_E^{(t)} = \mathcal{V}^{(n)}_{x\mapsto 1_E(x)} - 2\,\mathcal{E}\,\mathcal{CV}^{(n)}_{(x\to x')\mapsto 1_E(x),(x\to x')\mapsto \Delta t(x,x')} + \mathcal{E}^2\,\mathcal{V}^{(n)}_{(x\to x')\mapsto \Delta t(x,x')}. \tag{3.15}$$

This is indeed the case, as we show in Appendix B.1.3, where we expand (3.14) via (3.9) and reorganize the resulting terms. $\mathcal{CV}^{(n)}_{x\mapsto f(x),x\mapsto g(x)}$ is defined via a natural extension of (3.8) and (3.9),

(3.16)

$$\mathcal{CV}^{(n)}_{x\mapsto f(x),x\mapsto g(x)} \equiv \lim_{n\to\infty}\frac{1}{n}\mathrm{cov}\left\{\sum_{k=0}^{n-1} f(x_k), \sum_{k=0}^{n-1} g(x_k)\right\}$$

$$= \mathrm{cov}\left\{f(x_0), g(x_0)\right\} + \sum_{k=1}^{\infty}\mathrm{cov}\left\{f(x_0), g(x_k)\right\} + \sum_{k=1}^{\infty}\mathrm{cov}\left\{f(x_k), g(x_0)\right\}. \tag{3.17}$$

Note that if f and g are vector-valued, $\mathcal{CV}^{(n)}_{x\mapsto f(x),x\mapsto g(x)}$ is analogous to a cross-covariance matrix and will in general be non-symmetric.

The first term of (3.15) is just the sequential variance that arises by ignoring the temporal aspect of the process, i.e. by setting $\Delta T = 1$. We call it the "Markov-only" term. Similarly, if one ignores the partitioning of the state-space by E by, say, setting $E = S$, there is no variance in $1_E(x)$. In that case, only the last term of (3.15), $\mathcal{V}^{(n)}_{(x\to x')\mapsto\Delta t(x,x')}$, is non-zero. We call this the "temporal" term. We call the remaining covariance term "mixed", since it only appears when both the temporal and partitioning aspects of the process are present and are inter-dependent.

3.4. Temporal Variance Growth Rate for a Toy Model

Before considering the TVGR for stochastic limit-cycle oscillators, we introduce a simple, discrete 2-state toy model that reveals the quality of each the Markov, the mixed, and the temporal components in (3.15) and how they interact to generate the non-monotonic behavior of $\mathcal{V}^{(t)}_E$. We take $x \in \{0,1\} = S$ and $E = \{1\}$. In order to fully specify the Markov-renewal process, we must prescribe the joint, conditional distribution for x_{k+1} and $\Delta T^S_k = T^S_{k+1} - T^S_k$ given x_k. For simplicity, we assume that ΔT^S_k is independent of x_k and that x_{k+1} is completely determined by ΔT^S_k. Equivalently, there is a subset $E_{\Delta T}$ of the positive real line, so that x_{k+1} corresponds to an event if and only if $\Delta T^S_k \in E_{\Delta T}$. Thus,

$$x_{k+1} = 1_{E_{\Delta T}}\left(\Delta T^S_k\right). \tag{3.18}$$

This very roughly captures what we expect to see in stochastic oscillators: the time to reach E will generally be distributed differently from the time to reach $S \setminus E$. Under these assumptions, x_j and x_k are independent when $j \neq k$. Likewise, the pairs ΔT_j^S and x_{k+1} as well as ΔT_j^S and ΔT_k^S are independent when $j \neq k$. In this simple scenario, only one term survives in each of the components of the TVGR, corresponding to the first terms in (3.9) and (3.17). Each of the Markov-only, mixed and temporal components of $\mathcal{V}_E^{(t)}$ can be evaluated easily, since x_k is a sequence of independent Bernoulli-distributed random variables governed by the probabilities $\mathcal{E}$ and $1 - \mathcal{E}$. We have via (3.9) and (3.17)

$$\begin{aligned}
\mathcal{V}^{(n)}_{x \mapsto 1_E(x)} &= \operatorname{var}\left\{1_E\left(x_0\right)\right\} \\
&= \mathcal{E}\left(1-\mathcal{E}\right) \\
\mathcal{CV}^{(n)}_{(x \to x') \mapsto 1_E(x), (x \to x') \mapsto \Delta t(x,x')} &= \operatorname{cov}\left\{1_E\left(x_1\right), \Delta t\left(x_0, x_1\right)\right\} \\
&= \mathcal{E}\left(1-\mathcal{E}\right)\left(\overline{\Delta t}_{(1)} - \overline{\Delta t}_{(0)}\right) \\
\mathcal{V}^{(n)}_{(x \to x') \mapsto \Delta t(x,x')} &= \operatorname{var}\left\{\Delta t\left(x_0, x_1\right)\right\} \\
&= \operatorname{var}\left\{\Delta T_0^S\right\},
\end{aligned}$$

where $\overline{\Delta t}_{(n)} = \mathrm{E}\left\{\Delta T_k^S \left| x_{k+1} = n \right.\right\}$ for $n = 0, 1$.

Recall that we are interested in the TVGR as a function of an input noise strength D_{in}. In our toy model, we assert that the effect of increased input noise will be to increase the variance of the inter-step times: we take ΔT_k^S to be a random variable with mean 1 and variance $\sigma^2 \equiv \frac{D_{in}}{2}$. Following (3.15) the TVGR can then be written simply as

$$\mathcal{V}_E^{(t)} = \mathcal{E}\left(1-\mathcal{E}\right)\left(1 - 2\mathcal{E}\left(\overline{\Delta t}_{(1)} - \overline{\Delta t}_{(0)}\right)\right) + \mathcal{E}^2 \frac{D_{in}}{2}, \tag{3.19}$$

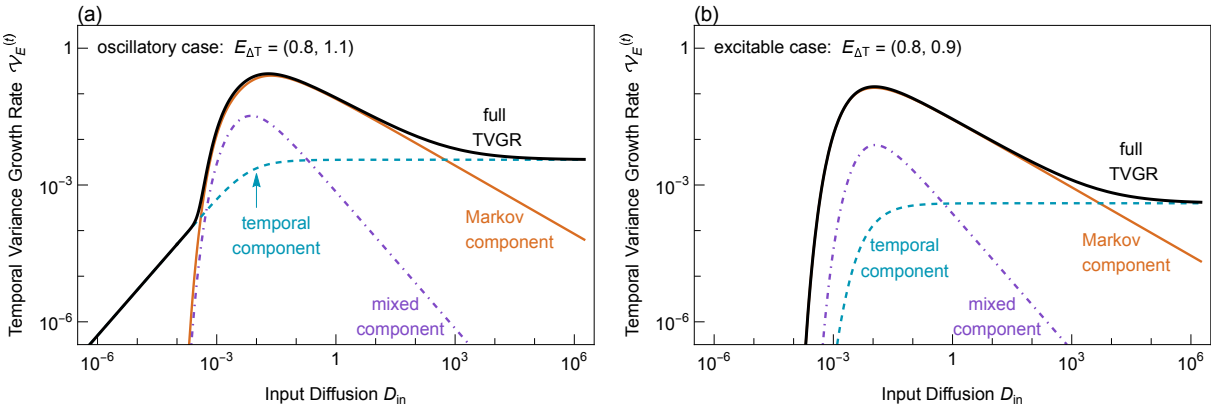

Figure 3.3. **Temporal variance growth rate $\mathcal{V}_E^{(t)}$ for the 2-state toy model (3.18).** The "Markov-only" and mixed components of the TVGR are non-monotonic and the temporal component is monotonic. In (a), the oscillatory case, their sum, the full TVGR, has the characteristic unruly quality: an initial linear growth, a strong nonlinear amplification, a maximum, and a subsequent decrease. The choice of an asymmetric interval $E_{\Delta T}$ about $\Delta T^S = 1$ is only made so that $\overline{\Delta t}_1 \neq \overline{\Delta t}_0$ and the mixed term is nonzero. It is not essential to the appearance of unruliness. In (b), the excitable case, the TVGR is missing the initial linear growth.

where $\mathcal{E}$, $\overline{\Delta t}_{(1)}$, and $\overline{\Delta t}_{(0)}$ each vary with D_{in}. When $E_{\Delta T}$ is a contiguous interval and ΔT_n^S is unimodally distributed, we expect that each will in fact vary monotonically with D_{in}. Then, because of the appearance of the Bernoulli variance $\mathcal{E}\,(1-\mathcal{E})$, we expect the Markov and mixed terms to be non-monotonic, but the temporal term to be monotonic. As a concrete example, we consider the case where ΔT_n is a Gaussian random variable and $E_{\Delta T}$ is an interval $(a,\, b)$ with $a < 1 < b$. Note that this choice of $E_{\Delta T}$ implies that events will be produced regularly when $D_{in} = 0$, since then $\Delta T_k^S = 1 \in E_{\Delta T}$ for all k. In this case, the temporal term is indeed monotonic, producing a roughly linear TVGR in the weak-input-noise regime (Figure 3.3a). And the Markov and mixed terms are non-monotonic and produce the strong and non-monotonic nonlinearity that is characteristic of what we call "unruliness" (compare Figures 3.1b and 3.3a).

Although we do not pursue it further in this work, we note that the toy model can also be placed in an "excitable" rather than an oscillatory regime by choosing $E_{\Delta T} = (a, b)$ with either $a < b < 1$ or $1 < a < b$. Here event-production requires a nonzero amount of noise, $D_{in} > 0$. In contrast with the oscillatory case, no initial linear growth appears in the TVGR, and the Markov-only component dominates at low noise (Figure 3.3b).

The pattern of behavior of the toy model in the oscillatory regime is mirrored in the TVGR for planar limit cycle oscillators, as we show in Section 3.6. There, as is the case here, the strong nonlinear rise in the TVGR roughly corresponds to $\mathcal{E}$ decreasing from 1 to $\frac{1}{2}$. In the oscillator, this reflects an increasing number of phase slips, i.e. oscillation cycles that miss the event subset E and replace events with non-events, or vice versa. The TVGR decreases again when $\mathcal{E}$ decreases past $\frac{1}{2}$, where the phase slips dominate the oscillations. Thus, the Bernoulli variance $\mathcal{E}(1 - \mathcal{E})$ underlies the seemingly counterintuitive increased regularity in the (lack of) events with increased input noise strength.

3.5. Temporal Variance Growth Rate for Limit Cycle Oscillators

We now apply the theory developed in Section 3.3 to the particular Markov renewal processes arising from Poincare map dynamics for oscillators. Here we first sketch the case of general oscillators, which is discussed in detail in Appendices B.2 and B.3. In Section 3.6 we then focus on the implications for unruliness in planar oscillators.

3.5.1. Noise-Driven Oscillators in Phase-Isostable Coordinates

We first write the limit-cycle oscillator in $d + 1$ dimensions, (3.1), in phase-isostable coordinates $(\phi, \psi_1, \psi_2, \ldots, \psi_d)$ (see [**84**] and Appendix B.2.1 for an extended discussion).

For weak noise and assuming small deviations ψ_i from the limit cycle, one has

$$d\phi = dt + \sqrt{2D_{in}}\vec{Z}(\phi)^T \mathbf{G}_{LC}(\phi)\, d\vec{W} + h.o.t. \tag{3.20}$$

$$d\psi_i = -\kappa_i \psi_i\, dt + \sqrt{2D_{in}}\vec{Y}_i(\phi)^T \mathbf{G}_{LC}(\phi)\, d\vec{W} + h.o.t. \quad i = 1 \ldots d\,. \tag{3.21}$$

Here, $\vec{Z}(\phi)$ is known as the phase response curve (PRC) and the $\vec{Y}_i(\phi)$ as isostable response curves (IRCs). They are the gradients of the phase and isostables coordinates, respectively, as functions of the state $\vec{y}$ evaluated at the point on the limit cycle with phase ϕ. Here, in the limit of small D_{in}, they give the responses of the phase and isostable variables to the perturbation $\sigma\mathbf{G}(\vec{y})\, d\vec{W}(t)$ that appears in (3.1). At lowest order, $\mathbf{G}(\vec{y})$ is replaced by $\mathbf{G}_{LC}(\Phi(\vec{y}))$, where $\mathbf{G}_{LC}(\phi)$ is the value that $\mathbf{G}$ has at the point on the limit cycle with phase ϕ. The higher order terms (*h.o.t.*'s) in (3.21) and in the following are $\mathcal{O}\left(D_{in}, \sqrt{D_{in}}\left|\vec{\psi}\right|, \left|\vec{\psi}\right|^2\right)$, where $\vec{\psi} \equiv (\psi_1, \psi_2, \ldots, \psi_d)$.

We note that, in the context stochastically-forced oscillators, recent studies have introduced alternative, noise-strength-dependent versions of phase [**68, 77**] and isostable coordinates [**61**]. Here in (3.20) and (3.21), we however make use of the phase-isostable coordinate system defined with respect to the deterministic dynamics (see (B.11) and B.12).

3.5.2. Linearized Poincare Map and First Passage Time

Turning now to the Poincare section and map, we represent the Poincare section S near the limit cycle by

$$\phi = \phi_S\left(\vec{\psi}\right) = \vec{m}_S^T\vec{\psi} + \mathcal{O}\left(\left|\vec{\psi}\right|^2\right). \tag{3.22}$$

Note that this section intersects the limit cycle at $\phi = 0$, and therefore taking $\vec{m}_S = 0$ chooses the isochron given by $\phi = 0$ as the Poincare section. In the general case, $\vec{m}_S \neq 0$, the Poincare section does not correspond to an isochron, and so the time ΔT_n^S between consecutive section crossings differs from the period of the periodic orbit by an amount that reflects the change in $\vec{\psi}$ between the crossings. One has to leading order $\mathrm{E}\left\{\vec{\psi}_{n+1} \middle| \vec{\psi}_n\right\} \sim \mathbf{\Lambda}\vec{\psi}_n$ and $\mathrm{E}\left\{\Delta T_n^S \middle| \vec{\psi}_n\right\} \sim 1 + \vec{\delta\mu}^T \vec{\psi}_n$, where Λ is the diagonal matrix with entries $e^{-\kappa_i}$ and $\vec{\delta\mu} \equiv m_S^T (\mathbf{I} - \mathbf{\Lambda})$. Taking into account the contribution of the noisy perturbations at first order, we expect

$$\begin{aligned} \Delta T_n^S &= 1 + \vec{\delta\mu}^T \vec{\psi}_n + \sqrt{2D_{in}}\zeta_n + h.o.t. \\ \vec{\psi}_{n+1} &= \mathbf{\Lambda}\vec{\psi}_n + \sqrt{2D_{in}}\vec{\eta}_n + h.o.t., \end{aligned} \tag{3.23}$$

where ζ_n and $\vec{\eta}_n$ are normally distributed random variables with mean 0, which, in general, will be correlated. Because the process is Markovian, ζ_n and $\vec{\eta}_n$ are i.i.d. across n. The (co)variances of ζ_n and $\vec{\eta}_n$ depend on $\vec{m}_s$, $\mathbf{\Lambda}$ and the PRC and IRCs. They are derived along with $\vec{\delta\mu}$ in Appendix B.3.1. Note that since we assume stability of the limit cycle, all Floquet multipliers that appear in the diagonal matrix $\mathbf{\Lambda}$ have magnitude less than 1. Then in the limit $n \to \infty$, where the process becomes stationary, we expect $\vec{\psi}_n$ to be distributed as

$$\vec{\psi}_n \overset{dist.}{\sim} \mathcal{N}\left(0, 2D_{in}\mathbf{\Gamma}_{ss}\right), \tag{3.24}$$

where $\mathcal{N}\left(\vec{\mu}, \mathbf{\Sigma}\right)$ is the multivariate normal distribution with mean $\vec{\mu}$ and covariance matrix $\mathbf{\Sigma}$. The "steady-state" matrix $\mathbf{\Gamma}_{ss}$ is found via (3.23) by the condition that the covariance

matrices for $\vec{\psi}_n$ and $\vec{\psi}_{n+1}$ are the same, namely $2D_{in}\mathbf{\Gamma}_{ss}$. Thus,

$$\text{var}\left\{\vec{\psi}_{n+1}\right\} = \text{var}\left\{\mathbf{\Lambda}\vec{\psi}_n\right\} + 2D_{in}\mathbf{\Gamma}$$

implies

$$2D_{in}\mathbf{\Gamma}_{ss} = 2D_{in}\left(\mathbf{\Lambda}\mathbf{\Gamma}_{ss}\mathbf{\Lambda} + \mathbf{\Gamma}\right)$$

where Γ is the covariance matrix of $\vec{\eta}_n$. Equivalently,

$$\mathbf{\Gamma}_{ss} = \sum_{k=0}^{\infty} \mathbf{\Lambda}^k \mathbf{\Gamma} \mathbf{\Lambda}^k. \tag{3.25}$$

Further, based on (3.23) and using elementary relationships between conditional and joint normal distributions, we show in Appendix B.3.2 that the joint, stationary distribution for $\vec{\psi}$ across multiple iterations of the map is given by

$$\begin{pmatrix} \vec{\psi}_{n+m} \\ \vec{\psi}_n \end{pmatrix} \overset{dist.}{\sim} \mathcal{N}\left(0, 2D_{in}\mathbf{C}_m\right);\ \mathbf{C}_m \equiv \begin{pmatrix} \mathbf{\Gamma}_{ss} & \Lambda^m\mathbf{\Gamma}_{ss} \\ \mathbf{\Gamma}_{ss}\Lambda^m & \mathbf{\Gamma}_{ss} \end{pmatrix}, \tag{3.26}$$

again in the limit $n \to \infty$.

3.5.3. Temporal Variance Growth Rate for Linearized Dynamics

We can compute the temporal variance growth rate using the distributions and statistic cited above and derived in Appendix B.3.2. We use (3.15), where the Markov state x is in this case the vector $\vec{\psi}$. The above linearization of the dynamics is essential for the computation; we make progress by leveraging the fact that all of the probability distributions are Gaussian. While we leave the full details of the derivation to SI Section

B.3.3, here we comment on results for each component of the TVGR, highlighting the dependence of each on D_{in}. In the SI and in the following, we take $N(\cdot; 0, \mathbf{\Sigma})$ to be the probability density function for the multivariate normal distribution with mean $\vec{0}$ and covariance matrix $\mathbf{\Sigma}$. Then, for example, (3.24) gives

$$\mathcal{E} = \int_E N(x; 0, 2D_{in}\mathbf{\Gamma}_{ss})\, dx. \tag{3.27}$$

From (3.9), we see that the Markov-only component $\mathcal{V}^{(n)}_{x\mapsto 1_E(x)}$ is made up of a series of integrals over $N(\cdot; 0, \sigma^2\mathbf{C}_m)$,

$$\mathcal{V}^{(n)}_{x\mapsto 1_E(x)} = \left(\mathcal{E} - \mathcal{E}^2\right) + 2\sum_{m=1}^{\infty}\left(\mathcal{E}_m - \mathcal{E}^2\right), \tag{3.28}$$

where

$$\mathcal{E}_m \equiv \int_{E\times E} N(z; 0, 2D_{in}\mathbf{C}_m)\, dz. \tag{3.29}$$

Note $\mathcal{E}_0 = \mathcal{E}$ and $\mathcal{E}_m \to \mathcal{E}^2$ as $m \to \infty$ (see the definition of $\mathbf{C}_m$ in (3.26)), so that in some sense $\mathcal{E}_m$ interpolates between $\mathcal{E}$ and $\mathcal{E}^2$ as m increases. Notoriously, closed form expressions for integrals like $\mathcal{E}_m$ are not known even if E is simple, e.g. a rectangular region. One must resort to numerical analysis or, as we do in Appendix B.4 , to bounding procedures to capture their behavior as a function of D_{in}.

The mixed component $\mathcal{CV}^{(n)}_{(x\to x')\mapsto 1_E(x'),(x\to x')\mapsto\Delta t(x,x')}$ involves both 1_E and $\Delta t\,(x,x')$. Under the linearized dynamics (3.23), $\Delta t\,(x,x')$ is linear in x and x' (see (B.52) in Appendix B.3.2). The mixed component, therefore, has at its essence the quantity

$$x_E \equiv \frac{1}{\mathcal{E}}\int_E xN\,(x;0,2D_{in}\mathbf{\Gamma}_{ss})\,dx, \tag{3.30}$$

which identifies the center of mass of the event subset E with respect to the steady-state density $N\,(\cdot;0,2D_{in}\Gamma_{ss})$. Indeed, the only dependence $\mathcal{CV}^{(n)}_{(x\to x')\mapsto 1_E(x'),(x\to x')\mapsto\Delta t(x,x')}$ has on D_{in} is via the product $x_E\mathcal{E}$ (see Appendix B.3.3). For many simple choices of E, x_E can be written in closed-form in terms of well-known functions, as we will do in Section 3.6. As a function of D_{in}, we expect the vector-valued x_E to increase in magnitude from the minimal distance between $x=0$ and E (when the density N is localized around $x=0$) and converge to the centroid of E as $D_{in}\to\infty$ (when the density conditioned on $x\in E$ approaches the uniform density).

The temporal factor, $\mathcal{V}^{(n)}_{(x\to x')\mapsto\Delta t(x,x')}$, does not involve the restriction of x to E; all integrals found here will be over the entire domain, and there will be no boundary-induced effects. Therefore this factor will only depend on D_{in} via an overall scaling by D_{in} (see Appendix B.3.3 for details).

In total,

$$\mathcal{V}^{(t)}_E = \left(\mathcal{E}-\mathcal{E}^2\right) + 2\sum_{m=1}^{\infty}\left(\mathcal{E}_m-\mathcal{E}^2\right) + \vec{b}^T x_E\mathcal{E}^2 + cD_{in}\mathcal{E}^2. \tag{3.31}$$

$\vec{b}^T$ and c and are found in Appendix B.3.3 to be

$$\vec{b}^T = 2\boldsymbol{\Gamma}_{ss}^{-1}\left(\mathbf{I}-\boldsymbol{\Lambda}\right)^{-1}\mathbf{Y}_G^T\vec{z}_G \equiv 2\boldsymbol{\Gamma}_{ss}^{-1}\left(\mathbf{I}-\boldsymbol{\Lambda}\right)^{-1}\int_0^1 \boldsymbol{\Lambda}^{1-s}\mathbf{Y}\left(s\right)^T\mathbf{G}_{LC}\left(s\right)\mathbf{G}_{LC}\left(s\right)^T\vec{Z}\left(s\right)ds, \tag{3.32}$$

where the columns of the matrix $\mathbf{Y}\left(\phi\right)$ are the IRCs $\vec{Y}_i\left(\phi\right)$, and

$$c = 2\vec{z}_G^T\vec{z}_G \equiv 2\int_0^1 \vec{Z}\left(s\right)^T\mathbf{G}_{LC}\left(s\right)\mathbf{G}_{LC}\left(s\right)^T\vec{Z}\left(s\right)ds, \tag{3.33}$$

respectively. As reflected in the above formulae, we define $\vec{z}_G$ to be an averaged version of the PRC $\vec{Z}\left(\phi\right)$ and the matrix $\mathbf{Y}_G$ to be an averaged version of the IRC matrix $\mathbf{Y}\left(\phi\right)$ (see Appendix B.2.2 for an extended discussion of the averaging procedure). $\vec{b}$ and c are therefore quantities that reflect the intrinsic oscillator dynamics: they do not vary with D_{in} or depend on the size and shape of E within the Poincare section. It will be the focus of the next section to develop an understanding of the E- and D_{in}-dependent elements $\mathcal{E}$, $\mathcal{E}_m$, and x_E and how they interact to form unruliness given different values of $\vec{b}$ and c.

Before moving on, we consider the limiting case where $E = S$. In this case, since there are no steps in the Markov renewal process that are not events, no steps x_k are distinguished from any other. We see this in the TVGR (3.31), where $x_E = 0$ and $\mathcal{E} = \mathcal{E}_m = 1$ when $E = S$. Only the final, temporal term, $cD_{in} = 2\vec{z}_G^T\vec{z}_G D_{in}$, survives, and so the effective diffusion coefficient $D_{\mathit{eff}} = \frac{1}{2}\mathcal{V}_E^{(t)}$ is precisely the linear phase diffusion D_{phase} (3.3) from the phase reduction theory.

More generally, when $x = 0 \in E \subset S$, both the Markov-only and mixed terms are eliminated as $D_{in} \to 0$. Specifically, $\mathcal{V}^{(n)}_{x \mapsto 1_E(x)}$ approaches 0 as $D_{in} \to 0$, since events are produced with high probability and little variance. The event center of mass x_E also approaches 0 as $D_{in} \to 0$. Once again, only the temporal term survives. Therefore when the event subsection intersects the limit cycle ($0 \in E$), D_{eff} coincides with D_{phase} in the weak noise limit. This shows that the effective diffusion coefficient D_{eff} is the result of an augmentation to the phase diffusion that takes effect above the weak-noise regime. This follows from the fact that the phase-isostable representation, (3.20) and (3.21), augments the standard phase reduction. Note that, in particular, oscillator models in $d+1$ dimensions are represented in the full TVGR (3.31) by $2d$ additional derived parameters, the d components of each Λ and $\vec{b}$, beyond the scalar c, which is associated with standard phase reduction.

3.6. Main Results: Planar Oscillators

We now narrow our focus to one-dimensional Poincare maps that arise from two-dimensional oscillators. In this case, b, x_E, Λ, and Γ_{ss} become scalar quantities and only one column appears in $Y_G \equiv \vec{y}_G$. From (3.25) and (B.46) one obtains

$$\Gamma_{ss} = \frac{1}{1-\Lambda^2}\Gamma = \frac{1}{1-\Lambda^2}\left\|\vec{y}_G\right\|^2 , \tag{3.34}$$

and the expression for b, (3.32), simplifies to

$$b = 2\left(1+\Lambda\right)\frac{\vec{z}_G^T\vec{y}_G}{\left\|\vec{y}_G\right\|^2} . \tag{3.35}$$

The Cauchy-Schwarz inequality applied to $\vec{z}_G^T \vec{y}_G$ shows that b^2 and $c = 2\left\|\vec{z}_G\right\|^2$ cannot be varied completely independently of each other,

$$1 \geq \frac{\left(\vec{z}_G^T \vec{y}_G\right)^2}{\left\|\vec{y}_G\right\|^2 \left\|\vec{z}_G\right\|^2} = \frac{b^2}{4(1+\Lambda)^2} \frac{\left\|\vec{y}_G\right\|^2}{\left\|\vec{z}_G\right\|^2} = \frac{1}{2}\frac{1-\Lambda}{1+\Lambda}\Gamma_{ss}\frac{b^2}{c}\,.$$

Thus, the $b-c$ parameter space is limited to a parabolic region that depends on Λ and Γ_{ss}

$$\frac{c}{\Gamma_{ss}} \geq \frac{1}{2}\frac{1-\Lambda}{1+\Lambda}b^2. \tag{3.36}$$

We will further limit our investigation to cases where E is a single, finite interval in the one-dimensional section S and where the limit cycle intersects E, so that events are produced regularly in the absence of noise. Since S is parameterized by the single number $x = \psi$, this is equivalent to taking

$$E = \left(w\left(\delta - \frac{1}{2}\right), w\left(\delta + \frac{1}{2}\right)\right), \tag{3.37}$$

where $-\frac{1}{2} < \delta < \frac{1}{2}$ and w is the width of the interval. The TVGR will then depend parametrically on w and δ as well as b, c, and Γ_{ss}:

$$\begin{aligned}\mathcal{V}_E^{(t)}\left(D_{in}; w, \delta, b, c, \Gamma_{ss}\right) &= \mathcal{V}_{x\mapsto 1_E(x)}^{(n)}\left(D_{in}; w, \delta, \Gamma_{ss}\right) + \\ &\quad + b\, x_E\left(D_{in}; w, \delta, \Gamma_{ss}\right)\, \mathcal{E}^2\left(D_{in}; w, \delta, \Gamma_{ss}\right) + \\ &\quad + c\, D_{in}\, \mathcal{E}^2\left(D_{in}; w, \delta, \Gamma_{ss}\right).\end{aligned} \tag{3.38}$$

The dependence of $\mathcal{V}_E^{(t)}\left(D_{in}; w, \delta, b, c, \Gamma_{ss}\right)$ on its various parameters can be simplified by noting their symmetry and scaling properties. Since, under linearization of the dynamics, the relevant joint densities are Gaussian and therefore symmetric, $\mathcal{E}$ and $\mathcal{E}_m$ are invariant and x_E is negated under the reflection of E about 0, $\delta \to -\delta$. And so the TVGR is invariant when b is also negated (see (3.38)), i.e.

$$\mathcal{V}_E^{(t)}\left(D_{in}; w, \delta, b, c, \Gamma_{ss}\right) = \mathcal{V}_E^{(t)}\left(D_{in}; w, -\delta, -b, c, \Gamma_{ss}\right).$$

Therefore we only need to consider intervals E, (3.37), with $0 \leq \delta < \frac{1}{2}$ so that $x_E \geq 0$. Note also that w can be absorbed into D_{in} in $\mathcal{E}$, $\mathcal{E}_m$, and x_E by a rescaling of b and c,

$$D_{in} \leftarrow \frac{D_{in}}{w^2},\ b \leftarrow wb, c \leftarrow w^2 c. \tag{3.39}$$

Thus, a larger event window is compensated for by larger noise. Likewise, without loss of generality, we could also effectively set $\Gamma_{ss} = 1$ by absorbing Γ_{ss} into D_{in} and correspondingly rescaling c (see (3.38)):

$$D_{in} \leftarrow \Gamma_{ss} D_{in}, c \leftarrow \frac{c}{\Gamma_{ss}}. \tag{3.40}$$

In total,

$$\mathcal{V}_E^{(t)}\left(D_{in}; w, \delta, b, c, \Gamma_{ss}\right) = \mathcal{V}_E^{(t)}\left(\frac{\Gamma_{ss}}{w^2} D_{in}; 1, \delta, wb, \frac{w^2}{\Gamma_{ss}} c, 1\right), \tag{3.41}$$

and so, for any $\Gamma_{ss} > 0$ and $w > 0$, $\mathcal{V}_E^{(t)}$ can be understood by considering the case $\Gamma_{ss} = w = 1$ and varying b and c. Therefore, we at times only consider the behavior of $\mathcal{E}\left(D_{in}; w, \delta, \Gamma_{ss}\right)$ and $x_E\left(D_{in}; w, \delta, \Gamma_{ss}\right)$ for $\Gamma_{ss} = w = 1$.

Our main goal is to understand the origin of the unruliness of $\mathcal{V}_E^{(t)}$, i.e. of its strong rise and subsequent decrease with increasing D_{in}. We therefore consider in detail the D_{in}-dependence of the three terms in (3.38). From (3.27) and (3.30) we find the exact formulae

$$\mathcal{E}\left(D_{in};w,\delta,\Gamma_{ss}\right)=\frac{1}{2}\left[\operatorname{erf}\left(\frac{w\left(\delta+\frac{1}{2}\right)}{2\sqrt{\Gamma_{ss}D_{in}}}\right)-\operatorname{erf}\left(\frac{w\left(\delta-\frac{1}{2}\right)}{2\sqrt{\Gamma_{ss}D_{in}}}\right)\right] \tag{3.42}$$

and

$$x_E\left(D_{in};w,\delta,\Gamma_{ss}\right)=\frac{1}{\mathcal{E}}\sqrt{\frac{\Gamma_{ss}D_{in}}{\pi}}\left(\exp\left(-\frac{w^2\left(\delta-\frac{1}{2}\right)^2}{4\Gamma_{ss}D_{in}}\right)-\exp\left(-\frac{w^2\left(\delta+\frac{1}{2}\right)^2}{4\Gamma_{ss}D_{in}}\right)\right). \tag{3.43}$$

For any $\delta\in\left[0,\frac{1}{2}\right)$, the event probability $\mathcal{E}$ is monotonically decreasing as a function of D_{in} (Figure 3.4a,b.i), reflecting the increased likelihood to miss the event sub-section E as the noise is increased. In parallel, the event center of mass x_E, which is a scalar for planar oscillators, monotonically increases from 0 to the centroid of E at $x=w\delta$, since the probability density becomes ever more homogeneous across E (Figure 3.4a,b.ii).

For the Markov-only term $\mathcal{V}^{(n)}_{x\mapsto 1_E(x)}$ no analytical expression is available for general Λ due to the appearance of the term $\mathcal{E}_m$ in (3.28). As we will see below in Section (3.6.2), in a limit where $\Lambda\to 0$ and $w\to 0$ simultaneously, $\mathcal{E}_m$ is given by $\mathcal{E}^2$ and therefore $\mathcal{V}^{(n)}_{x\mapsto 1_E(x)}=\mathcal{E}\left(1-\mathcal{E}\right)$ is non-monotonic in D_{in} (Figure 3.4b.iii). As in the toy model, the mixed term $x_E\mathcal{E}^2$ has a maximum at intermediate values of D_{in} and the temporal term $D_{in}\mathcal{E}^2$ increases monotonically, reaching a constant value asymptotically (Figure 3.4b.iv and b.v).

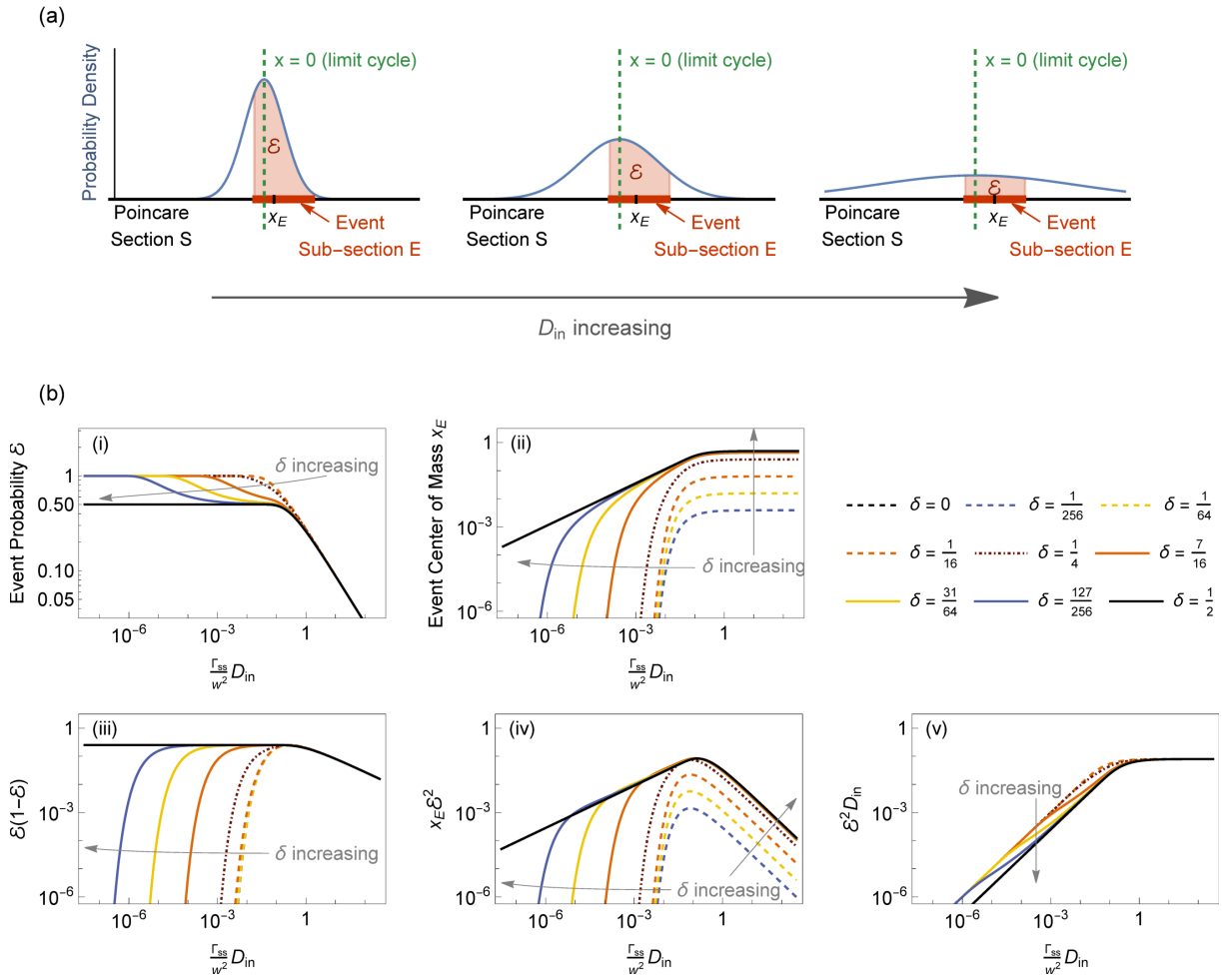

Figure 3.4. **The D_{in}-dependent Elements of the TVGR $\mathcal{V}_E^{(t)}$.** (a): A schematic depiction of the steady-state distribution on the Poincare section S (black line) for $\delta > 0$ and various D_{in}. $\mathcal{E}$ is the probability mass associated with E (red shading above the red interval), and x_E is its center of mass. (b): Plots of the TVGR elements as a function D_{in} for various values of δ. The gray arrows indicate roughly how the graphs change with increasing δ; all converge to the solid black curve as $\delta \to \frac{1}{2}$. The top row shows the "raw" elements, $\mathcal{E}$ and x_E. Note that x_E is identically 0 for $\delta = 0$. In the quasi-renewal limit (see Section 3.6.2), the TVGR is a linear combination of the components plotted in (b.iii) and (b.v), $\mathcal{E}(1-\mathcal{E})$ and $D_{in}\mathcal{E}^2$.

Overall, the D_{in}-dependence of $\mathcal{V}_E^{(t)}$ depends on the balance between the first two and the third term in (3.38). Thus, b and c, which represent the specific dynamics

of the oscillator via its averaged PRC and IRC (see (3.32) and (3.33)), are the primary parameters in the TVGR. We will show in the following sections that for any choice of the interval E containing 0, unruliness appears in a set of finite measure in the b-c parameter space. After giving a concrete criterion for unruliness, we start our investigation in Section 3.6.2 with a limiting case of the dynamics where the TVGR can be treated exactly. In the subsequent sections we then provide analytical estimates for the Markov-only term that allow us to identify sufficient conditions on b and c to guarantee unruliness of $\mathcal{V}_E^{(t)}$, resulting in the phase diagram Figure 3.7.

3.6.1. The Criterion for Unruliness

The main goal of this chapter is to demonstrate widespread unruliness in the event output of noisy oscillators. To make progress, we therefore need to make the criterion for unruliness concrete and quantitative. At the same time, unruliness is in some sense a qualitative phenomenon, where $\mathcal{V}_E^{(t)}$ as a function of D_{in} is initially linear, has a *strong* monotonic rise, attains a maximum, and then eventually decreases. Any quantitative test for unruliness should, therefore, only be interpreted as a rough "rule of thumb". We briefly consider three criteria, before settling on one. First, the most obvious feature of unruliness as seen in Figures 3.1b and 3.3a is the maximum in the TVGR. A criterion could therefore simply be that $\mathcal{V}_E^{(t)}$ has a local maximum. This is unfortunately too generous, since a local maximum does not imply a strong rise. For instance, in the simple quasi-renewal case discussed in Section 3.6.2 below, $\mathcal{V}_E^{(t)}$ has a local maximum for any value of $c \geq 0$, but not all cases shown in Figure 3.5 should qualify as unruly. A better criterion is motivated by the visual appearance of graphs of sums of functions on a log-scale

plot. On the log-scale, addition $f + g$ becomes - very roughly - the maximum operation $\max(\log f, \log g)$. This can be seen, e.g., in Figure 3.3a, where the total TVGR (black) closely follows the larger of the temporal component (blue, dashed) and the Markov component (orange, solid). Following that line of reasoning, the total TVGR would have a prominent maximum if, in the vicinity of the maximum of the non-monotonic component $\mathcal{V}^{(t)}_{E,non\text{-}monotonic} \equiv \mathcal{V}^{(n)}_{x\mapsto 1_E(x)} + bx_E\mathcal{E}^2$, that component dominated the monotonic, temporal component $\mathcal{V}^{(t)}_{E,monotonic} \equiv cD_{in}\mathcal{E}^2$. I.e.

$$\mathcal{V}^{(t)}_{E,non\text{-}monotonic}(D^{(max)}_{in}) > \mathcal{V}^{(t)}_{E,monotonic}(D^{(max)}_{in}) \quad \text{with } \frac{\partial}{\partial D_{in}}\mathcal{V}^{(t)}_{E,non\text{-}monotonic}(D_{in})\Big|_{D_{in}=D^{(max)}_{in}} = 0\,. \tag{3.44}$$

(3.44) is a reasonable criterion that effectively separates those cases which we would reasonably classify as unruly or not unruly. But we found that the analysis of (3.44) becomes unnecessarily complicated for something meant to be "just a rule of thumb". We therefore opt for a yet stricter criterion that is more amiable to analysis and interpretation. Given the monotonic saturation of $\mathcal{V}^{(t)monotonic}_{E}$ for $D_{in} \to \infty$ (see Figure 3.4b.v), for $\mathcal{V}^{(t)}_{E}$ to be classified as unruly we require that the *maximum* of $\mathcal{V}^{(t)}_{E,non\text{-}monotonic}$ is greater than the *asymptotic* value (and upper bound) of $\mathcal{V}^{(t)}_{E,monotonic}$, which is also the asymptotic value of $\mathcal{V}^{(t)}_{E}$:

$$\max_{D_{in}}\left\{\mathcal{V}^{(t)}_{E,non\text{-}monotonic}\right\} > \lim_{D_{in}\to\infty}\mathcal{V}^{(t)}_{E,monotonic}. \tag{3.45}$$

This criterion guarantees a prominent maximum in the TVGR; it indeed picks out those cases that are strongly unruly (Figure 3.5).

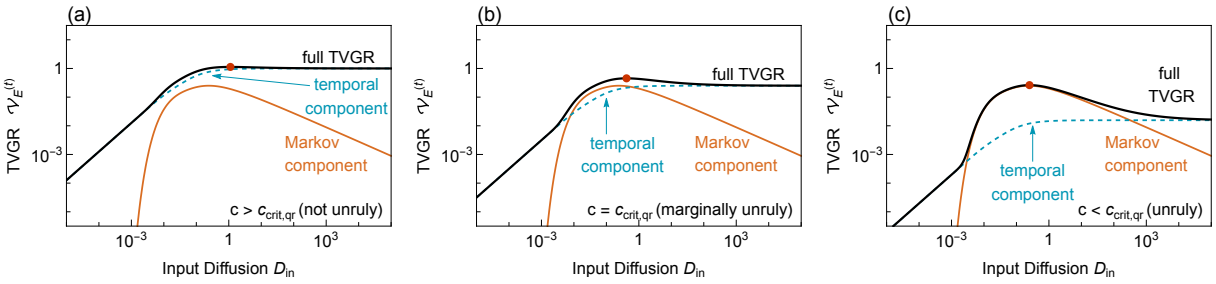

Figure 3.5. **Unruliness requires c to be sufficiently small.** Sample TVGRs in the quasi-renewal limit that are (a) not unruly with c above the unruliness threshold $c_{crit,r}$, (b) marginally unruly with c at the threshold, and (c) unruly with c below the threshold (3.50). The presence of a local maximum (red circle) does not distinguish these three cases. In the plots, we take $\delta = \frac{1}{4}$, but they are qualitatively representative for $\delta \in \left[0, \frac{1}{2}\right)$.

3.6.2. The Quasi-Renewal Case ($\Lambda \to 0$ with $w \to 0$)

We first consider the simple limiting case of strong contraction to the limit cycle, $\Lambda \to 0$. In this limit, we expect the steady-state variance on the section, Γ_{ss}, to go to 0. Indeed, with (3.34) and the analysis in SI Section B.2.2.3,

$$\Gamma_{ss} \sim \|\vec{y}_G\|^2 \sim -\frac{1}{\log \Lambda}.$$

For fixed values of D_{in}, w, and δ, we have then $\mathcal{E}\left(D_{in}; w, \delta\right) \to 1$, $\mathcal{E}_m\left(D_{in}; w, \delta\right) \to 1$, and $x_E\left(D_{in}; w, \delta\right) \to 0$; crossings of the Poincare section are always at $x = 0$, i.e. on the limit cycle. There are no phase slips, since events are produced at every crossing. This is unsurprising: for strong contraction the standard phase reduction is an accurate representation of the dynamics, and the process is effectively a renewal process. Indeed, $\mathcal{V}_E^{(t)}$ reduces in this limit to its phase-reduction value, $\mathcal{V}_E^{(t)} \to cD_{in} = 2D_{phase}$.

To obtain a non-trivial result in the limit of strong contraction, we need to take the width of E to 0, as well, in order to keep the probability of phase slips $1 - \mathcal{E}\left(D_{in}; w, \delta\right)$

unchanged. That requires

$$w \propto \sqrt{\Gamma_{ss}} \propto \sqrt{-\frac{1}{\log \Lambda}} \tag{3.46}$$

as $\Lambda \to 0$ (see (3.42)). In this case, it is clear from (3.43) that $x_E \to 0$ still, but the limiting behavior of $\mathcal{E}_m$ requires some care. Applying the definitions of E from (3.37) and $\mathbf{C}_m$ from (3.26), (3.29) becomes

$$\mathcal{E}_m = \int_{w\left(\delta-\frac{1}{2}\right)}^{w\left(\delta+\frac{1}{2}\right)} \int_{w\left(\delta-\frac{1}{2}\right)}^{w\left(\delta+\frac{1}{2}\right)} N\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}; 2D_{in}\Gamma_{ss}\begin{pmatrix} 1 & \Lambda^m \\ \Lambda^m & 1 \end{pmatrix}\right) dx_1 dx_2. \tag{3.47}$$

The dominant contribution to the covariance matrix for $\Lambda \ll 1$ is

$$2D_{in}\mathbf{C}_m \sim 2D_{in}\Gamma_{ss}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

and thus

$$\mathcal{E}_m \to \left[\int_{w\left(\delta-\frac{1}{2}\right)}^{w\left(\delta+\frac{1}{2}\right)} N\left(x; 2D_{in}\Gamma_{ss}\right) dx\right]^2 = \mathcal{E}^2. \tag{3.48}$$

In total (3.28) becomes $\mathcal{V}^{(n)}_{x\mapsto 1_E(x)} = \mathcal{E}\left(1-\mathcal{E}\right)$, reflecting the fact that there are no correlations in events from one step to the next. We therefore call this limiting case, given by $\Lambda \to 0$ along with (3.46), the "quasi-renewal" case. Since $x_E = 0$, the quasi-renewal TVGR,

$$\mathcal{V}^{(t)}_{E,qr}\left(D_{in}; c\right) = \mathcal{E}\left(1-\mathcal{E}\right) + cD_{in}\mathcal{E}^2, \tag{3.49}$$

is independent of b. Here the TVGR has essentially the same form as in the toy model discussed in Section 3.4 (cf. (3.19)) and unruliness arises if c is small enough for the non-monotonic Markov-only term to dominate over the temporal term. It can be shown that $D_{in}\mathcal{E}^2 \to \frac{w^2}{4\pi\Gamma_{ss}}$ independently of a as $D_{in} \to \infty$. Therefore, since $\mathcal{E}(1-\mathcal{E})$ has a maximum value of $\frac{1}{4}$, $\mathcal{V}^{(t)}_{E,qr}$ is unruly according to (3.45) if

$$c < c_{crit,qr} = \pi \frac{\Gamma_{ss}}{w^2}. \tag{3.50}$$

Sample TVGRs that satisfy (3.50) are indeed strongly unruly (Figure 3.5).

3.6.3. Symmetric Intervals E ($\delta = 0$, $\Lambda > 0$)

There are two effects as Λ is perturbed from 0: the mixed term $bx_E\mathcal{E}^2$ becomes generally non-zero and $\mathcal{E}_m$ deviates from $\mathcal{E}^2$. As a stepping stone to the most general case, we first consider only symmetric intervals given by $\delta = 0$. In that case, the event center of mass x_E is 0, and we can consider $\mathcal{E}_m$ alone in the question of how (3.50) generalizes.

In order to address unruliness in the TVGR in this case, we will need to understand the infinite sum in (3.28), $\sum_{m=1}^{\infty}(\mathcal{E}_m - \mathcal{E}^2)$. Bounding $\mathcal{E}_m$ in such a way that the sum over m can be carried out and is well behaved for all $\Lambda \in [0,1)$ and all $\delta \in \left[0, \frac{1}{2}\right)$ turns out to be quite involved. We show the full details in Appendix B.4 where we find that

$$0 \leq \mathcal{E}_m - \mathcal{E}^2 \leq \Lambda^m \mathcal{E}(1-\mathcal{E}),$$

and so

$$\mathcal{E}(1-\mathcal{E}) = \mathcal{V}^{(n)}_{x\mapsto 1_E(x)}(\Lambda = 0) \leq \mathcal{V}^{(n)}_{x\mapsto 1_E(x)} < \frac{1+\Lambda}{1-\Lambda}\mathcal{E}(1-\mathcal{E}), \tag{3.51}$$

via a geometric series. Thus, a lower bound on the non-monotonic part of the TVGR for $\Lambda > 0$ is found by replacing it with its $\Lambda = 0$ counterpart, that is by replacing $\mathcal{V}^{(n)}_{x\mapsto 1_E(x)}$ with $\mathcal{E}\,(1-\mathcal{E})$. Therefore, if the maximum of $\mathcal{E}\,(1-\mathcal{E})$ exceeds the asymptotic value $\frac{cw^2}{4\pi\Gamma_{ss}}$ of $cD_{in}\mathcal{E}^2$, then so will the maximum value of $\mathcal{V}^{(n)}_{x\mapsto 1_E(x)}$, and the unruliness criterion (3.45) will certainly be satisfied. So the quasi-renewal condition for unruliness, (3.50), can only underestimate the range of c for which unruliness appears. At the same time, the upper bound in (3.51) guarantees that for

$$c > \frac{1+\Lambda}{1-\Lambda} c_{crit,qr} \tag{3.52}$$

$\mathcal{V}^{(t)}_E\,(D_{in}; w, \delta = 0, b, c)$ is certainly not unruly.

Thus, despite the uncertainty involved in being able to only bound $\mathcal{V}^{(n)}_{x\mapsto 1_E(x)}$, we can make the following statements

$$\mathcal{V}^{(t)}_E\,(D_{in}; w, \delta = 0, b, c) \text{ is certainly unruly when } c < c_{crit,qr} \tag{3.53}$$

$$\mathcal{V}^{(t)}_E\,(D_{in}; w, \delta = 0, b, c) \text{ is certainly not unruly when } c > \frac{1+\Lambda}{1-\Lambda} c_{crit,qr} \tag{3.54}$$

(3.55)

$$\mathcal{V}^{(t)}_E\,(D_{in}; w, \delta = 0, b, c) \text{ may be unruly or not unruly for } c_{crit,qr} < c < \frac{1+\Lambda}{1-\Lambda} c_{crit,qr}\,.$$

When we consider $\delta \neq 0$ in the following, b and c will both be relevant parameters. The above statements will generalize to three regions of the b-c parameter space in which $\mathcal{V}^{(t)}_E$ is 1) certainly not unruly, 2) certainly unruly, and 3) may be either.

3.6.4. The General Case

Finally, we consider $\mathcal{V}_E^{(t)}$ in full generality for $\delta \neq 0$ and $\Lambda > 0$. Due to the asymmetry $\delta \neq 0$, the mixed term $bx_E\mathcal{E}^2$ does not drop out and one has to deal with two non-monotonic terms. Obtaining sharp, general conditions for the onset of unruliness is therefore quite difficult. We therefore content ourselves with sufficient conditions.

Before considering more detailed criteria, (3.56)-(3.58) below, we first note that as long as $b \geq 0$ and $\delta \geq 0$, the mixed term $bx_E\mathcal{E}^2$ is non-negative and non-monotonic (Figure 3.4b.iv) and can not detract from unruliness. So the criterion (3.53), which applied in the case that $\delta = 0$, will also guarantee unruliness when $\delta \geq 0$ and $b \geq 0$:

$$\mathcal{V}_E^{(t)}\left(D_{in}; w, \delta, b, c\right) \text{ is certainly unruly when } c < c_{crit,qr} \text{ and } b \geq 0\,.$$

Thus, *at least a finite region* of the parameter space is unruly, bounded by $c = c_{crit,qr} = \pi\frac{\Gamma_{ss}}{w^2}$ above, $b = 0$ to the left and to the right by $b = \sqrt{2\frac{c}{\Gamma_{ss}}\frac{1+\Lambda}{1-\Lambda}}$, which follows from the physicality condition (3.36).

But, moreover, for $b > 0$ the mixed term $bx_E\mathcal{E}^2$ and the Markov-only term $\mathcal{V}^{(n)}_{x\mapsto 1_E(x)}$ *both* contribute to the non-monotonicity of $\mathcal{V}_E^{(t)}$ (Figure 3.6a,b). We therefore choose as a broader, sufficient condition for unruliness that either of these terms by itself dominate the monotonic temporal term $cD_{in}\mathcal{E}^2$. And for $b < 0$ unruliness is guaranteed if the mixed term is small enough to not undermine the dominance of the Markov-only term over the temporal term (Figure 3.6c). We therefore require that the mixed term is dominated by the temporal term, which in turn is dominated by the Markov-only term. These sufficient

conditions can be summarized as either (I) the mixed term dominates,

$$\max_{D_{in}} \left\{ b x_E \mathcal{E}^2 \right\} > \lim_{D_{in}\to\infty} \mathcal{V}_E^{(t)} = c\frac{w^2}{4\pi\Gamma_{ss}}, \tag{3.56}$$

or (II) the Markov-only term dominates,

$$\max_{D_{in}} \left\{ \mathcal{V}^{(n)}_{x\mapsto 1_E(x)} \right\} > \lim_{D_{in}\to\infty} \mathcal{V}_E^{(t)} = c\frac{w^2}{4\pi\Gamma_{ss}}, \tag{3.57}$$

as long as b is not strongly negative,

$$-b < c \min_{D_{in}} \left\{ \frac{D_{in}}{x_E} \right\}. \tag{3.58}$$

If (3.58) is not satisfied, a more refined analyses would be required to determine unruliness. In the following we classify the unruliness in these situations, where $\mathcal{V}_E^{(t)}$ could even have a local minimum, as "unclear" (Figure 3.6d).

From (3.56)-(3.58), there will be four boundaries for unruliness in the parameter space. One boundary, below which unruliness certainly occurs, follows from (3.56). Because of the range of values that $\mathcal{V}^{(n)}_{x\mapsto 1_E(x)}$ can take on (see (3.51)), two more follow from (3.57): the boundary $c = \pi\frac{\Gamma_{ss}}{w^2}$ below which unruliness certainly occurs and the boundary $c = \frac{1+\Lambda}{1-\Lambda}\pi\frac{\Gamma_{ss}}{w^2}$ below which unruliness may possibly occur. These are generalizations of (3.53) and (3.55) for $\delta \geq 0$, and they apply as long as the condition on b, (3.58), is satisfied. Thus, the two regions corresponding to certain and possible unruliness will be cut short by the fourth boundary $-b = c\min_{D_{in}} \left\{ \frac{D_{in}}{x_E} \right\}$ beyond which the negative mixed term could diminish the unruliness.

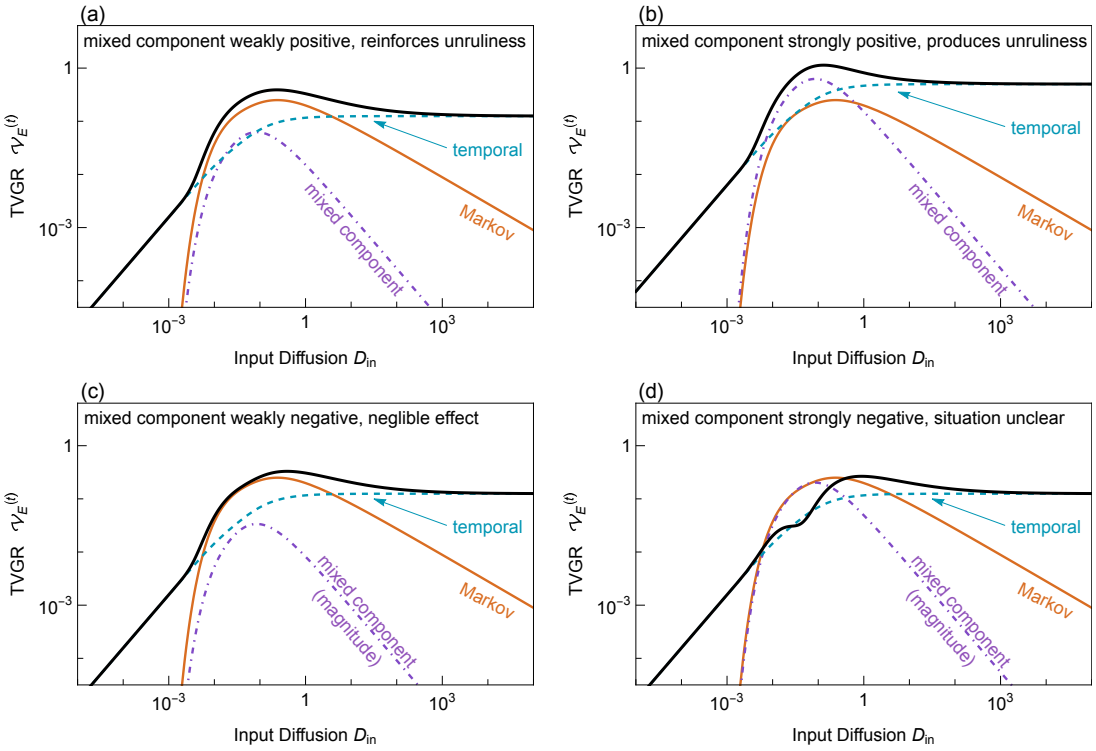

Figure 3.6. **The mixed component can promote or detract from unruliness.** Sample lower bounds (from (3.51)) of TVGRs that are unruly (a) because the Markov component dominates over the asymptotic value of the TVGR and (b) because b is positive and large and the mixed component dominates. (c): b is negative, but the mixed term is not large enough to detract significantly from the unruliness. (d): a situation where the typical unruly quality is lost and a local minimum appears because b is sufficiently large and negative. This situation is outside the scope of our analysis.

In order to represent all of the possibilities graphically in a single phase diagram (Figure 3.7), we minimize the number of parameters that need to be considered. We have already shown in (3.41) that the event interval width w and the steady-state variance coefficient Γ_{ss} can be absorbed by rescaling D_{in}, b, and c. In the following we make this rescaling explicit by defining

$$b_\Lambda \equiv \frac{1-\Lambda}{1+\Lambda} wb \qquad \text{and} \qquad c_\Lambda \equiv \frac{1-\Lambda}{1+\Lambda} w^2 \frac{c}{\Gamma_{ss}}. \tag{3.59}$$

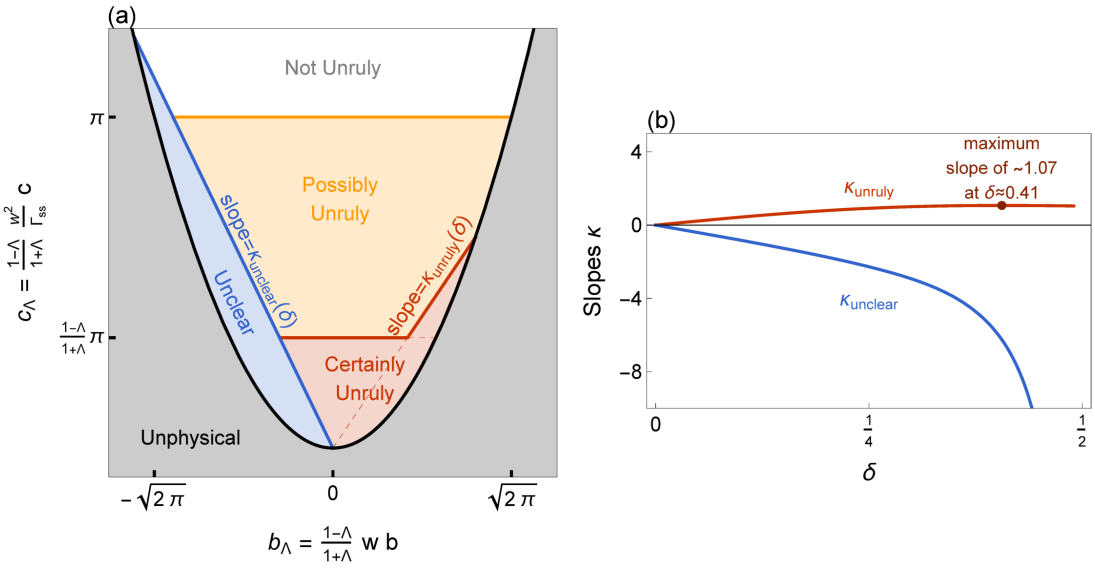

Figure 3.7. **Dependence of unruliness on b_Λ and c_Λ (cf. (3.59)).** (a): The b_Λ-c_Λ phase diagram, representative for all $\Lambda > 0$ and $0 < \delta < \frac{1}{2}$. The physical portion of the diagram is divided into four regions, reflecting TVGRs that are not unruly (white), certainly unruly (red), and may be either (orange). The fourth, blue region corresponds to situations that may be different altogether, with a local minimum possibly appearing in the TVGR. The boundaries between physical regions are all piece-wise linear, and the pieces are either fixed ($c_\Lambda = \pi$), vary with Λ ($c_\Lambda = \frac{1-\Lambda}{1+\Lambda}\pi$), or have slopes that vary with δ ($c_\Lambda = \kappa(\delta)\, b_\Lambda$). (b): The δ-dependent slopes κ_{unruly} and $\kappa_{uncertain}$.

To be precise, the noise strength should also be explicitly scaled, $D_{in} \leftarrow \frac{\Gamma_{ss}}{w^2} D_{in}$. But that just shifts the noise level at which features in the TVGR, like unruliness, occur, and does not change whether they appear or not. We choose to include also a Λ-dependent factor in this rescaling, since it largely frees us from considering Λ explicitly. The physicality condition (3.36) is simply $c_\Lambda \geq \frac{1}{2} b_\Lambda^2$. And the threshold beyond which unruliness cannot occur, (3.52), becomes $c_\Lambda = \pi$, for *any* value of $\Lambda > 0$. The tradeoff is that the threshold in c_Λ below which unruliness certainly does occur, $c_\Lambda = \frac{1-\Lambda}{1+\Lambda}\pi$, depends on Λ. The other criteria, (3.56) and (3.58), remain independent of Λ. They correspond to linear boundaries

that pass through the origin in the b_Λ-c_Λ plane (Figure 3.7a), with δ-dependent slopes given by

$$\kappa_{unclear}(\delta) = -\max_{D_{in}} \left\{ \frac{x_E(D_{in};\delta)}{D_{in}} \right\} \tag{3.60}$$

and

$$\kappa_{unruly}(\delta) = 4\pi \max_{D_{in}} \left\{ x_E(D_{in};\delta)\, \mathcal{E}^2(D_{in};\delta) \right\} , \tag{3.61}$$

respectively (Figure 3.7b). Note that in the case $\delta = 0$, which was discussed in Section 3.6.3, $\kappa_{unclear} = \kappa_{unruly} = 0$: the phase diagram is divided into horizontal bands corresponding to (3.53)-(3.55). And, as $\delta \to \frac{1}{2}$, $\kappa_{unclear}$ diverges: in this limit, the entire left hand of the phase diagram becomes "unclear".

It is worthwhile considering what happens when the width of the event interval w is changed. Varying w while keeping the remaining parameters fixed amounts to tracing out parabolic curves in the (b_Λ, c_Λ)-plane that are defined by $b_\Lambda = \mathcal{B}\sqrt{c_\Lambda}$ (cf. (3.59)) with $\mathcal{B}$ a constant (dashed lines in Figure 3.8a). Depending on $\mathcal{B}$ different scenarios arise (Figure 3.8b). Examples of the traces corresponding to the different regimes in Figure 3.8b are shown as dashed lines in Figure 3.8a. Considering the limit $w \to 0$, $\mathcal{V}_E^{(t)}$ "certainly" becomes unruly if $b > 0$ (scenario (i)), while for $b < 0$ the outcome is "unclear" (scenarios (ii)-(iv)). As w is increased from 0, $\mathcal{V}_E^{(t)}$ can become "certainly unruly" (scenario (ii)) or "possibly unruly" (scenarios (ii) and (iii)) even when $b < 0$. The various transitions define the values $w^*_{unclear}$, $w^*_{certain}$, and $w^*_{possible}$, respectively, which are determined in Appendix B.5.

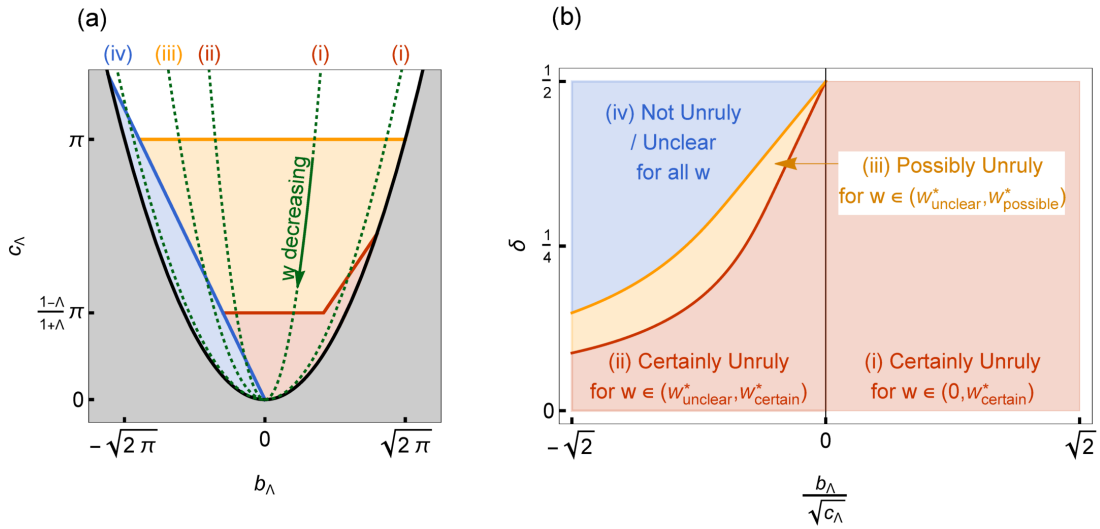

Figure 3.8. **Dependence of unruliness on the event interval E.** (a): The effect of varying the interval width, w, in the b_Λ-c_Λ phase diagram is to trace out parabolic curves of constant $\frac{b_\Lambda}{\sqrt{c_\Lambda}}$. Four qualitatively different behaviors appear, depending on $\frac{b_\Lambda}{\sqrt{c_\Lambda}}$: (i) unruliness will certainly appear for some interval $w \in (0, w^*_{certain})$, (ii) unruliness will certainly appear for $w \in (w^*_{unclear}, w^*_{certain})$, (iii) unruliness could possibly appear for $w \in (w^*_{unclear}, w^*_{possible})$, and (iv) the oscillator is either not unruly or the situation is unclear for all w. (b): A phase diagram over $\frac{b_\Lambda}{\sqrt{c_\Lambda}}$ and δ summarizes the appearance of the four different unruliness behaviors (i)-(iv).

In summary, since the physically accessible parameter range for planar oscillators is given by

$$-1 \leq \frac{1}{\sqrt{2}} \frac{b_\Lambda}{\sqrt{c_\Lambda}} = \frac{\vec{z}_G^T \vec{y}_G}{\|\vec{z}_G\| \, \|\vec{y}_G\|} \leq 1 \tag{3.62}$$

(see (3.36) and (3.59)), and all finite event intervals E that intersect the limit cycle satisfy $0 \leq \delta < \frac{1}{2}$ up to reflection, Figure 3.8b captures all planar oscillators and their relevant event intervals. It shows that for any value of $\mathcal{B} = b_\Lambda/\sqrt{c_\Lambda}$, which characterizes the

internal dynamics of the oscillator, there is a range of event intervals, characterized by δ and w, for which the TVGR $\mathcal{V}_E^{(t)}$ is certainly unruly.

3.6.5. Example: Hopf Normal Form

One might rightly complain that while the b_Λ-c_Λ plane (Figure 3.7a) shows a substantial region of unruliness, b and c are derived parameters, whose relationship to the parameters that originally appear in a given oscillator model are not clear. Likewise, while *some* choice of E produces unruliness in any oscillator, that choice may not be of interest in a given model or application. In short, it is not clear that the unruly region is reachable by reasonable choices of the model's parameters and a reasonable choice of interval E. Here we address this concern by investigating the prototypical limit-cycle oscillator. We compute the reduced parameters b and c for the unfolded Hopf normal form, also known as the Stuart-Landau oscillator, with added noise,

$$d\vec{y} = 2\pi \left[\begin{pmatrix} \epsilon & -\alpha \\ \alpha & \epsilon \end{pmatrix} \vec{y} - \begin{pmatrix} \beta & -\gamma \\ \gamma & \beta \end{pmatrix} |\vec{y}|^2 \, \vec{y} \right] dt + \sigma \mathbf{G} \, d\vec{W}.$$

In this coordinate system, the limit cycle is a circle centered at the origin with radius $\sqrt{\frac{\epsilon}{\beta}}$ and has a period of $\alpha - \frac{\epsilon\gamma}{\beta}$. Without loss of generality, we rescale $\vec{y}$ and t so that the radius and period are both fixed at 1. We can equivalently set $\beta = \epsilon$ and $\alpha = \gamma + 1$.

We also make three simplifying choices of parameters in our analysis. Despite these choices, the oscillator will show unruliness over a wide range of its remaining natural parameters. First, we set $\gamma = 0$, eliminating the amplitude-dependence of the deterministic

angular velocity. This yields what we call the "Hopf oscillator",

$$d\vec{y} = 2\pi \left[\begin{pmatrix} \epsilon & -1 \\ 1 & \epsilon \end{pmatrix} \vec{y} - \begin{pmatrix} \epsilon & 0 \\ 0 & \epsilon \end{pmatrix} |\vec{y}|^2 \vec{y} \right] dt + \sigma \mathbf{G}\, d\vec{W}, \tag{3.63}$$

for $\epsilon > 0$. We also choose the Poincare section to be along the positive y_1 axis, and leave the event interval E unspecified for now, investigating below what is required of it to produce unruliness. Lastly, we allow the additive noise to be correlated in the y_1 and y_2 components but have the same marginal variance in each; we parameterize the covariance matrix as

$$\sigma^2 \mathbf{G}\mathbf{G}^T = \sigma^2 \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix},$$

where $\rho \in [-1, 1]$ is the correlation coefficient between the components.

In order to rewrite the system in phase-isostable coordinates, (3.21), we first note that, with the choice $\gamma = 0$, the geometric angle $\theta = \arctan \frac{y_2}{y_1}$ evolves with a constant rate when the oscillator is unforced ($\sigma = 0$). So the isochrons are radial lines and we will take $\phi = \frac{\theta}{2\pi}$ as the phase coordinate. The radial coordinate $r = \sqrt{y_1^2 + y_2^2}$ evolves by

$$\dot{r} = 2\pi\epsilon r \left(1 - r^2\right).$$

$\psi \sim r - 1$ is the isostable coordinate to first order approximation, since it has linear dynamics near the limit cycle in the absence of noise,

$$\dot{\psi} = -4\pi\epsilon\psi + \mathcal{O}\left(\psi^2\right)$$

(cf. (3.21)). We therefore have $\kappa = 4\pi\epsilon$ and $\Lambda = \mathrm{e}^{-4\pi\epsilon}$. The PRC $\vec{Z}(\phi)$ and the IRC $\vec{Y}(\phi)$ follow by evaluating the appropriate gradients on the limit cycle:

$$\begin{aligned}
\vec{Z}(\phi) &= \left.\frac{d\Phi}{d\vec{y}}\right|_{r=1} = \frac{1}{2\pi}\begin{pmatrix} -\sin(2\pi\phi) \\ \cos(2\pi\phi) \end{pmatrix} \\
\vec{Y}(\phi) &= \left.\frac{d\Psi}{d\vec{y}}\right|_{r=1} = \begin{pmatrix} \cos(2\pi\phi) \\ \sin(2\pi\phi) \end{pmatrix},
\end{aligned}$$

which, once averaged, yield via (B.25)-(B.27),

$$\begin{aligned}
\vec{z}_G^T\vec{z}_G &= \frac{1}{4\pi^2} \\
\vec{z}_G^T\vec{y}_G &= \frac{1-\mathrm{e}^{-4\pi\epsilon}}{8\pi^2}\frac{\epsilon\rho}{1+\epsilon^2} \\
\vec{y}_G^T\vec{y}_G &= \frac{1-\mathrm{e}^{-8\pi\epsilon}}{8\pi\epsilon}\frac{1-2\epsilon\rho+4\epsilon^2}{1+4\epsilon^2}.
\end{aligned}$$

Then from (3.35) and (3.33),

$$b = \frac{2}{\pi}\frac{\epsilon^2}{1+\epsilon^2}\frac{1+4\epsilon^2}{1-2\epsilon\rho+4\epsilon^2}\rho \tag{3.64}$$

$$\frac{c}{\Gamma_{ss}} = \frac{4}{\pi}\epsilon\frac{1+4\epsilon^2}{1-2\epsilon\rho+4\epsilon^2} \tag{3.65}$$

and from (3.59) and (3.62),

$$\frac{c_\Lambda}{b_\Lambda} = 2w\frac{1+\epsilon^2}{\epsilon\rho}$$

and

$$\frac{b_\Lambda}{\sqrt{c_\Lambda}} = \frac{1}{\sqrt{\pi}}\sqrt{\frac{1-\mathrm{e}^{-4\pi\epsilon}}{1+\mathrm{e}^{-4\pi\epsilon}}}\frac{\epsilon^{3/2}}{1+\epsilon^2}\sqrt{\frac{1+4\epsilon^2}{1-2\epsilon\rho+4\epsilon^2}}\,\rho.$$

Note that $\frac{b_\Lambda}{\sqrt{c_\Lambda}}$ has the same sign as ρ and is monotonically increasing in ρ.

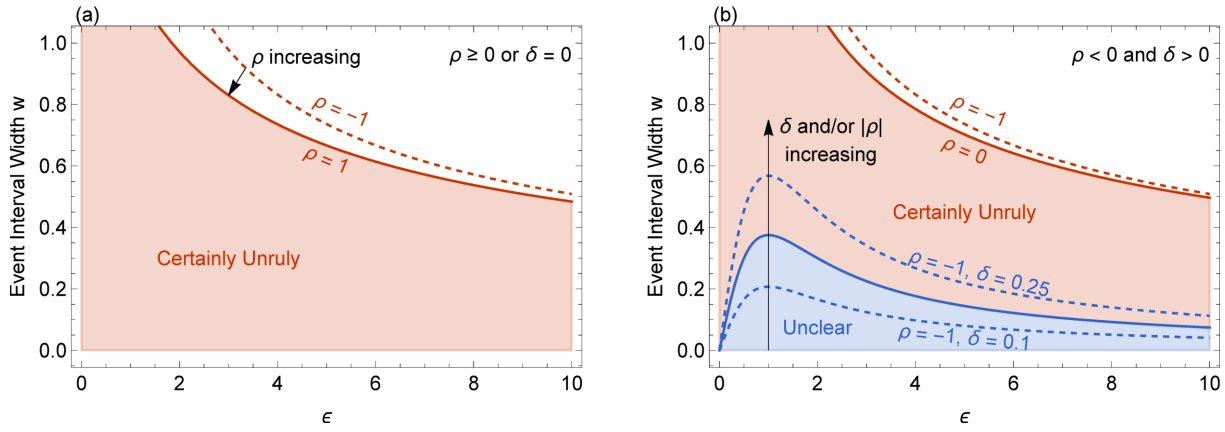

Figure 3.9. **Dependence of unruliness on ϵ and w in the Hopf Oscillator.** Regions of certain unruliness appear prominently in the Hopf Oscillator. (a): For $\rho \geq 0$ or $\delta = 0$, there is no unclear region and certain unruliness appears when either w or ϵ take on moderate or small values. (b): When both $\rho < 0$ and $\delta > 0$, the unclear region appears for small w. Regions of possible unruliness are not shown.

Recall that there are two thresholds beyond which unruliness will certainly occur: (3.56) and (3.57), which correspond to the line with slope $\kappa_{unruly}(\delta)$ and the lower horizontal line segment, respectively, in Figure 3.7a. For the Hopf oscillator, it turns out that the second criterion, $c < c_{crit} = \pi \frac{\Gamma_{ss}}{w^2}$ with $\frac{c_\Lambda}{b_\Lambda} > \kappa_{unclear}(\delta)$, is sufficient to describe the onset of certain unruliness. That is because $\frac{b_\Lambda}{\sqrt{c_\Lambda}}$ and $\kappa_{unruly}(\delta)$ are both bounded above in such a way that the first criterion is met only if the second is already satisfied. In particular, as w is decreased, the parameterized curve $(b_\Lambda(w), c_\Lambda(w))$ always intersects the line $c = c_{crit}$ before the line $\frac{c_\Lambda}{b_\Lambda} = \kappa_{unruly}(\delta)$. (This is similar to the left parabolic curve labelled by "(i)" in Figure 3.8a and in contrast with the rightmost curve.) We therefore need only consider how the Hopf oscillator's parameters ϵ, ρ, δ, and w interact with the boundary $c = c_{crit}$ for certain unruliness and the boundary $\frac{c_\Lambda}{b_\Lambda} = \kappa_{unclear}(\delta)$ for the "unclear" region.

For the Hopf oscillator, the second unruliness criterion (3.57) becomes

$$\frac{4}{\pi}\epsilon\frac{1+4\epsilon^2}{1-2\epsilon\rho+4\epsilon^2}=\frac{c}{\Gamma_{ss}}<\frac{c_{crit}}{\Gamma_{ss}}=\frac{\pi}{w^2}$$

or

$$w<\frac{\pi}{2}\sqrt{\frac{1-2\epsilon\rho+4\epsilon^2}{\epsilon\left(1+4\epsilon^2\right)}},$$

which is satisfied for moderate and small values of each w and ϵ (red region, Figure 3.9a and b). There is no "unclear" region when $\rho \geq 0$, since the mixed component of the TVGR is positive and can only enhance the unruliness. In this case, there is a prominent portion of the parameter space which shows certain unruliness, and it is realized with very reasonable parameter choices, e.g. $w \lesssim \frac{1}{2}$ with $\epsilon < 10$. Recall that in Hopf oscillator has a limit cycle of radius 1; only event interval widths w that are less than 1 are sensible. The parameter ϵ measures the strength of contraction to the limit cycle, and corresponds to a damping time constant of $\frac{1}{2\epsilon}$. Since the limit cycle period is fixed to be 1, without further, specific information about the system, ϵ should be "expected" to be $\mathcal{O}(1)$. Figure 3.9a shows that for such choices of w and ϵ, the TVGR is largely unruly.

When $\rho < 0$ and $\delta \neq 0$, $\frac{b_\Lambda}{\sqrt{c_\Lambda}}$ and, therefore, the mixed component of the TVGR is negative, a region where the situation is unclear appears in the ϵ-w plane. The nature of the TVGR is unclear when (3.58) is not satisfied. That happens when

$$2w\frac{1+\epsilon^2}{\epsilon\rho}=\frac{c_\Lambda}{b_\Lambda}<\kappa_{unclear}\left(\delta\right)$$

or

$$w<\frac{1}{2}\frac{\epsilon}{1+\epsilon^2}\rho\kappa_{unclear}\left(\delta\right).$$

Since $\kappa_{unclear}(\delta)$ is negative and increasing in magnitude with a, the corresponding region grows with δ and $|\rho|$ (Figure 3.9b). For moderate asymmetry δ in the event interval, there is still a prominent region of certain unruliness even in this case.

3.6.6. Comparison of Theoretical Bounds, Numerics, and Simulation

We also make use the Hopf oscillator example to empirically validate our theory by comparing with numerics and simulations. We consider two parameter sets: $(\epsilon = 1, \rho = 0, \delta = 0)$ and $(\epsilon = 0.01, \rho = 1, \delta = 0.25)$. The first corresponds to "typical" parameter choices for the Hopf oscillator: $\epsilon = 1$ corresponds to an $\mathcal{O}(1)$ time constant for the contraction to the limit cycle that is comparable to the limit cycle period, $\rho = 0$ to isotropic additive noise, and $\delta = 0$ to the event subset E placed symmetrically around the limit cycle. Comparing the theory and simulation using (in addition) the second, non-trivial parameter set validates our derivation of the TVGR (3.31) and the bounds (3.51) we place on it in the case of planar oscillators.

We consider the following computations of the TVGR:

(1) The theoretical bounds (see (3.31) and (3.51)),

$$\left(\mathcal{E} - \mathcal{E}^2\right) + bx_E\mathcal{E}^2 + \frac{c}{2}\sigma^2\mathcal{E}^2 \leq \mathcal{V}_E^{(t)} \leq \frac{1+\Lambda}{1-\Lambda}\left(\mathcal{E} - \mathcal{E}^2\right) + bx_E\mathcal{E}^2 + \frac{c}{2}\sigma^2\mathcal{E}^2.$$

(2) Numerical approximation of the theoretical results via a truncation of the series in $\mathcal{V}^{(n)}_{x\mapsto 1_E(x)}$ to $M = 50$ terms (cf. (3.28)),

$$\mathcal{V}_E^{(t)} \approx \left(\mathcal{E} - \mathcal{E}^2\right) + 2\sum_{m=1}^{M}\left(\mathcal{E}_m - \mathcal{E}^2\right) + bx_E\mathcal{E}^2 + \frac{c}{2}\sigma^2\mathcal{E}^2,$$

where the double integral (3.47) for $\mathcal{E}_m$ is computed exactly along one dimension and via numerical integration in the other. Note that, like the theoretical bounds, this numerical estimate is subject to the first order asymptotic approximations in the noise strength and isostable coordinate developed in Section 3.5 and the corresponding Appendices B.2 and B.3.

(3) Simulation of the Hopf oscillator model with 8192 independent noise realizations. We use the Euler-Maruyama time-stepping scheme with a step size of $\Delta t = 10^{-6}$ for a total time of $t_{run} = 900$ (after a burn-in period of 100) and determine the times T_k^E at which each realization crosses the line segment E via linear interpolation between the Euler-Maruyama steps. We then empirically estimate the asymptotic TVGR as (see the definitions of "Event Rate" and "Event Dispersion Rate" in Table 3.1)

$$\mathcal{V}_E^{(t)} \approx n^2 \frac{\hat{\text{var}}\left\{T_n^E\right\}}{\hat{\text{E}}\left\{T_n^E\right\}^3},$$

where the hatted, sample statistics are computed across the 8192 realizations, and n is the largest value such that all of the realizations produce at least n events within a time of t_{run}. For weak noise, the above empirical quantity seems to converge much faster than the more obvious empirical estimate of $\frac{1}{t_{run}}\hat{\text{var}}\left\{N_{t_{run}}^E\right\}$.

For the first parameter set ($\epsilon = 1, \rho = 0, \delta = 0$), note that $\Lambda = \text{e}^{-4\pi\epsilon} \approx 3.4 \times 10^{-6}$ is quite small and so the theoretical upper and lower bounds and the numerical approximation offer nearly identical results. We focus on the accuracy of those predictions in comparison with simulations as a function of the noise strength D_{in} and the event subset size $|E| = w$. For smaller event intervals, the unruliness phenomenon arises at weaker noise strengths (see (3.41) for the theoretical effective scaling of D_{in} by w^2). Since the

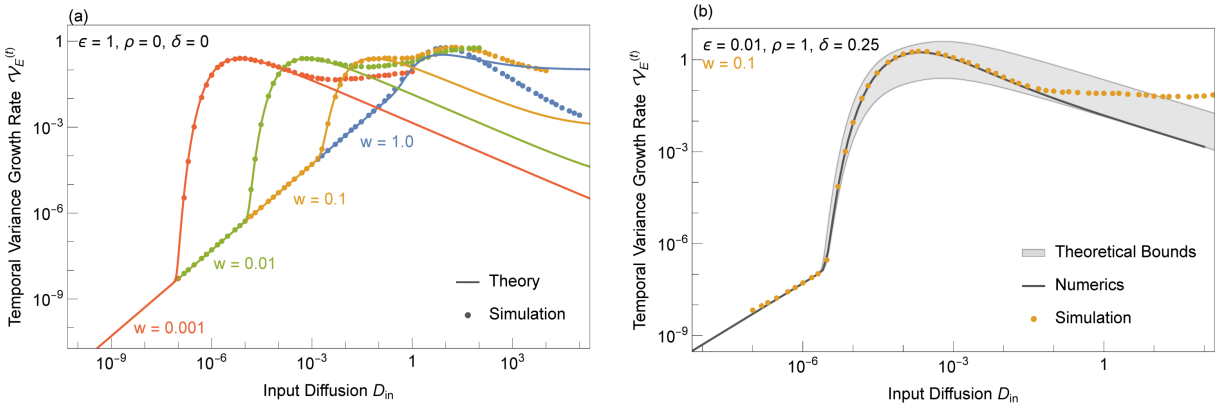

Figure 3.10. **Comparison of theory, numerics, and simulation for the Hopf Oscillator.** (a): The TVGR for the Hopf oscillator as a function of the input noise strength D_{in} for a variety of event subset sizes w and $\epsilon = 1$, $\rho = 0$, and $\delta = 0$. Since $\Lambda = \mathrm{e}^{-4\pi\epsilon} \approx 3.4 \times 10^{-6}$ is so small in this case, the theoretical upper bound, the lower bound, and the numerical approximation are indistinguishable at this scale. Only the theoretical lower bounds (solid curves) are plotted in comparison with the simulation results (points). (b): The TVGR for $w = 0.01$ and $\epsilon = 0.01$, $\rho = 1$, and $\delta = 0.25$. The shaded region shows the range of the theoretical estimates, the solid line the numerical approximation and the points the simulation results.

theory is accurate in the limit of weak noise, we expect that the theory will better capture the unruliness in the graph of $\mathcal{V}_E^{(t)}(D_{in}; w)$ for smaller w. This is indeed the case (Figure 3.10a); for $w = 0.001$ and $w = 0.01$, the theory quantitatively recovers the nonlinear rise and local maximum in $\mathcal{V}_E^{(t)}(D_{in}; w)$, which the standard phase reduction can not capture even qualitatively. In all cases, the Markov renewal theory offers accuracy at noise levels orders of magnitude greater than where the linear estimate $\mathcal{V}_{E,linear}^{(t)} = cD_{in}$ from the standard phase reduction theory applies. Note that the Euler-Maruyama scheme captures the stochastic forcing with a term with magnitude of order $\sqrt{\Delta t D_{in}}$. For the scheme to be accurate it must be able to resolve the passage through E, and the simulation should

not be trusted when $\sqrt{\Delta t D_{in}} \gtrsim w$ or, equivalently, $D_{in} \gtrsim 10^6 w^2$. In Figure 3.10, we have restricted the simulation results for each w accordingly.

The second parameter set ($\epsilon = 0.01, \rho = 1, \delta = 0.25$) tests the accuracy of the theory when the mixed term in the TVGR is non-zero (since $\delta \neq 0$ and $\rho \neq 0$) and when the lower bound, upper bound, and numerical approximation are distinguishable (since $\Lambda = \mathrm{e}^{-4\pi \cdot 0.01} \approx 0.88$ is significant). Naturally, the numerical estimate will lie between the theoretical upper and lower bound. But we should not necessarily expect the numerical approximation and the simulation to coincide, since the numerical estimate is subject to the first order approximation in the noise strength and isostable coordinate (Section 3.5). Nevertheless, for an event interval width of $w = 0.1$, both the theoretical bound and numerics recover the unruly quality of the TVGR found via simulation (Figure 3.10b).

3.7. Extensions

In this section we briefly investigate situations in which alternative types of unruliness appear: a semi-infinite event interval E and oscillators with higher dimensional state spaces. We proceed heuristically, only considering the limiting, quasi-renewal case, where the TVGR is given by (3.49). As we saw for planar oscillators with a finite interval E, the $\mathbf{\Lambda} \neq 0$ case is significantly more complicated but often produces qualitatively similar results.

3.7.1. Semi-infinite E

In Section 3.6, we have only considered the cases where E is a finite subset of S. There is another natural case to consider: E semi-infinite. For a planar oscillator, we might for

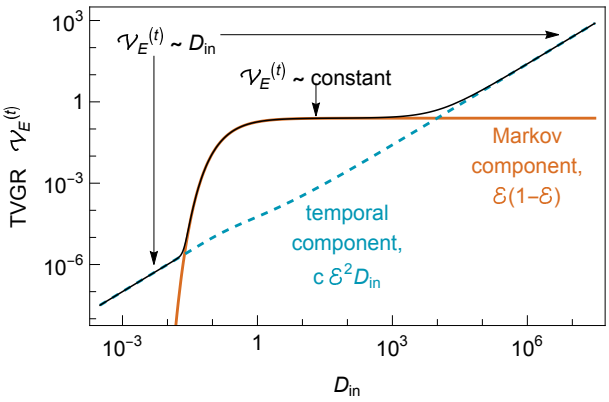

Figure 3.11. **An alternative, monotonic, form of "unruliness" for a semi-infinite interval E.** Shown are the two components of the quasi-renewal TVGR $\mathcal{V}_{E,qr}^{(t)}$ and their total for $c = 10^{-4}$. We make this extreme choice and c for illustrative purposes: each component is dominant for some range of D_{in}, and thus this plot shows all of the possible trends that can appear in $\mathcal{V}_{E,qr}^{(t)}$. $c\mathcal{E}^2 D_{in}$ will always be dominant for small and very large noise, producing a linear trend. But if c is relatively small, $\mathcal{E}(1-\mathcal{E})$ will then produce the characteristically strong nonlinear rise for moderate noise strengths. It is possible that $\mathcal{E}(1-\mathcal{E})$ then contributes to a range over moderately large input noise strengths where $\mathcal{V}_E^{(t)}$ is roughly constant.

example take $E = (x_{min}, \infty)$, so that events occur when the crossing position is above a threshold x_{min}. This is perhaps the case most relevant to neural oscillators, where only a sufficiently large voltage peak may be classified as an "output event".

In the quasi-renewal limit, Λ and x_{min} are jointly taken to 0 such that $x_{min} \propto \sqrt{\Gamma_{ss}} \propto \sqrt{-\frac{1}{\log \Lambda}}$. Then, as was the case with the finite interval E (cf. Section 3.6.2),

$$\mathcal{E} = \frac{1}{2}\left[1 - \text{erf}\left(\frac{x_{min}}{2\sqrt{\Gamma_{ss} D_{in}}}\right)\right]$$

is left invariant, $x_E \to 0$, and $\mathcal{E}_m \to \mathcal{E}^2$ so that $\mathcal{V}_{E,qr}^{(t)}(D_{in}; c) = \mathcal{E}(1-\mathcal{E}) + cD_{in}\mathcal{E}^2$. The proportionality constant between $x_{\min}$ and $\sqrt{\Gamma_{ss}}$ can be absorbed into D_{in}, so that $\mathcal{E} = \frac{1}{2}\left[1 + \text{erf}\left(\frac{1}{2\sqrt{D_{in}}}\right)\right]$, and we find that $\mathcal{E}$ is monotonically decreasing with D_{in}. But, unlike

the situation with a finite interval E, $\mathcal{E}(1-\mathcal{E})$ is monotonic as well (Figure 3.11, orange curve). Thus, overall $\mathcal{V}_E^{(t)}$ will not be non-monotonic. Instead, we find an alternative "unruly" quality for semi-infinite intervals E: a linear trend for low D_{in}, a substantial nonlinear rise for moderate D_{in}, and a monotonically increasing return to a linear trend for large D_{in} (Figure 3.11, black curve). Recall that, in contrast, for finite E the TVGR $\mathcal{V}_E^{(t)}$ is asymptotically constant for large D_{in}.

3.7.2. Oscillators of Dimension $d+1>2$

Considering the general analysis in Section 3.5, perhaps the most obvious generalization is an extension to higher-dimensional oscillators. This becomes a challenge immediately: the probability distributions of interest, though they might be Gaussian (Section 3.5.2), are all multi-dimensional. For an arbitrary choice of the event subset E, the integrals $\mathcal{E}_m$ (3.29) and x_E (3.30) will not in general be found exactly or easily bounded. In this section, we briefly consider two choices for E for which the calculation in the $\Lambda \to 0$ limit is tractable. They serve as prototypical examples and may give insights into the general unruliness phenomena in higher dimensions.

3.7.2.1. Semi-infinite E. The discussion of Section 3.7.1 applies directly to higher-dimensional oscillators when one takes E to be a semi-infinite region bounded by a hyperplane: $E = \left\{ \vec{\psi} \middle| \vec{a}^T \vec{\psi} > x_{min} \right\}$. In that case, the dimensions orthogonal to the vector $\vec{a}$ can be "integrated out" reducing the analysis to the one dimensional case. And, since all probability densities are Gaussian, the marginal distributions along the coordinate y defined by $y = \vec{a}^T \vec{\psi}$ are Gaussian as well. We must note, however, that this analysis

only covers the case where the boundary of E is a hyperplane in the phase-isostable coordinates. This may not be a natural choice for every oscillator model. Even so, the high-dimensional neural mixed-mode oscillator (Figure 3.1b) has a linear trend for large D_{in} (as is predicted here for a semi-infinite E, cf. Figure 3.11). In that model, any voltage peak large enough qualifies as a "spike" event, i.e. E is bounded by a hyperplane orthogonal to the voltage axis, which, in general, is not linear in phase-isostable coordinates. We speculate that the semi-infinite quality, rather than the specific choice of the boundary of E is responsible for the large-noise linear trend.

3.7.2.2. Finite E with $0 \in E$. As the prototypical example for finite E, take E to be an ellipsoid centered at $\vec{\psi} = 0$, shaped like the level-sets of the steady-state probability density $\mathcal{N}(0, 2D_{in}\mathbf{\Gamma}_{ss})$ on the d-dimensional section S. In this case, due to the symmetry in E, the event center of mass $x_E = 0$. In the quasi-renewal limit, we take Λ and the size of E to 0 jointly so that $\mathcal{E}$ is fixed. With the appropriate constants absorbed into D_{in}, we find

$$\mathcal{E} = 1 - \frac{\Gamma\left(\frac{d}{2}, \frac{1}{4D_{in}}\right)}{\Gamma\left(\frac{d}{2}, 0\right)},$$

where Γ is the incomplete gamma function. Since $0 \in E$, the event probability $\mathcal{E}$ approaches 1 for small D_{in}, as before. For large D_{in}, $\mathcal{E}$ goes like $D_{in}^{-\frac{d}{2}}$. And so, in contrast with the planar case, the temporal term $c\mathcal{E}^2 D_{in} \sim cD_{in}^{1-d}$ is actually asymptotically decreasing for state-space dimension $d + 1 > 2$. The result is a yet different form of unruliness, where the TVGR is dominated by $\mathcal{E}(1 - \mathcal{E})$ at moderate and large noise (Figure 3.12). The prominent maximum at moderate noise will appear so long as c is sufficiently small (we set $c = 0.1$ in Figure 3.12). Note that in the large noise limit $\mathcal{E}(1 - \mathcal{E}) \sim \mathcal{E}$

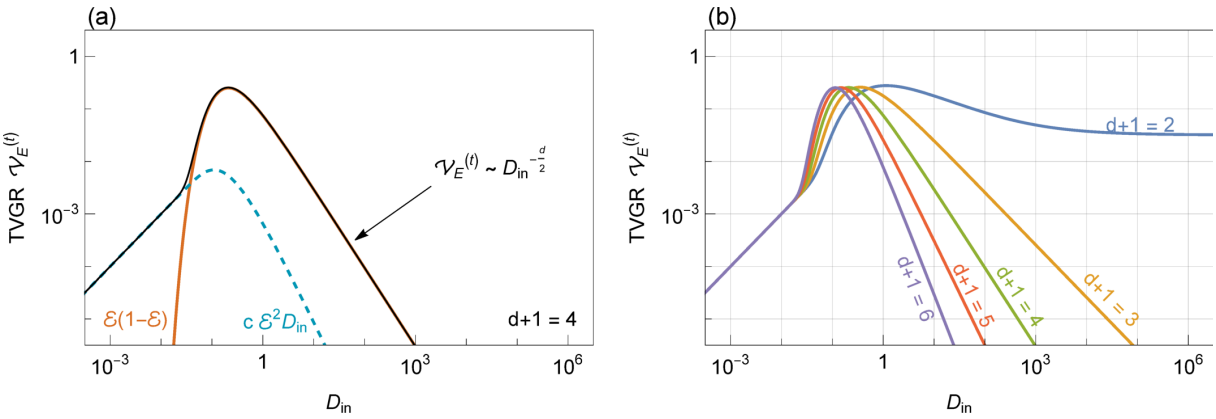

Figure 3.12. **Alternative forms of unruliness in higher dimensions with symmetric E.** Shown is the quasi-renewal TVGR $\mathcal{V}_{E,qr}^{(t)}$ with $c = 0.1$. (a): The two non-zero components of the TVGR for oscillator dimension $d+1 = 4$. For $d+1 > 3$, the TVGR is dominated by $\mathcal{E}(1-\mathcal{E})$ for large D_{in}. (b): The TVGR for various d. For $d+1 > 2$, it is asymptotically decreasing.

goes like $D_{in}^{-\frac{d}{2}}$, which dominates D_{in}^{1-d} when $d+1 > 3$. For all $d+1 > 2$, the TVGR is asymptotically decreasing.

Consider for the moment a similar ellipsoid E, which is, however, asymmetric, such that the center of mass of E is non-zero, $x_E \neq 0$. Since the mixed term in the TVGR is $\vec{b}\cdot x_E \mathcal{E}^2$, the contribution of the asymmetry to $\mathcal{V}_E^{(t)}$ in the large-noise limit is subdominant, $\mathcal{O}(\mathcal{E}^2)$. Therefore we expect that the small- and large- noise asymptotic trends and the alternative form of unruliness found for the unit ball E will be reproduced for mild asymmetry in E. We speculate that it is also reproduced for other reasonable choices of finite E, e.g. those that are convex.

3.8. Discussion

In this work, we have demonstrated that beyond the linear regime predicted by the standard phase-reduction analysis, the event-based, diffusive response of noise-driven

limit-cycle oscillators can be "unruly": with increasing noise strength it can exhibit an enormous nonlinear amplification and a subsequent decrease for yet stronger noise. Such behavior was observed in Chapter 2 in numerical simulations of mixed-mode oscillations. Here we extracted the origin of such unruly behavior by considering the point process that arises from crossings of a Poincare section S on which a subset E is distinguished as event-producing. Since we make use of the phase-isostable representation, the work presented here covers a large class of systems with stable limit cycles. Indeed, for generic planar oscillators and E a finite interval, we have argued that unruliness appears in a finite region of the natural parameter space, and, via a linearization of the dynamics, we have shown this explicitly. We found that unruliness can appear even in a simple, prototypical oscillator driven by additive noise. And, through that example, we confirmed quantitatively the predictions of the point-process approach by direct numerical simulations.

The specification of the event surface E as a proper subset of the Poincare section S is essential to our results; it partitions the points on the section S into events and non-events. Using the linearization of the Poincare map (Section 3.5.2) but excluding the partitioning, there would be no unruliness (see the discussion at the end of Section 3.5.3). What justifies introducing such a partitioning? Recall that our inspiration follows from the mixture of large-amplitude voltage spikes and small-amplitude STOs that appear in neural mixed-mode oscillators (Figure 3.1a). There, only the spikes correspond to the action potentials through which the neurons interact with each other; therefore the spikes are the events of interest. Indeed, in Chapter 2, we found that replacing the standard phase diffusion coefficient D_{phase} in the Fokker-Planck equation (1.2) with the event-based

effective diffusion coefficient $D_{\mathit{eff}} = \frac{1}{2}\mathcal{V}_E^{(t)}$, which is unruly and substantially larger, was essential and sufficient to capture the dynamics of a globally-coupled coupled population of noisy neural mixed-mode oscillators. That replacement enables (1.2) to reproduce even the bistability and non-trivial population dynamics that appear in the full simulations (see Chapter 2). Notably, because the effective diffusion coefficient D_{eff} is a property of noise-response of individual oscillators, we could measure it efficiently via simulations of a single, uncoupled oscillator and then apply the result to the population dynamics.

The designation of events by E can be interpreted more generally as a relevant read-out of an oscillator's dynamics. Examples can also be found outside of neuroscience. For instance, in synthetic biology the repressilator is captured experimentally via large amplitude oscillations in the read-out fluorescence [**18**]. And, in fact, a recent, mathematical study of a repressilator model makes use of a finite Poincare section (equivalently, a finite subset E of a section S) to capture the distribution of times between oscillation peaks [**59**]. In contrast with the original repressilator model, Potapov et. al.'s version supports bistability between a fixed point and a limit cycle via a subcritical Hopf bifurcation. That leads to noise-induced small amplitude oscillations near the fixed point and noisy large amplitude oscillations near the limit cycle, with stochastic transitions between them. While those authors did not consider it, the observed variability in the oscillator output resulting from the transitions could be quantified using the effective diffusion coefficient D_{eff}.

Regular transitions between oscillatory and quiescent phases in the absence of noise are also found in a wide range of natural systems. These "bursters" often arise from the coupling of a bistable oscillator to an additional slow mode [**63**]. When the large

oscillations are identified as events and the system is forced with noise, our results suggest that the event variability of the bursters often could be unruly. Indeed, in a numerical study [**35**], we found such unruliness in a burster arising in an unfolding of the Hopf-zero bifurcation normal form (see e.g. [**27**] for a discussion of the unfolding of this codimension-2 bifurcation). The effect of noise in both bursting and mixed-mode oscillator systems is of considerable theoretical interest [**50, 76, 5**]. It would be an important extension of our work to go beyond the computational results of Chapter 2 and capture theoretically the effective diffusion coefficient in such systems, where, in particular, oscillations of variable amplitude appear already without noise. The analysis is likely to be more complicated than what we have covered here: when the noiseless limit cycle shows both small and large oscillation amplitudes, typical Poincare sections would be pierced by it multiple times. The technical elements of our derivation in Section 3.5 and Appendix B.3 will then be more involved, with a key challenge being a reduced representation of the more complicated Poincare map and first passage time.

In this work, we have focused on the variance in the number of events appearing in a given time interval normalized by that time interval in order to quantify the variability in events produced by oscillators. This "temporal variance growth rate" (TVGR), which is proportional to the effective diffusion coefficient, is just one of a many statistical attributes of a point process. The distributions of the individual time intervals between events, e.g. inter-spike-intervals (ISIs) for neuronal oscillators, are also of broad interest. And, interestingly, for certain oscillators the coefficient of variation of the ISIs can be non-monotonic as a function of input noise strength [**55, 13**]. Our event-based framework may prove useful in that context as well, offering efficient models for complex (e.g.

multi-modal) ISIs distributions and complementing existing, theoretical [**55**] and phenomenological [**13**] approaches. Besides the TVGR and the ISI coefficient of variation, other natural, variance-like, summary statistics have been previously investigated in the literature. Some, like the dispersion of event times and the Fano factor, are related to the TVGR by factors of the mean event interval in the long-time limit (Table 3.1). Thus, the unruly behavior in the TVGR we characterized here may well be related to the so-called coherence resonance and incoherence maximization other authors have found in the Fano factor [**58, 48**]. It would be an interesting extension of our work to characterize the mean event interval and thus connect all of the statistics that appear in Table 3.1 in application to limit-cycle oscillators. We note that in much of the literature (e.g. in [**60, 33**]), the TVGR and the related quantities $\mathcal{V}^{(n)}_{x\mapsto f(x)}$ and $\mathcal{V}^{(t)}_{(x,x',\Delta T)\mapsto w(x,x',\Delta T)}$, which we considered in Sections 3.3 and B.1, appear in the context of central limit theorems. Informally, those theorems indicate that those mean and variance statistics are largely sufficient in describing the event point process over long time scales.

We provided a general decomposition of the TVGR, (3.15), and applied it to a simple toy model (Section 3.4). Notably, the toy model does not include any memory between states at one step and the next, but nevertheless shows unruliness. Thus, the Markov-renewal process (and the memory across steps it offers) is not strictly necessary; a renewal process would be sufficient to generate unruliness in the TVGR. Indeed, the analyses of coherence resonance and incoherence maximization found in the literature often assume renewal dynamics, e.g. by considering only the marginal distribution of single inter-event time intervals. However, the Poincare map dynamics for limit cycle oscillators are generally Markovian, reflecting the deterministic dynamics of the underlying limit

cycle. Moreover, those non-renewal dynamics are not only widespread but may also play a functional role, for instance, in the context of neuroscience, in information processing in the brain [**24, 2, 62**]. One of our contributions in this work is an examination of the TVGR in its entirety within the Markov renewal framework, including the effect of correlations between successive states of the process. In our analysis of planar oscillators, we carefully bound the size of those effects (see (3.51) and SI Section B.4) and find that the correlations contribute positively towards the non-monotonic, Markov-only component of the TVGR (see (3.51)), and therefore may enhance its unruly quality.

The formula we derive for the TVGR (3.15) is exact, but we apply it to limit cycle oscillators in Section 3.5 and the corresponding Appendix B.3 via a linearization of the dynamics. Even though the linearization is an approximation, we anticipate that generically there will be a degree of structural stability as nonlinear effects are "turned on". In particular, we expect that the event probability $\mathcal{E}$ continues to be a monotonically decreasing function of the noise strength D_{in} (as in Figure 3.4a,b.i), leading to the non-monotonicity in the Markov-only component and the TVGR overall. And, so, we anticipate that the unruliness we have discussed here will appear in many (weakly) nonlinear oscillator models.

In real, nonlinear systems the events of interest and the non-events may well be produced in different regions of phase space, which may have very different dynamics. The partitioning of S by E is then a first step in taking this qualitative difference into account. The boundary ∂E constitutes a dynamical separatrix of some type (see [**65, 31, 6**] for specific examples and analyses). In general, we would not expect the Poincare map to vary smoothly across such a boundary, and the linearization we have considered in this

work fails to be an accurate representation. It would be an interesting extension of this work to consider a piece-wise linear approximation to the Poincare map that is segmented at ∂E. That would be analogous to piece-wise linear approximations to the differential equations of nonlinear systems that have been used to study, e.g., coherence and stochastic resonance [**47**].

Additionally, one may consider a relaxation of the hard threshold in the partitioning of events and non-events. In the broader context of Markov renewal reward processes that we discuss in Section 3.3, there is no particular reason to limit an asymptotic variance growth rate calculation to rewards given by the indicator function 1_E. In this view, 1_E is just one of many observation functions on the Poincare section that can be used to probe the nonlinear, stochastic dynamics of the oscillator. We would expect the unruly behavior that we have investigated in this work to be reflected in some way also in the asymptotic variance growth rates for other observation functions.

Finally, we return to our original motivation: a reduced model whose event statistics (Table 3.1) agree with those of the limit-cycle oscillator in the long-time limit. In considering the events, we eschewed the phase variable in favor of a point process. Notably, extending the phase reduction by considering isostable variables, (3.20)-(3.21), does not offer corrections to the phase dynamics at first order in the noise strength (compare (3.2) and (B.13)). In contrast, since the event subset E distinguishes different amplitudes, the phase-isostable representation along with E *does* offer an enhanced approximation of the event point process statistics even at linear order. Our work offers one example where first identifying within the dynamics a non-renewal point process and then capturing the point process as a reduced Markov renewal-reward process proved beneficial. There may

be many other circumstances in which complex noise-driven nonlinear dynamics can be captured effectively by Markov-renewal point processes.

References

[1] Aminzare, Z., Holmes, P., and Srivastava, V. On Phase Reduction and Time Period of Noisy Oscillators. In *2019 IEEE 58th Conference on Decision and Control (CDC)* (Nice, France, Dec. 2019), IEEE, pp. 4717–4722.

[2] Avila-Akerberg, O., and Chacron, M. J. Nonrenewal spike train statistics: causes and functional consequences on neural coding. *Experimental Brain Research 210*, 3-4 (May 2011), 353–371.

[3] Bathellier, B., Lagier, S., Faure, P., and Lledo, P. M. Circuit properties generating gamma oscillations in a network model of the olfactory bulb. *J. Neurophysiol. 95*, 4 (April 2006), 2678–2691.

[4] Bender, C. M., and Orszag, S. A. *Advanced Mathematical Methods for Scientists and Engineers I*. Springer New York, New York, NY, 1999.

[5] Berglund, N., Gentz, B., and Kuehn, C. From Random Poincaré Maps to Stochastic Mixed-Mode-Oscillation Patterns. *Journal of Dynamics and Differential Equations 27*, 1 (Mar. 2015), 83–136.

[6] Berglund, N., and Landon, D. Mixed-mode oscillations and interspike interval statistics in the stochastic FitzHugh–Nagumo model. *Nonlinearity 25*, 8 (July 2012), 2303–2335.

[7] Brea, J. N., Kay, L. M., and Kopell, N. J. Biophysical model for gamma rhythms in the olfactory bulb via subthreshold oscillations. *Proc Natl Acad Sci U S A 106*, 51 (Dec 2009), 21954–21959.

[8] Brown, S., Joseph, J., and Stopfer, M. Encoding a temporally structured stimulus with a temporally structured neural representation. *Nat Neurosci 8*, 11 (Nov 2005), 1568–1576.

[9] Buzsaki, G., and Schomburg, E. W. What does gamma coherence tell us about inter-regional neural communication? *Nature Neuroscience 18*, 4 (Apr. 2015), 484–489.

[10] Buzsaki, G., and Wang, X.-J. Mechanisms of gamma oscillations. *Annual Review of Neuroscience, Vol 35 35* (2012), 203–225.

[11] Castejon, O., and Guillamon, A. Phase-Amplitude Response Functions for Transient-State Stimuli. *Journal of Mathematical Neuroscience 3* (Aug. 2013), 13.

[12] Cox, D. R. *The statistical analysis of series of events*. Methuen, London, 1966.

[13] Craft, M., and Ly, C. The Effects of Background Noise on a Biophysical Model of Olfactory Bulb Mitral Cells. *Bulletin of Mathematical Biology 84*, 10 (Oct. 2022), 107.

[14] Desmaisons, D., Vincent, J. D., and Lledo, P. M. Control of action potential timing by intrinsic subthreshold oscillations in olfactory bulb output neurons. *J. Neurosci. 19*, 24 (December 1999), 10727–10737.

[15] Desroches, M., Guckenheimer, J., Krauskopf, B., Kuehn, C., Osinga, H., and Wechselberger, M. Mixed-mode oscillations with multiple time scales. *SIAM Review 54* (2012), 209.

[16] Dörfler, F., and Bullo, F. Synchronization in complex networks of phase oscillators: A survey. *Automatica 50*, 6 (June 2014), 1539–1564.

[17] Eaton, M. L. *Multivariate Statistics: A Vector Space Approach*. Institute of Mathematical Statistics, 2007.

[18] Elowitz, M. B., and Leibler, S. A synthetic oscillatory network of transcriptional regulators. *Nature 403*, 6767 (Jan. 2000), 335–338.

[19] Ermentrout, B. Complex dynamics in winner-take-all neural nets with slow inhibition. *Neural Netw. 5*, 3 (August 1992), 415–431.

[20] Ermentrout, B., and Terman, D. H. *Mathematical foundations of neuroscience*. No. v. 35 in Interdisciplinary applied mathematics. Springer, New York, 2010.

[21] Ermentrout, B., and Wechselberger, M. Canards, clusters, and synchronization in a weakly coupled interneuron model. *Siam Journal on Applied Dynamical Systems 8*, 1 (2009), 253–278.

[22] Ermentrout, G. B., and Terman, D. H. *Foundations of Mathematical Neuroscience*. 2010.

[23] Ernst, U., Pawelzik, K., and Geisel, T. Delay-induced multistable synchronization of biological oscillators. *Phys. Rev. E 57*, 2 (February 1998), 2150–2162.

[24] Farkhooi, F., Strube-Bloss, M. F., and Nawrot, M. P. Serial correlation in neural spike trains: Experimental evidence, stochastic modeling, and single neuron variability. *Physical Review E 79*, 2 (Feb. 2009), 021905.

[25] Folks, J. L., and Chhikara, R. S. The Inverse Gaussian Distribution and its Statistical Application—A Review. *Journal of the Royal Statistical Society: Series B (Methodological) 40*, 3 (1978), 263–275.

[26] Gabbiani, F., and Cox, S. J. *Mathematics for neuroscientists*, 1. ed.f ed. Elsevier, Acad. Press, Amsterdam, 2010.

[27] Guckenheimer, J. On a codimension two bifurcation. In *Dynamical Systems and Turbulence, Warwick 1980* (Berlin, Heidelberg, 1981), D. Rand and L.-S. Young, Eds., Lecture Notes in Mathematics, Springer, pp. 99–142.

[28] Gunawan, R., and Doyle, F. J. Isochron-Based Phase Response Analysis of Circadian Rhythms. *Biophysical Journal 91*, 6 (Sept. 2006), 2131–2141.

[29] Hansel, D., Mato, G., and Meunier, C. Clustering and slow switching in globally coupled phase oscillators. *Phys. Rev. E 48*, 5 (November 1993), 3470–3477.

[30] Hansel, D., Mato, G., and Meunier, C. Synchrony in excitatory neural networks. *Neural Comput. 7*, 2 (March 1995), 307–337.

[31] Hitczenko, P., and Medvedev, G. S. Bursting Oscillations Induced by Small Noise. *SIAM Journal on Applied Mathematics 69*, 5 (2009), 1359–1392. Publisher: Society for Industrial and Applied Mathematics.

[32] Hitczenko, P., and Medvedev, G. S. The Poincare map of randomly perturbed periodic motion. *Journal of Nonlinear Science 23*, 5 (Oct. 2013), 835–861.

[33] Jones, G. L. On the Markov chain central limit theorem. *Probability Surveys 1*, 0 (2004), 299–320.

[34] Kao, E. P. C. *An Introduction to Stochastic Processes*. Courier Dover Publications, Dec. 2019.

[35] Karamchandani, A., and Riecke, H. *unpublished*.

[36] Karamchandani, A. J., Graham, J. N., and Riecke, H. Pulse-coupled mixed-mode oscillators: Cluster states and extreme noise sensitivity. *Chaos: An Interdisciplinary Journal of Nonlinear Science 28*, 4 (Apr. 2018), 043115.

[37] Kawamura, Y., Nakao, H., Arai, K., Kori, H., and Kuramoto, Y. Phase synchronization between collective rhythms of globally coupled oscillator groups: Noisy identical case. *Chaos 20*, 4 (December 2010), 043109.

[38] Kay, L. M., and Beshel, J. A beta oscillation network in the rat olfactory system during a 2-alternative choice odor discrimination task. *J Neurophysiol 104*, 2 (Jun 2010), 829–839.

[39] Kilpatrick, Z. P., and Ermentrout, B. Sparse gamma rhythms arising through clustering in adapting neuronal networks. *PLoS Comput Biol 7*, 11 (Nov 2011), e1002281.

[40] Kilpatrick, Z. P., and Ermentrout, B. Sparse Gamma Rhythms Arising through Clustering in Adapting Neuronal Networks. *PLoS Computational Biology 7*, 11 (Nov. 2011), e1002281.

[41] Kuramoto, Y. *Chemical Oscillations, Waves, and Turbulence*, vol. 19 of *Springer Series in Synergetics*. Springer Berlin Heidelberg, Berlin, Heidelberg, 1984.

[42] Kvalheim, M. D., and Revzen, S. Existence and uniqueness of global Koopman eigenfunctions for stable fixed points and periodic orbits. *Physica D: Nonlinear Phenomena 425* (Nov. 2021), 132959.

[43] Ladenbauer, J., Augustin, M., Shiau, L., and Obermayer, K. Impact of adaptation currents on synchronization of coupled exponential integrate-and-fire neurons. *PLoS Comput Biol 8*, 4 (2012), e1002478.

[44] Li, G., and Cleland, T. A. A coupled-oscillator model of olfactory bulb gamma oscillations. *PLoS computational biology 13* (Nov. 2017), e1005760.

[45] Lin, K. K., Wedgwood, K. C. A., Coombes, S., and Young, L.-S. Limitations of perturbative techniques in the analysis of rhythms and oscillations. *Journal of Mathematical Biology 66*, 1-2 (Jan. 2013), 139–161.

[46] Lin, K. K., and Young, L.-S. Shear-induced chaos. *Nonlinearity 21*, 5 (Mar. 2008), 899–922.

[47] Lindner, B., and Schimansky-Geier, L. Coherence and stochastic resonance in a two-state system. *Physical Review E 61*, 6 (June 2000), 6103–6110.

[48] Lindner, B., Schimansky-Geier, L., and Longtin, A. Maximizing spike train coherence or incoherence in the leaky integrate-and-fire model. *Physical Review E 66*, 3 (Sept. 2002), 031916.

[49] Liu, C., Weaver, D. R., Strogatz, S. H., and Reppert, S. M. Cellular construction of a circadian clock: Period determination in the suprachiasmatic nuclei. *Cell 91*, 6 (December 1997), 855–860.

[50] Longtin, A. Autonomous stochastic resonance in bursting neurons. *Physical Review E 55*, 1 (Jan. 1997), 868–876.

[51] Meng, J. H., and Riecke, H. Synchronization by uncorrelated noise: interacting rhythms in interconnected oscillator networks. *Scientific Reports 8* (May 2018), 6949.

[52] Michaels, D. C., Matyas, E. P., and Jalife, J. Mechanisms of sinoatrial pacemaker synchronization - a new hypothesis. *Circulation Research 61*, 5 (Nov. 1987), 704–714.

[53] Moca, V. V., Nikolic, D., Singer, W., and Muresan, R. C. Membrane resonance enables stable and robust gamma oscillations. *Cereb Cortex* (Oct 2012).

[54] Monga, B., and Moehlis, J. Augmented phase reduction for periodic orbits near a homoclinic bifurcation and for relaxation oscillators. *Nonlinear Theory and Its Applications, IEICE 12*, 1 (2021), 103–116.

[55] Nesse, W. H., Negro, C. A. D., and Bressloff, P. C. Oscillation Regularity in Noise-Driven Excitable Systems with Multi-Time-Scale Adaptation. *Physical Review Letters 101*, 8 (Aug. 2008), 088101. Publisher: American Physical Society.

[56] Neville, K. R., and Haberly, L. B. Beta and gamma oscillations in the olfactory system of the urethane-anesthetized rat. *J. Neurophysiol. 90*, 6 (December 2003), 3921–3930.

[57] Osinski, B. L., and Kay, L. M. Granule cell excitability regulates gamma and beta oscillations in a model of the olfactory bulb dendrodendritic microcircuit. *J. Neurophys. 116*, 2 (2016), 522–539.

[58] Pikovsky, A. S., and Kurths, J. Coherence Resonance in a Noise-Driven Excitable System. *Physical Review Letters 78*, 5 (Feb. 1997), 775–778.

[59] Potapov, I., Zhurov, B., and Volkov, E. Multi-stable dynamics of the non-adiabatic repressilator. *Journal of The Royal Society Interface 12*, 104 (Mar. 2015), 20141315.

[60] Pyke, R., and Schaufele, R. Limit Theorems for Markov Renewal Processes. *The Annals of Mathematical Statistics 35*, 4 (1964), 1746–1764.

[61] Pérez-Cervera, A., Lindner, B., and Thomas, P. J. Isostables for Stochastic Oscillators. *Physical Review Letters 127*, 25 (Dec. 2021), 254101. Publisher: American Physical Society.

[62] Ramlow, L., and Lindner, B. Interspike interval correlations in neuron models with adaptation and correlated noise. *PLoS computational biology 17*, 8 (Aug. 2021), e1009261.

[63] Rinzel, J. A Formal Classification of Bursting Mechanisms in Excitable Systems. In *Mathematical Topics in Population Biology, Morphogenesis and Neurosciences: Proceedings of an International Symposium held in Kyoto, November 10–15, 1985*, E. Teramoto and M. Yumaguti, Eds., Lecture Notes in Biomathematics. Springer, Berlin, Heidelberg, 1987, pp. 267–281.

[64] Rodrigues, F. A., Peron, T. K. D. M., Ji, P., and Kurths, J. The Kuramoto model in complex networks. *Physics Reports-Review Section of Physics Letters 610* (Jan. 2016), 1–98.

[65] Rowat, P. Interspike interval statistics in the Stochastic Hodgkin-Huxley model: coexistence of gamma frequency bursts and highly irregular firing. *Neural Computation 19*, 5 (May 2007), 1215–1250.

[66] Schultheiss, N. W., and Butera, R. J. *Phase Response Curves in Neuroscience*. Springer, 2012.

[67] Schultheiss, N. W., Prinz, A. A., and Butera, R. J., Eds. *Phase Response Curves in Neuroscience*. Springer New York, New York, NY, 2012.

[68] Schwabedal, J. T. C., and Pikovsky, A. Phase Description of Stochastic Oscillations. *Physical Review Letters 110*, 20 (May 2013), 204102.

[69] Schwalger, T., and Lindner, B. Theory for serial correlations of interevent intervals. *The European Physical Journal Special Topics 187*, 1 (Sept. 2010), 211–221.

[70] Schwemmer, M. A., and Lewis, T. J. The theory of weakly coupled oscillators. *Phase Response Curves In Neuroscience: Theory, Experiment, and Analysis 6* (2012), 3–31.

[71] Schwemmer, M. A., and Lewis, T. J. The Theory of Weakly Coupled Oscillators. In *Phase Response Curves in Neuroscience: Theory, Experiment, and Analysis*, N. W. Schultheiss, A. A. Prinz, and R. J. Butera, Eds., Springer Series in Computational Neuroscience. Springer, New York, NY, 2012, pp. 3–31.

[72] Shinomoto, S., and Kuramoto, Y. Phase Transitions in Active Rotator Systems. *Progress of Theoretical Physics 75*, 5 (May 1986), 1105–1110.

[73] Smith, W. L. Regenerative stochastic processes. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences 232*, 1188 (Oct. 1955), 6–31.

[74] Stiefel, K. M., and Ermentrout, G. B. Neurons as oscillators. *Journal of Neurophysiology 116*, 6 (Dec. 2016), 2950–2960.

[75] Strogatz, S. H., and Mirollo, R. E. Stability of incoherence in a population of coupled oscillators. *Journal of Statistical Physics 63*, 3-4 (May 1991), 613–635.

[76] Su, J., Rubin, J., and Terman, D. Effects of noise on elliptic bursters. *Nonlinearity 17*, 1 (Jan. 2004), 133–157.

[77] Thomas, P. J., and Lindner, B. Asymptotic Phase for Stochastic Oscillators. *Physical Review Letters 113*, 25 (Dec. 2014), 254101.

[78] Wang, X.-J. Neurophysiological and computational principles of cortical rhythms in cognition. *Physiol. Rev. 90*, 3 (July 2010), 1195–1268.

[79] Wedgwood, K. C., Lin, K. K., Thul, R., and Coombes, S. Phase-Amplitude Descriptions of Neural Oscillator Models. *The Journal of Mathematical Neuroscience 3*, 1 (2013), 2.

[80] Whittington, M. A., Traub, R. D., Kopell, N., Ermentrout, B., and Buhl, E. H. Inhibition-based rhythms: experimental and mathematical observations on network dynamics. *Int. J. Psychophysiol. 38* (2000), 315.

[81] Wiesenfeld, K., Colet, P., and Strogatz, S. H. Synchronization transitions in a disordered Josephson series array. *Phys. Rev. Lett. 76*, 3 (January 1996), 404–407.

[82] Wilson, D. Isostable reduction of oscillators with piecewise smooth dynamics and complex Floquet multipliers. *Physical Review E 99*, 2 (Feb. 2019), 022210.

[83] Wilson, D., and Ermentrout, B. An Operational Definition of Phase Characterizes the Transient Response of Perturbed Limit Cycle Oscillators. *SIAM Journal*

on Applied Dynamical Systems 17, 4 (Jan. 2018), 2516–2543. Publisher: Society for Industrial and Applied Mathematics.

[84] Wilson, D., and Moehlis, J. Isostable reduction of periodic orbits. *Physical Review E 94*, 5 (Nov. 2016), 052213.

[85] Winfree, A. T. *The Geometry of Biological Time*. Springer Berlin Heidelberg, Berlin, Heidelberg, 1980.

[86] York, R. A., and Compton, R. C. Quasi-optical power combining using mutually synchronized oscillator arrays. *IEEE Transactions on Microwave Theory and Techniques 39*, 6 (June 1991), 1000–1009.

[87] Yoshimura, K., and Arai, K. Phase reduction of stochastic limit cycle oscillators. *Physical Review Letters 101* (Oct. 2008), 154101.

[88] Zhang, M. A., Wiederhecker, G. S., Manipatruni, S., Barnard, A., McEuen, P., and Lipson, M. Synchronization of micromechanical oscillators using light. *Physical Review Letters 109*, 23 (Dec. 2012), 233906.

[89] Çinlar, E. Markov Renewal Theory: A Survey. *Management Science 21*, 7 (1975), 727–752.

APPENDIX A

Supplementary Information for Chapter 2

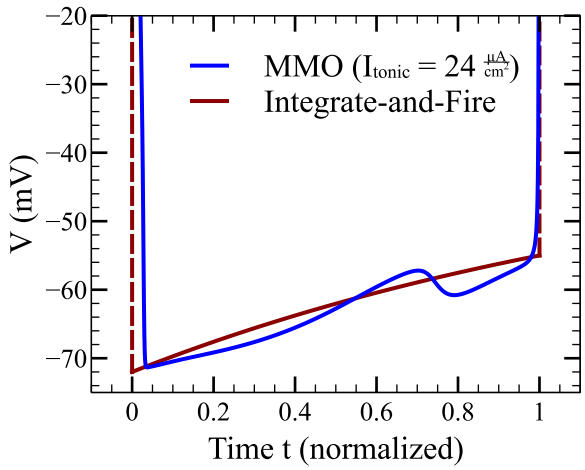

Figure A.1. The integrate-and-fire model chosen for comparison: the parameters of the leaky integrate-and-fire model used in Fig. 2.8 are chosen to match the voltage range of the MMO.

Currents:

Capacitive	Leak	Synaptic
$C_m^{(IF)} = 1 \frac{\mu F}{cm^2}$	$E_{leak}^{(IF)} = -38\text{mV}$ $g_{leak} = \log 2 \frac{\text{s·mS}}{\text{cm}^2} \cdot \text{f}_{mitral}$	$E_{syn} = -70\text{mV}$

Spiking:

$$V_{reset} = -72\text{mV}$$
$$V_{threshold} = -55\text{mV}$$

Table A.1. Integrate-and-fire model parameters.

To contrast the MMOs of the mitral cell model in the context of synaptic interactions, we make use of a leaky integrate-and-fire model:

$$C_m^{(IF)} \frac{dV}{dt} = g_{leak}^{(IF)} \left(E_{leak}^{(IF)} - V \right) + g_{syn} \left(E_{syn} - V \right),$$

and when $V = V_{threshold}$, V is set to V_{reset}. g_{syn} is given by Eq. (2.4) like the mitral cell, and we choose parameters as in Table A.1 so that the firing frequency f_{mitral}, the minimum voltage, and the spiking "threshold voltage" approximately match that of the mitral cell model (Fig. A.1).

APPENDIX B

Supplementary Information for Chapter 3

B.1. Derivation of the Event TVGR Formula

In this appendix section, we offer two derivations of a generalization of (3.11), (B.8) below, wherein the temporal variance growth rate $\mathcal{V}^{(t)}_{(x,x',\Delta T)\mapsto w(x,x',\Delta T)}$ (TVGR) is related to a sequential variance growth rate $\mathcal{V}^{(n)}_{(x\to x')\mapsto f_w(x,x')}$ (SVGR), and a derivation of (3.15), where the TVGR is decomposed into three components. In Section B.1.1, we offer a non-rigorous, plausibility argument that motivates the form of (3.14). In Section B.1.2 we give a more careful derivation that is only valid for Markov chains with discrete state spaces and yet yields the same formula. Between the plausibility argument given in B.1.1, the derivation of limited scope given in B.1.2, and the empirical evidence we provide in Section 3.6.6, we are content to take (3.14) as fact for the further analyses presented in this work.

B.1.1. The TVGR Formula by Rough Analogy

(3.7) shows that the event TVGR is associated with the function $1_E(x)$ on the Markov renewal process. We will consider more generally any "reward" w on the Markov renewal process, which is typically taken to be a (possibly random) function of the current state x, the next state x', and the time interval ΔT between them. The temporal variance growth rate associated with w is therefore

$$\mathcal{V}^{(t)}_{(x\to x',\Delta T)\mapsto w(x,x',\Delta T)} \equiv \lim_{t\to\infty} \frac{1}{t} \mathrm{var} \left\{ \sum_{k=1}^{N_t} w\left(x_k, x_{k+1}, \Delta T_k\right) \right\}. \tag{B.1}$$

We recall that the random variable ΔT_k will in general depend on x_k and x_{k+1}, and, once conditioned on x_k and x_{k+1}, ΔT_k is independent of x_j, $j \neq k, k+1$. The same is true of the reward $w\left(x_k, x_{k+1}, \Delta T_k\right)$ conditioned on x_k and x_{k+1}. We can imagine,

therefore, constructing the Markov process x_k first, independently of the time intervals and reward, and then add to the process the random time intervals ΔT_k and the random reward w. To that end, we take $\Delta t\left(x_k, x_{k+1}\right)$ to be a random function that has the same distribution as ΔT_k when conditioned on x_k and x_{k+1} and rewrite $w\left(x_k, x_{k+1}, \Delta T_k\right)$ as $w\left(x_k, x_{k+1}, \Delta t\left(x_k, x_{k+1}\right)\right)$:

$$\mathcal{V}^{(t)}_{(x\to x',\Delta T)\mapsto w(x,x',\Delta T)} = \lim_{t\to\infty}\frac{1}{t}\mathrm{var}\left\{\sum_{k=1}^{N_t} w\left(x_k, x_{k+1}, \Delta t\left(x_k, x_{k+1}\right)\right)\right\}. \tag{B.2}$$

As compared with (B.1), (B.2) is a step closer to the generalized SVGR (3.10), since the quantity in the sum is a random function of the states x_j alone. This will be helpful in the following, where we find a formula for $\mathcal{V}^{(t)}_{(x,x',\Delta T)\mapsto w(x,x',\Delta T)}$ in terms of a SVGR.

In Section B.1.2, we offer a derivation of such a formula for the special case of a Markov renewal process on a discrete state space. But here we find the same formula ((B.8), below) in a more general setting by analogy. To set up the analogy, first compare the SVGR of a Markov process x_k with that of an i.i.d. process y_k. For independent samples y_k,

$$\frac{1}{n}\mathrm{var}\left\{\sum_{k=1}^{n} f\left(y_k\right)\right\} = \mathrm{var}\left\{f\left(y\right)\right\}, \tag{B.3}$$

while in general for large n,

$$\frac{1}{n}\mathrm{var}\left\{\sum_{k=1}^{n} f\left(x_k\right)\right\} \sim \mathcal{V}^{(n)}_{x\mapsto f(x)} \neq \mathrm{var}\left\{f\left(x\right)\right\}. \tag{B.4}$$

So the SVGR $\mathcal{V}^{(n)}_{x\mapsto f(x)}$ can be thought of as the asymptotic, effective variance of $f\left(x\right)$ that takes into account the fact that x_k and x_l are interdependent. Now, in the analogy, the

SVGR (B.4) is to the i.i.d variance (B.3) as the TVGR $\mathcal{V}^{(t)}_{(x,x',\Delta T)\mapsto w(x,x',\Delta T)}$ for w is to a TVGR $\mathcal{V}^{(t)}_{\Delta T\mapsto w(\Delta T)}$ for a *renewal-reward* process:

$$\mathcal{V}^{(n)}_{x\mapsto f(x)} \ : \ \mathrm{var}\left\{f\left(y\right)\right\} \ :: \ \mathcal{V}^{(t)}_{(x,x',\Delta T)\mapsto w(x,x',\Delta T)} \ : \ \mathcal{V}^{(t)}_{\Delta T\mapsto w(\Delta T)}. \tag{B.5}$$

As compared with a Markov renewal-reward process, a renewal-reward process has no state space and the time intervals are i.i.d. The reward can only be a random function of the time interval, and therefore the corresponding temporal variance growth rate is

$$\mathcal{V}^{(t)}_{\Delta T\mapsto w(\Delta T)} \equiv \lim_{t\to\infty}\frac{1}{t}\mathrm{var}\left\{\sum_{k=1}^{N_t} w\left(\Delta T_k\right)\right\}. \tag{B.6}$$

Renewal-reward theory [**73**] offers the required formula for (B.6),

$$\mathcal{V}^{(t)}_{\Delta T\mapsto w(\Delta T)} = \frac{1}{\mathrm{E}\left\{\Delta T\right\}}\mathrm{var}\left\{w\left(\Delta T\right) - \frac{\mathrm{E}\left\{w\left(\Delta T\right)\right\}}{\mathrm{E}\left\{\Delta T\right\}}\Delta T\right\}. \tag{B.7}$$

Note that like the i.i.d variance $\mathrm{var}\left\{f\left(y\right)\right\}$ and in contrast with the function variance $\mathcal{V}^{(n)}_{x\mapsto f(x)}$ (cf. (3.9)), (B.7) only includes the variance evaluated over one step of the process, reflecting the fact that the steps in the renewal-reward process are independent.

To complete the analogy (B.5) and find $\mathcal{V}^{(t)}_{(x,x',\Delta T)\mapsto w(x,x',\Delta T)}$, we must replace the variance var in (B.7) with its asymptotic, effective counterpart $\mathcal{V}^{(n)}$ that takes into account the non-renewal quality of Markov renewal processes (compare (B.3) and (B.4)). We also generalize the function $w\left(\Delta T\right)$ as $w\left(x,x',\Delta t\left(x,x'\right)\right)$, concluding that

$$\mathcal{V}^{(t)}_{(x,x',\Delta T)\mapsto w(x,x',\Delta T)} = \frac{1}{\mathrm{E}\left\{\Delta T\right\}}\mathcal{V}^{(n)}_{(x\to x')\mapsto f_w(x,x')}, \tag{B.8}$$

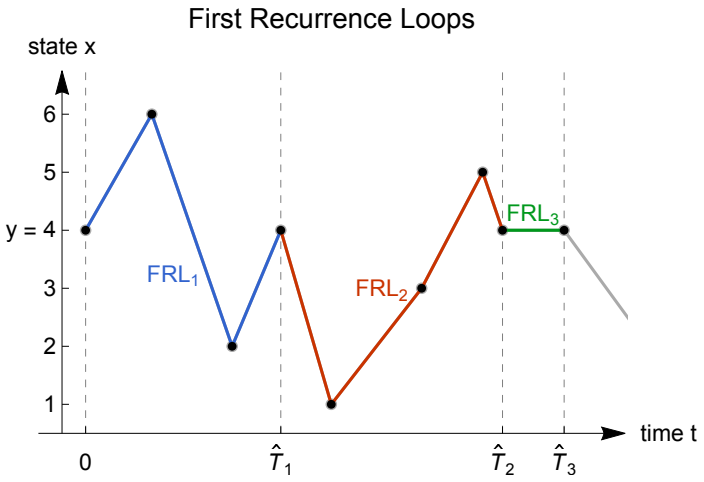

Figure B.1. A schematic depiction of Cinlar's analysis for discrete-state Markov-renewal processes [**89**]. The "first recurrence loops" (FRLs) each with initial and final state y are distinguished by color. The time durations $\Delta\hat{T}_l = \hat{T}_{l+1} - \hat{T}_l$ of the FRLs are independent and identically distributed; the visits to y form the renewal process $\hat{T}_l$.

where

$$f_w(x, x') \equiv w(x, x', \Delta t(x, x')) - \frac{\mathrm{E}\{w(x, x', \Delta T)\}}{\mathrm{E}\{\Delta T\}} \Delta t(x, x').$$

Note that the left hand side of (B.8) is the TVGR-like quantity, defined in (B.1) with respect to the time t, while the right hand side includes a function variance, defined via (3.8) and (3.10) with respect to the discrete steps n. We make use of (B.8) and the formula for function variances, (3.9), to evaluate the reward TVGR $\mathcal{V}^{(t)}_{(x,x',\Delta T)\mapsto w(x,x',\Delta T)}$.

For the case of interest in this work, the reward is the indicator function for events, $w(x, x', \Delta T) = 1_E(x)$, and (B.8) yields (3.11).

B.1.2. The TVGR Formula for Discrete State Spaces

Here we derive (B.8) for Markov renewal processes for a discrete rather than continuous state space. We begin by referring to Cinlar, [**89**], who approaches the Markov-renewal-reward process by connecting it to a renewal-reward process. He considers a partitioning of the sequence x_k for $k \in \{0, 1, \ldots\}$ into what we call "first recurrence loops" (FRLs), which are all the subsequences that start at the initial state $y = x_0$ and end at the same state y without visiting it in between (Figure B.1). Note that this requires the state space to be discrete, since otherwise the return to state y happens with probability 0. Once the first FRL back to y is complete, the next one is a completely independent realization of a FRL starting at y, reflecting the Markov character of the underlying process. Thus, the time $\Delta\hat{T}_l$ elapsed between subsequent visits to y (during the l^{th} FRL, FRL_l) along with the reward $\hat{W}_l$ accumulated during those intervals is a renewal-reward process. Renewal-reward theory then gives [**73**][1]

$$\mathcal{V}^{(t)}_{(x,x',\Delta T)\mapsto w(x,x',\Delta T)} = \frac{1}{\mathrm{E}\left\{\Delta\hat{T}_1\right\}} \mathrm{var}\left\{\hat{W}_1 - \frac{\mathrm{E}\left\{\hat{W}_1\right\}}{\mathrm{E}\left\{\Delta\hat{T}_1\right\}}\Delta\hat{T}_1\right\}, \tag{B.9}$$

where $\Delta\hat{T}_1 = \sum_{k=1}^{\hat{n}_1} \Delta T_k = T_{\hat{n}_1}$ and $\hat{W}_1 = \sum_{k=1}^{\hat{n}_1} w\left(x_k, x_{k+1}, \Delta T_k\right)$, defining $\hat{n}_l$ to be the (random) number of steps in the l^{th} FRL. Since the FRLs parameterized by l are independent, the choice to set $l = 1$ in (B.9) is as good as any. $\mathrm{E}\left\{\Delta\hat{T}_1\right\}$ is the mean first recurrence time for state y. It is proportional to the expected transition time $\mathrm{E}\{\Delta T\}$ across all pairs of states,

$$\mathrm{E}\left\{\Delta\hat{T}_1\right\} = p\,\mathrm{E}\{\Delta T\},$$

[1] In [**89**], the factor of $\mathrm{E}\left\{\Delta\hat{T}_1\right\}^{-1}$ is missing. It appears correctly in Equation 5.2.5 of [**73**].

where the constant of proportionality $p \equiv \frac{1}{\pi(y)} = \mathrm{E}\{\hat{n}_1\}$ is the mean first recurrence time as measured in steps [**89**]. Similarly, $\mathrm{E}\left\{\hat{W}_1\right\} = p\mathrm{E}\{w\}$, so that ${}^{\mathrm{E}\{\hat{W}_1\}}/_{\mathrm{E}\{\Delta\hat{T}_1\}} = {}^{\mathrm{E}\{w\}}/_{\mathrm{E}\{\Delta T\}}$.

We are now equipped to meet our goal - to rewrite $\mathcal{V}^{(t)}_{(x,x',\Delta T)\mapsto w(x,x',\Delta T)}$ as a SVGR that can be computed via (3.9). We first note that the definition of the TVGR for a general reward, (B.1), covers the generalized SVGR case, $\mathcal{V}^{(n)}_{(x\to x')\mapsto f(x,x')}$, considered in Section 3.3.1. Namely, when the Markov renewal process is defined so that the intervals ΔT_k are always 1,

$$\begin{aligned} \mathcal{V}^{(n)}_{(x\to x')\mapsto f(x,x')} &= \mathcal{V}^{(t)}_{(x,x',\Delta T)\mapsto f(x,x')} \\ &= \frac{1}{\mathrm{E}\{\hat{n}_1\}}\mathrm{var}\left\{\hat{F}_1 - \mathrm{E}\{f\}\,\hat{n}_1\right\}, \end{aligned} \tag{B.10}$$

where $\Delta\hat{T}_1$ is replaced by $\hat{n}_1$ and $\hat{F}_1 = \sum_{k=1}^{\hat{n}_1} f(x_k, x_{k+1})$ is determined from f like $\hat{W}_1$ is from w. Now consider the randomly-valued function

$$f_w(x,x') \equiv w(x,x',\Delta t(x,x')) - \frac{\mathrm{E}\{w\}}{\mathrm{E}\{\Delta T\}}\Delta t(x,x'),$$

where $\Delta t(x,x')$ is such that $\Delta t(x_{k-1},x_k)$ has the same distribution as ΔT_k conditioned on x_{k-1} and x_k. Note that

$$\begin{aligned} \hat{W}_1 - \frac{\mathrm{E}\{w\}}{\mathrm{E}\{\Delta T\}}\Delta\hat{T}_1 &= \sum_{k=1}^{\hat{n}_1}\left[w(x_{k-1},x_k,\Delta T_k) - \frac{\mathrm{E}\{w\}}{\mathrm{E}\{\Delta T\}}\Delta T_k\right] \\ \hat{F}_{w1} - \mathrm{E}\{f_w\}\,\hat{n}_1 &= \sum_{k=1}^{\hat{n}_1}\left[f_w(x_{k-1},x_k) - \mathrm{E}\{f_w\}\right]. \end{aligned}$$

And, since $\mathrm{E}\{f_w\} = 0$, $f_w(x_{k-1}, x_k) - \mathrm{E}\{f_w\}$ has the same distribution as $w(x_{k-1}, x_k, \Delta T_k) - \frac{\mathrm{E}\{w\}}{\mathrm{E}\{\Delta T\}}\Delta T_k$ when each is conditioned on x_{k-1} and x_k. In turn, since $\hat{n}_1$ is completely determined by the sequence of states, $x_0, x_1, \ldots$, the sums $\hat{F}_{w1} - \mathrm{E}\{f_w\}\hat{n}_1$ and $\hat{W}_1 - \frac{\mathrm{E}\{w\}}{\mathrm{E}\{\Delta T\}}\Delta\hat{T}_1$ conditioned on $x_0, x_1, \ldots$ have the same mean and variance. It then follows from the law of total variance that $\mathrm{var}\left\{\hat{F}_{w1} - \mathrm{E}\{f_w\}\hat{n}_1\right\} = \mathrm{var}\left\{\hat{W}_1 - \frac{\mathrm{E}\{w\}}{\mathrm{E}\{\Delta T\}}\Delta\hat{T}_1\right\}$. Thus, comparing (B.9) and (B.10) yields (B.8).

As mentioned earlier, the reasoning Cinlar and we present regarding the FRLs - in particular the hatted quantities that appear in (B.9) and subsequent equations - are sensical only for Markov chains with finitely-many and perhaps countably-many states. However, the rough argument by analogy presented in Section B.1.1 does not require a restriction of the state-space and yet yields the same result, suggesting a more general validity.

B.1.3. Decomposition of the TVGR

In this section we expand the temporal variance growth rate, $\mathcal{V}_E^{(t)} = \mathcal{V}_{(x\to x')\mapsto(1_E(x)-\mathcal{E}\Delta t(x,x'))}$, to produce (3.15). From (3.9), we have

$$\mathcal{V}_E^{(t)} = \mathrm{var}\{1_{E,0} - \mathcal{E}\Delta t_0\} + 2\sum_{k=1}^{\infty}\mathrm{cov}\{1_{E,0} - \mathcal{E}\Delta t_0, 1_{E,k} - \mathcal{E}\Delta t_k\},$$

where $1_{E,k} \equiv 1_E(x_k)$ and $\Delta t_k = \Delta t(x_k, x_{k+1})$. Note that since $1_E(x') - \mathcal{E}\Delta t(x, x')$ is a scalar function, we can and have combined the two sums that appear in (3.9). Then by the standard sum properties of variance and covariance,

$$\mathrm{var}\{1_{E,0} - \mathcal{E}\Delta t_0\} = \mathrm{var}\{1_{E,0}\} - 2\mathcal{E}\mathrm{cov}\{1_{E,0}, \Delta t_0\} + \mathcal{E}^2\mathrm{var}\{\Delta t_0\}$$

and

$$\begin{aligned}&\mathrm{cov}\left\{1_{E,0}-\mathcal{E}\Delta t_0, 1_{E,k}-\mathcal{E}\Delta t_k\right\}\\&=\mathrm{cov}\left\{1_{E,0},1_{E,k}\right\}-\mathcal{E}\mathrm{cov}\left\{1_{E,0},\Delta t_k\right\}-\mathcal{E}\mathrm{cov}\left\{\Delta t_0,1_{E,k}\right\}+\mathcal{E}^2\mathrm{cov}\left\{\Delta t_0,\Delta t_k\right\}\end{aligned}$$

The first terms in each of the above two lines give

$$\mathrm{var}\left\{1_{E,0}\right\}+2\sum_{k=1}^{\infty}\mathrm{cov}\left\{1_{E,0},1_{E,k}\right\}=\mathcal{V}^{(n)}_{x\mapsto 1_E(x)},$$

and similarly the last terms give $\mathcal{E}^2\mathcal{V}^{(n)}_{(x\to x')\mapsto\Delta t(x,x')}$. The remaining terms - those with pre-factors of $\mathcal{E}$ - are

$$\begin{aligned}-2\mathcal{E}\left[\mathrm{cov}\left\{1_{E,0},\Delta t_0\right\}+\sum_{k=1}^{\infty}\mathrm{cov}\left\{1_{E,0},\Delta t_k\right\}+\sum_{k=1}^{\infty}\mathrm{cov}\left\{\Delta t_0,1_{E,k}\right\}\right]\\=-2\mathcal{E}\,\mathcal{CV}^{(n)}_{(x\to x')\mapsto 1_E(x'),(x\to x')\mapsto\Delta t(x,x')},\end{aligned}$$

using the definition of $\mathcal{CV}$ given in (3.17). Together, these give the decomposition cited in (3.15).

B.2. Averaged System in Phase-Isostable Coordinates

In order to capture the linearized Poincare map dynamics, we make use of the phase and isostable coordinates [**84**], which we find reflect the oscillator dynamics simply and facilitate a comparison to the standard phase reduction. Here in Section B.2.1 we discuss the phase-isostable coordinate system, and in Section B.2.2 we introduce a form of averaging to further simplify the equations.

B.2.1. Phase-Isostable Coordinates

The phase-isostable coordinates are based on isochrons, equi-phase surfaces that foliate the basin of attraction of the limit cycle. The isochrons can be thought of as level sets of the so-called asymptotic phase function $\Phi(\vec{y})$, which extends the definition of a phase coordinate ϕ on the limit cycle to the basin of attraction. In turn, the isostable coordinates ψ_i, defined via level sets of functions $\Psi_i(\vec{y})$, parameterize the isochrons in such a way that $\Psi_i(\vec{y}) = 0$ identifies the limit cycle. Thus, the $\psi_i = \Psi_i(\vec{y})$ can be thought of as amplitude coordinates. Φ and Ψ_i are defined so that that flows $\vec{y}_0(t)$ of the unperturbed system (i.e. (3.1) with $D_{in} = 0$) satisfy

$$\frac{d}{dt}\Phi(\vec{y}_0(t)) = 1 \tag{B.11}$$

$$\frac{d}{dt}\Psi_i(\vec{y}_0(t)) = -\kappa_i \Psi_i(\vec{y}_0(t)). \tag{B.12}$$

Note we have implicitly defined time such that the limit cycle has period 1 in the absence of perturbations; the phase ϕ goes through a unit interval in a complete cycle. Note also that when the constants κ_i have positive real parts, the isostable functions relax towards the limit cycle given by $\psi_i = 0$ as $t \to \infty$. This reflects the fact that the limit cycle is stable. In this work, we will limit our consideration to the cases where κ_i are real and positive, in which case the functions Ψ_i are also real-valued. We refer the reader to [**84**] for a detailed introduction to isostable coordinates and to [**82**] for a treatment of the situations where some κ_i are complex. We also note that, since they satisfy the linear (B.12), the isostable functions Ψ_i are eigenfunctions of the Koopman operator. But not all Koopman eigenfunctions produce isostable coordinates. Isostables correspond

in particular to the so-called "principal" eigenfunctions, which are non-degenerate: they have non-zero gradients on the limit cycle. See [**42**] for a rigorous investigation of when the principal eigenfunctions exist and are unique.

The non-degeneracy of the isostables comes into play when we consider the noisy perturbative forcing of the oscillator (3.1). In the change of coordinates from $\vec{y}$ to phase and isostables, the gradients of Φ and Ψ_i gives the responses of the phase and isostables to the forcing:

$$\begin{aligned} d\phi &= dt + \sigma \left(\vec{\nabla}\Phi\left(\vec{y}\right)\right)^T \mathbf{G}\left(\vec{y}\right) d\vec{W} + \mathcal{O}\left(\sigma^2\right) \\ d\psi_i &= -\kappa_i \psi_i \, dt + \sigma \left(\vec{\nabla}\Psi_i\left(\vec{y}\right)\right)^T \mathbf{G}\left(\vec{y}\right) d\vec{W} + \mathcal{O}\left(\sigma^2\right) \quad i = 1 \ldots d \,, \end{aligned}$$

where $\sigma \equiv \sqrt{2D_{in}}$. (3.20) and (3.21) then follow from assumption that the deviation $\vec{\psi} \equiv (\psi_1, \psi_2, \ldots, \psi_d)$ from the limit cycle is $\mathcal{O}(\sigma)$, and, therefore, at first order approximation the gradients can be replaced with their values on the limit cycle. Thus, in (3.20) and (3.21),

$$\begin{aligned} \vec{Z}\left(\phi\right) &= \vec{\nabla}\Phi\left(\vec{y}\right)\Big|_{\vec{y}_{LC}(\phi)} \\ \vec{Y_i}\left(\phi\right) &= \vec{\nabla}\Psi_i\left(\vec{y}\right)\Big|_{\vec{y}_{LC}(\phi)} , \end{aligned}$$

where $\vec{y}_{LC}\left(\phi\right)$ is the point on the limit cycle with phase ϕ. Note that, since we work to first order in σ, the result is the same under either the Ito or Stratonovich interpretation of (3.1).

B.2.2. Averaging

For small σ, we expect that the accumulation of the noisy perturbation is slow in comparison with the fast oscillation of ϕ. Therefore, within the first order of approximation, we use a form of averaging to simplify the ϕ-dependent terms $\vec{Z}(\phi)^T \mathbf{G}(\phi)$ and $\vec{Y}_i^T(\phi)\mathbf{G}_{LC}(\phi)$ that appear in (3.20) and (3.21). The end result of our averaging are the equations

$$d\phi = dt + \sigma \vec{z}_G^T \, d\vec{W}(t) + h.o.t. \tag{B.13}$$

$$d\psi_i = -\kappa_i \psi_i \, dt + \sigma e^{\kappa_i (1-\mathrm{mod}_1 \phi)} \vec{y}_{G,i}^T \, d\vec{W}(t) + h.o.t., \tag{B.14}$$

where $\mathrm{mod}_1(\phi) = \phi - \lfloor \phi \rfloor$, a 1-periodic function, is the fractional part of ϕ. In Section B.2.2.1 we derive expressions for the constant vectors $\vec{z}_G$ and $\vec{y}_{G,i}$, and in Section B.2.2.2 we explain the (admittedly odd) appearance of $e^{-\kappa_i \mathrm{mod}_1(\phi)}$ in (B.14). We consider a limiting case in Section B.2.2.3 and offer a brief discussion of this unusual form of averaging in Section B.2.2.4.

B.2.2.1. Averaged PRC and IRCs. In order to apply averaging, we first rewrite the equations for ψ_i, (3.21), using an integrating factor,

$$d\left(e^{\kappa_i t}\psi_i\right) = \sigma e^{\kappa_i t} \vec{Y}_i(\phi)^T \mathbf{G}_{LC}(\phi) \, d\vec{W}(t) + h.o.t., \tag{B.15}$$

where, as in Section 3.5, the higher order terms ($h.o.t.$'s) are $\mathcal{O}\left(\sigma^2, \sigma\left|\vec{\psi}\right|, \left|\vec{\psi}\right|^2\right)$. Averaging (B.17) is non-trivial, since $e^{\kappa_i t}$ is non-periodic, and we will in fact not average in the usual, long-time-scale sense. We will be interested in the solutions $\phi(t)$ and $\psi_i(t)$ for times $t \in \mathbb{Z}$ that are roughly the times of crossings of the Poincare section. So, instead, by "averaging" we here mean replacing $\vec{Z}(\phi)^T \mathbf{G}_{LC}(\phi)$ and $\vec{Y}_i(\phi)^T \mathbf{G}_{LC}(\phi)$ with the constant

real vector $\vec{z}_G^T$ and a simplified expression $\mathrm{e}^{\kappa_i(1-\mathrm{mod}_1\phi)}\vec{y}_{G,i}^T$ (their respective "averages") in such a way that the distribution of the solutions ϕ and ψ_i remains unchanged at integer values of t.

Consider initial conditions $\phi(0) = 0 + \mathcal{O}\left(\sigma, \left|\vec{\psi}\right|\right)$ and $\psi_i(0) = \psi_{i,0}$, which will be relevant to the Poincare map dynamics. For both the unaveraged and averaged equations, $\phi = t + \mathcal{O}\left(\sigma, \left|\vec{\psi}\right|\right)$. From the unaveraged equations (3.20) and (B.15), we obtain linear, time-varying equations,

$$d\phi = dt + \sigma \vec{Z}(t)^T \mathbf{G}_{LC}(t)\, d\vec{W}(t) + h.o.t. \tag{B.16}$$

$$d\left(\mathrm{e}^{\kappa_i t}\psi_i\right) = \sigma \mathrm{e}^{\kappa_i t}\vec{Y}_i(t)^T \mathbf{G}_{LC}(t)\, d\vec{W}(t) + h.o.t., \tag{B.17}$$

at first order approximation. Similarly, from the averaged equations (B.13) and (B.14),

$$d\phi = dt + \sigma\, \vec{z}_G^T d\vec{W}(t) + h.o.t. \tag{B.18}$$

$$\begin{aligned} d\left(\mathrm{e}^{\kappa_i t}\psi_i\right) &= \sigma \mathrm{e}^{\kappa_i(t+1-\mathrm{mod}_1(t))}\vec{y}_{G,i}^T\, d\vec{W}(t) + h.o.t. \\ &= \sigma \mathrm{e}^{\kappa_i(\lfloor t \rfloor+1)}\vec{y}_{G,i}^T\, d\vec{W}(t) + h.o.t. \end{aligned} \tag{B.19}$$

The solutions of both the unaveraged system (B.16) and (B.17),

$$\phi(t) = \phi_0 + t + \sigma \int_0^t \vec{Z}(s)^T \mathbf{G}_{LC}(s)\, d\vec{W}(s) + h.o.t. \tag{B.20}$$

$$\mathrm{e}^{\kappa_i t}\psi_i(t) = \psi_{i,0} + \sigma \int_0^t \mathrm{e}^{\kappa_i s}\vec{Y}_i(s)^T \mathbf{G}_{LC}(s)\, d\vec{W}(s) + h.o.t., \tag{B.21}$$

and the averaged system (B.18) and (B.19),

$$\phi(t) = \phi_0 + t + \sigma \vec{z}_G^T \vec{W}(t) + h.o.t. \tag{B.22}$$

$$e^{\kappa_i t}\psi_i(t) = \psi_{i,0} + \sigma \int_0^t e^{\kappa_i(\lfloor s \rfloor + 1)} \vec{y}_{G,i}^T d\vec{W}(s) + h.o.t., \tag{B.23}$$

are Gaussian processes at lowest order. Both sets of equations are in agreement already about the means, $\mathrm{E}\{\phi(t)\} = t$ and $\mathrm{E}\{\psi_i(t)\} = e^{-\kappa_i t}\psi_{i,0}$. What remains is to match the variances and covariances, $\mathrm{var}\{\phi(t)\}$, $\mathrm{cov}\{\phi(t), \psi_i(t)\}$, and $\mathrm{cov}\{\psi_i(t), \psi_j(t)\}$, between the two sets of solutions.

It is sufficient to match after one period, i.e. at $t = 1$: since $\mathbf{G}_{LC}(s)$, $\vec{Y}_i(s)$ and $\vec{Z}(s)$ are 1-periodic, their influence on the (co)variances from $t = 0$ to $t = 1$ is similar to that between $t = l$ and $t = l + 1$ for $l \in \mathbb{Z}$. Indeed, the (co)variances at integer values of t only depend on their values at $t = 1$ and a time-dependent prefactor that depends parametrically at most on κ_i. For $\mathrm{var}\{\phi(t)\}$, this is easy to see, since the amount of variance that accumulated in each period is the same:

$$\begin{aligned}
\mathrm{var}\{\phi(t)\} &= \sigma^2 \int_0^t \vec{Z}(s)^T \mathbf{G}_{LC}(s) \mathbf{G}_{LC}(s)^T \vec{Z}(s)\, ds \\
&= \sigma^2 t \int_0^1 \vec{Z}(s)^T \mathbf{G}_{LC}(s) \mathbf{G}_{LC}(s)^T \vec{Z}(s)\, ds \\
&= t\mathrm{var}\{\phi(1)\}, \ (t \in \mathbb{Z}),
\end{aligned}$$

and so the aforementioned prefactor for var $\{\phi(t)\}$ is t. For cov $\{\phi(t), \psi_i(t)\}$, the prefactor turns out to be $\frac{1-e^{-\kappa_i t}}{1-e^{-\kappa_i}}$:

$$\begin{aligned}
\text{cov}\{\phi(t), \psi_i(t)\} &= \sigma^2 e^{-\kappa_i t} \int_0^t e^{\kappa_i s} \vec{Z}(s)^T \mathbf{G}_{LC}(s) \mathbf{G}_{LC}(s)^T \vec{Y}_i(s)\, ds \\
&= \sigma^2 e^{-\kappa_i t} \sum_{l=0}^{t-1} e^{\kappa_i l} \int_0^1 e^{\kappa_i s} \vec{Z}(s)^T \mathbf{G}_{LC}(s) \mathbf{G}_{LC}(s)^T \vec{Y}_i(s)\, ds \\
&= \frac{1-e^{-\kappa_i t}}{1-e^{-\kappa_i}} \text{cov}\{\phi(1), \psi_i(1)\} \ (t \in \mathbb{Z}).
\end{aligned}$$

Similarly, cov $\{\psi_i(t), \psi_j(t)\} = \frac{1-e^{-(\kappa_i+\kappa_j)t}}{1-e^{-(\kappa_i+\kappa_j)}} \text{cov}\{\psi_i(1), \psi_j(1)\}$ for $t \in \mathbb{Z}$. Since the time-dependent prefactors t, $\frac{1-e^{-\kappa_i t}}{1-e^{-\kappa_i}}$ and $\frac{1-e^{-(\kappa_i+\kappa_j)t}}{1-e^{-(\kappa_i+\kappa_j)}}$ are independent of $\mathbf{G}_{LC}(t)$, $\vec{Y}_i(t)$ and $\vec{Z}(t)$, they are unchanged when $\vec{Y}_i(t)^T \mathbf{G}_{LC}(t)$ and $\vec{Z}(t)^T \mathbf{G}_{LC}(t)$ are averaged. Therefore, we need only to match the (co)variances of the unaveraged and averaged solutions at a single integer value of t. We choose $t = 1$.

At $t = 1$, the expression for the averaged ψ_i, (B.23), simplifies to

$$e^{\kappa_i} \psi_i(1) \sim \psi_{i,0} + \frac{\sigma}{\sqrt{\kappa_i}} e^{\kappa_i} \vec{y}_{G,i}^T \vec{W}(s), \tag{B.24}$$

and the (co)variances of interest for averaged solutions, (B.22) and (B.24) are

$$\begin{aligned}
\text{var}\{\phi(1)\} &= \sigma^2 \vec{z}_G^T \vec{z}_G \\
\text{cov}\{\phi(1), \psi_i(1)\} &= \sigma^2 \vec{z}_G^T \vec{y}_{G,i} \\
\text{cov}\{\psi_i(1), \psi_j(1)\} &= \sigma^2 \vec{y}_{G,i}^T \vec{y}_{G,j}.
\end{aligned}$$

Define $\sigma^2\mathbf{\Xi}$ as the covariance matrix for $(\phi(1);\psi_1(1);\psi_2(1);\ldots;\psi_d(1))$ with entries $\sigma^2\xi_{i,j}$ as they follow from the unaveraged solutions, (B.20) and (B.21). Then for the averaged and unaveraged version of the (co)variances to match, we require that

$$\vec{z}_G^T\vec{z}_G = \xi_{1,1} \equiv \int_0^1 \vec{Z}(s)^T \mathbf{G}_{LC}(s)\,\mathbf{G}_{LC}(s)^T \vec{Z}(s)\,ds \tag{B.25}$$

$$\vec{z}_G^T\vec{y}_{G,i} = \xi_{1,i+1} \equiv \mathrm{e}^{-\kappa_i}\int_0^1 \mathrm{e}^{\kappa_i s}\vec{Z}(s)^T \mathbf{G}_{LC}(s)\,\mathbf{G}_{LC}(s)^T \vec{Y}_i(s)\,ds \tag{B.26}$$

$$\vec{y}_{G,i}^T\vec{y}_{G,j} = \xi_{i+1,j+1} \equiv \mathrm{e}^{-(\kappa_i+\kappa_j)}\int_0^1 \mathrm{e}^{(\kappa_i+\kappa_j)s}\vec{Y}_i(s)^T \mathbf{G}_{LC}(s)\,\mathbf{G}_{LC}(s)^T \vec{Y}_j(s)\,ds, \tag{B.27}$$

for $1 \le i,j \le d$. Equivalently, we require that $\mathbf{\Xi}$ can be decomposed as $\mathbf{\Xi} = \mathbf{R}_G^T\mathbf{R}_G$, where

$$\mathbf{R}_G = \left(\begin{array}{ccccc} \vec{z}_G, & \vec{y}_{G,1}, & \vec{y}_{G,2}, & \cdots, & \vec{y}_{G,d} \end{array}\right). \tag{B.28}$$

Since $\mathbf{\Xi}$ is proportional to a covariance matrix, it is positive semi-definite. The decomposition is therefore possible, but it is not unique. $\mathbf{\Xi}$ is invariant under orthogonal transformations $\mathbf{O}$ of $\mathbf{R}_G$, $\mathbf{R}_G \mapsto \mathbf{O}\mathbf{R}_G$. But $\mathbf{O}$ only has the effect of rotating or reflecting the independent components of $d\vec{W}(t)$. Indeed, as can be seen in Section (B.3) below, the temporal variance growth rate only depends on $\mathbf{\Xi}$, not on the specific choice of vectors $\vec{z}_G$ and $\vec{y}_{G,i}$.

B.2.2.2. Limitations of a Constant IRC Average. We now justify the odd inclusion of $\mathrm{e}^{\kappa_i(1-\mathrm{mod}_1\phi)}$ in the average of $\vec{Y}_i(\phi)^T\mathbf{G}_{LC}(\phi)$. Consider an equivalence class of all 1-periodic, vector-valued functions $\vec{Y}_i(\phi)^T\mathbf{G}_{LC}(\phi)$ that produce the same (co)variance statistics for ϕ and ψ_i. In the averaging discussed above, we have chosen the 1-periodic

vector $\mathrm{e}^{\kappa_i(1-\mathrm{mod}_1\phi)}\vec{y}_{G,i}$ as a simple representative of that equivalence class. Would a constant vector $\vec{y}_G^{constant}$ not be a simpler choice? In other words, why can one not replace $\vec{Y}_i(\phi)^T \mathbf{G}_{LC}(\phi)$ with a constant vector $(\vec{y}_G^{constant})^T$, like we do for $\vec{Z}(\phi)^T \mathbf{G}_{LC}(\phi)$?

As it turns out, constant averages can not capture the full range of possible correlations between ϕ and the ψ_i. Consider a planar oscillator where $\vec{Z}(\phi)^T \mathbf{G}_{LC}(\phi) = \vec{z}_G$ and $\vec{Y}_1(\phi)^T \mathbf{G}_{LC}(\phi) = \vec{y}_{G,1}^{constant}$ are constant. The correlation coefficient between $\phi(1)$ and $\psi_1(1)$ is (following (B.25)-(B.27))

$$\begin{aligned}
\rho_{\phi(1),\psi_1(1)} &= \frac{\mathrm{cov}\left\{\phi(1),\psi_1(1)\right\}}{\sqrt{\mathrm{var}\left\{\phi(1)\right\}\mathrm{var}\left\{\psi_1(1)\right\}}} \\
&= \frac{\mathrm{e}^{-\kappa_1}\int_0^1 \mathrm{e}^{\kappa_1 s}ds}{\sqrt{\left(\int_0^1 ds\right)\left(\mathrm{e}^{-2\kappa_1}\int_0^1 \mathrm{e}^{2\kappa_1 s}ds\right)}}\frac{\vec{z}_G^T\vec{y}_{G,1}^{constant}}{\|\vec{z}_G\|\left\|\vec{y}_{G,1}^{constant}\right\|} \\
&= \sqrt{\frac{2}{\kappa_1}\frac{\mathrm{e}^{\kappa_1}-1}{\mathrm{e}^{\kappa_1}+1}}\frac{\vec{z}_G^T\vec{y}_{G,1}^{constant}}{\|\vec{z}_G\|\left\|\vec{y}_{G,1}^{constant}\right\|}.
\end{aligned}$$

The second factor is the cosine of the angle between $\vec{z}_G$ and $\vec{y}_{G,1}^{constant}$ and can take on values between -1 and 1, but the first factor is strictly less than 1 for $\kappa_1 > 0$. The choice of constant averages, $\vec{z}_G$ and $\vec{y}_{G,1}^{constant}$, only accommodates correlation coefficients between $-\sqrt{\frac{2}{\kappa_1}\frac{\mathrm{e}^{\kappa_1}-1}{\mathrm{e}^{\kappa_1}+1}}$ and $\sqrt{\frac{2}{\kappa_1}\frac{\mathrm{e}^{\kappa_1}-1}{\mathrm{e}^{\kappa_1}+1}}$. It would not be a problem if no choice of $\vec{Z}(\phi)^T \mathbf{G}_{LC}(\phi)$ and $\vec{Y}_1(\phi)^T \mathbf{G}_{LC}(\phi)$ yields a correlation coefficient outside of that range: $\vec{z}_G$ and $\vec{y}_{G,1}^{constant}$ could be made to produce the same statistics. But counterexamples are easy to find. Consider the simple scenario where

$$\vec{Z}(\phi)^T \mathbf{G}_{LC}(\phi) = \vec{Y}_1(\phi)^T \mathbf{G}_{LC}(\phi) = \begin{pmatrix} \sin(2\pi\phi) & 0 \end{pmatrix}.$$

The correlation coefficient is

$$\begin{aligned}\rho_{\phi(1),\psi_1(1)} &= \frac{e^{-\kappa_1}\int_0^1 e^{\kappa_1 s}\sin^2(2\pi s)\,ds}{\sqrt{\left(\int_0^1 \sin^2(2\pi\phi)\,ds\right)\left(e^{-2\kappa_1}\int_0^1 e^{2\kappa_1 s}\sin^2(2\pi s)\,ds\right)}} \\ &= \frac{8\pi\sqrt{4\pi^2+\kappa_1^2}}{16\pi^2+\kappa_1^2}\sqrt{\frac{2}{\kappa_1}\frac{e^{\kappa_1}-1}{e^{\kappa_1}+1}}\end{aligned}$$

which exceeds $\sqrt{\frac{2}{\kappa_1}\frac{e^{\kappa_1}-1}{e^{\kappa_1}+1}}$ when κ_1 is between 0 and about 17.8.

Using the averages $\vec{z}_G$ and $e^{\kappa_i(1-\mathrm{mod}_1\phi)}\vec{y}_{G,i}$, the correlation coefficient $\rho_{\phi(1),\psi_i(1)}$ is - by design - just the cosine of the angle between $\vec{z}_G$ and $\vec{y}_{G,i}$. Similarly, $\rho_{\psi_i(1),\psi_j(1)}$ is the cosine of the angle between $\vec{y}_{G,i}$ and $\vec{y}_{G,j}$. The averages $\vec{z}_G$ and $e^{\kappa_i(1-\mathrm{mod}_1\phi)}\vec{y}_{G,i}$ thus cover all possible correlation values and can be made to represent any $\vec{Z}(\phi)^T\mathbf{G}_{LC}(\phi)$ and $\vec{Y}_i(\phi)^T\mathbf{G}_{LC}(\phi)$.

B.2.2.3. Averages in the Quasi-Renewal Limit. The analysis of the temporal variance growth rate is greatly simplified in the $\Lambda \to 0$ ($\kappa_i \to \infty$) limit. Here we discuss the behavior of $\vec{z}_G^T\vec{z}_G$, $\vec{z}_G^T\vec{y}_{G,i}$, and $\vec{y}_{G,i}^T\vec{y}_{G,j}$ (given in (B.25)-(B.27)) in that "quasi-renewal" limit. $\vec{z}_G^T\vec{z}_G$ does not depend on κ_i and is unchanged. $\vec{z}_G^T\vec{y}_{G,i}$ and $\vec{y}_{G,i}^T\vec{y}_{G,j}$ on the other hand both go to 0 as $\kappa_i \to \infty$. The asymptotic form of the integrals in (B.26) and (B.27) can be found via Laplace's method (see e.g. Section 6.4 of [**4**] for a pedagogical treatment). Both $\vec{z}_G^T\vec{y}_{G,i}$ and $\vec{y}_{G,i}^T\vec{y}_{G,j}$ are of the form

$$I(\kappa) = \int_0^1 e^{\kappa(s-1)}g(s)\,ds$$

where $g(s)$ is 1-periodic and κ play the role of κ_i or $\kappa_i+\kappa_j$. Since $e^{\kappa(s-1)}$ is maximized at $s=1$, we replace $g(s)$ with $g(1)=g(0)$ at first order approximation. To the same

order of approximation, since the integrand is exponentially small for $s < 1$, the integral can be taken over s from $-\infty$ to 1, yielding

$$I(\kappa) \sim \left(\int_{-\infty}^{1} \mathrm{e}^{\kappa(s-1)} \right) g(0) = \frac{1}{\kappa} g(0) .$$

For $\vec{z}_G^T \vec{y}_{G,i}$ and $\vec{y}_{G,i}^T \vec{y}_{G,j}$, this becomes (cf. (B.26) and (B.27))

$$\begin{aligned} \vec{z}_G^T \vec{y}_{G,i} &\sim \frac{1}{\kappa_i} \vec{Z}(0)^T \mathbf{G}_{LC}(0) \mathbf{G}_{LC}(0)^T \vec{Y}_i(0) \\ \vec{y}_{G,i}^T \vec{y}_{G,j} &\sim \frac{1}{\kappa_i + \kappa_j} \vec{Y}_i(0)^T \mathbf{G}_{LC}(0) \mathbf{G}_{LC}(0)^T \vec{Y}_j(0) . \end{aligned}$$

In particular, in a planar oscillator, where there is a single vector $\vec{y}_G$ and $\Lambda = \mathrm{e}^{-\kappa}$ is a scalar, $\|\vec{y}_G\|^2$ goes like $-\frac{1}{\log \Lambda}$, a fact the we make use of in Section 3.6.2.

B.2.2.4. Some Comments About Averaging. As we discussed above, averaging in the usual long-time-scale sense is not desirable, since - in deriving the Poincare map dynamics - we integrate (B.13) and (B.14) over an $\mathcal{O}(1)$ time interval. The ϕ-dependent factor $\mathrm{e}^{\kappa_i(1-\mathrm{mod}_1\phi)}$ is required to capture the joint statistics of $\phi(t)$ and $\psi_i(t)$ at $t \sim 1$, and, in some sense, $\vec{z}_G^T$ and $\mathrm{e}^{\kappa_i(1-\mathrm{mod}_1\phi)} \vec{y}_{G,i}^T$ are the simplest PRC and IRCs that can reproduce the statistics from any given oscillator model. This reflects the fact that additional time scales introduced by the appearance of κ_i in the isostable equations, (3.21), remain relevant over the $\mathcal{O}(1)$ time interval. In fact, the initial condition $\phi(0) \sim 0$ also remains relevant. $\vec{y}_{G,i}$ not only varies with κ_i but also depends on our choice of where $\phi = 0$: a redefinition of ϕ by a shift would not leave $\vec{y}_{G,i}$ invariant (see (B.26) and (B.27)). Since we choose $\phi = 0$ to identify the point where the Poincare section intersects the limit cycle, this ultimately means that the averaged IRCs $\vec{y}_{G,i}$ depend the location of that intersection.

The standard phase reduction with the standard averaging, (3.2) and (3.3), follows under the assumption of strong contraction to the limit cycle ($\kappa_i \gg 1$). In that case $\psi_i \sim 0$ and can be ignored. Note that the phase diffusion $D_{phase} = \vec{z}_G^T \vec{z}_G D_{in}$ is invariant under shifts of the phase variable (see (B.25)), and it is thus independent of where the Poincare section intersects the limit cycle.

We must also note that the averaging step we have presented here is not strictly necessary in our derivation of the event temporal variance growth rate, which could follow directly from the linearized, but unaveraged system, (3.20) and (3.21). But, by averaging, we reveal what minimal, extra information about the oscillator beyond the standard phase reduction is required to capture events statistically in the long-time limit. The averaged quantities give an interpretation to the oscillator-dependent parameters that appear in the formula for the TVGR (3.31): $\vec{b}$ roughly measures the overlap between the averaged PRC $\vec{z}_G$ and each of the averaged IRCs $\vec{y}_{G,i}$ (see (3.32)), and c measures the magnitude of the PRC (see (3.33)).

B.3. Temporal Variance Growth Rate for Limit-cycle Oscillators

Here we detail the derivation of the TVGR for limit cycle oscillators, starting from the averaged equations, (B.13) and (B.14). In Section B.3.1, we adapt the derivation given by Hitczenko and Medvedev in [**32**] of the linearized Poincare map and first passage time, (3.23), for the averaged system. We derive the associated joint distributions and statistics for the positions on the Poincare section and the first passage times in Section B.3.2. In Section B.3.3, we use those distributions and statistics to produce an expression for the temporal variance growth rate (3.31).

B.3.1. Linearized Stochastic Poincare Map

As in Hitczenko and Medvedev's work (hereafter referred to as "HM"), our derivation of the first passage time and linearized Poincare map, (3.23), follows from the approximate solution in a moving coordinates representation of the noise-driven limit cycle oscillator. While HM use amplitude coordinates $\vec{\rho}$ that are orthogonal to the limit cycle, we use isostable coordinates that parameterize isochrons, which in general are not orthogonal to the limit cycle. We first contrast the two choices, noting the simplicity of the system in isostable coordinates. In the moving orthogonal coordinates, the system to lowest order is (Equations 2.8-2.9 of HM)

$$d\theta \sim \left[1+\vec{a}\left(\theta\right)^{T}\vec{\rho}\right]dt+\sigma\vec{h}\left(\theta\right)^{T}d\vec{W}\left(t\right) \tag{B.29}$$

$$d\vec{\rho} \sim \mathbf{R}\left(\theta\right)\vec{\rho}+\sigma\mathbf{H}\left(\theta\right)^{T}d\vec{W}\left(t\right), \tag{B.30}$$

while in isostable coordinates with averaging, it is (B.13) and (B.14), or

$$d\phi \sim dt+\sigma\vec{z}_{G}^{T}d\vec{W}\left(t\right) \tag{B.31}$$

$$d\vec{\psi} \sim -\mathbf{K}\vec{\psi}\,dt+\sigma\mathrm{e}^{\mathbf{K}(1-\mathrm{mod}_{1}\phi)}\mathbf{Y}_{G}^{T}d\vec{W}\left(t\right), \tag{B.32}$$

where $\sigma\equiv\sqrt{2D_{in}}$, $\mathbf{K}$ is the diagonal matrix with entries κ_i, and $\mathbf{Y}_G$ is the matrix with columns $\vec{y}_{G,i}$. Note the inclusion of the vector $\vec{a}$ in (B.29) that captures shear in the flow in the vicinity of the limit cycle. Since the shape of the isochrons takes the shear into account, no such term appears in phase equation (B.31) under phase-isostable coordinates. The dynamics of the amplitude coordinates are also simpler. In the absence of perturbative inputs, the isostable coordinates are decoupled, evolving by a constant

diagonal matrix $\mathbf{K}$ (cf. the generally non-diagonal $\mathbf{R}(\theta_t)$ in (B.29) that varies with the phase variable θ). Since we apply averaging beforehand, the noise terms in (B.31) and (B.32) are simplified. We note that over the course of HM's derivation, the averaging calculation presented in Appendix B.2.2, as it applies to the time interval $t \in [0, 1]$, is done implicitly. In the end, the results we obtain by averaging first are equivalent to those found by applying HM's work to the unaveraged system (3.21). Formally, the averaged system, (B.31) and (B.32), which are the starting point of our analysis, are a special case of (B.29) and (B.30). HM study (B.29) and (B.30) in their full generality, and so their results apply to isostable coordinates - up to one limitation, discussed below - simply by making the following substitutions:

$$\theta \to \phi,\ \vec{\rho} \to \vec{\psi} \tag{B.33}$$

$$\vec{a}(\theta) \to 0,\ \mathbf{R}(\theta) \to -\mathbf{K},\ h(\theta) \to z_G, \mathbf{H}(\theta) \to \mathbf{Y}_G \mathrm{e}^{\mathbf{K}(1-\mathrm{mod}_1\phi)}. \tag{B.34}$$

Recall, however, that we define a general Poincare section by $\phi = \phi_S\left(\vec{\psi}\right) \sim \vec{m}_S^T\vec{\psi}$ (see the discussion around (3.22)). In contrast, HM's analysis is limited to the case where the Poincare section is given by $\theta = 0$, which in their work is to be interpreted as a surface orthogonal to the limit cycle. The replacements (B.33) and (B.34) would reflect the choice $\phi_S\left(\vec{\psi}\right) = 0$, for which the Poincare section is an isochron. One could extend HM's rigorous work, generalizing it for arbitrary Poincare sections in moving orthogonal coordinates. We opt instead to present a less formal derivation for arbitrary sections in the simpler phase-isostable coordinates.

Mirroring HM, we begin with the solution of (B.31) and (B.32) with the initial condition restricted to the section: $\vec{\psi}(t_n) = \vec{\psi}_n$ and $\phi(t_n) = \vec{m}_S^T \vec{\psi}_n$. The subscript n reflects that this starting point is the n^{th} crossing of the section. The equation for ϕ, (B.31), simply can be integrated, yielding

$$\phi(t) \sim \quad \mathrm{mod}_1 \left(\vec{m}_S^T \vec{\psi}_n + t - t_n + \sigma \vec{z}_G^T \left(\vec{W}(t) - \vec{W}(t_n) \right) \right) \tag{B.35}$$

to first order in σ and $\vec{\psi}$. We rewrite (B.32) using an integrating factor of $\mathrm{e}^{\mathbf{K}(t-t_n)}$ and (B.35),

$$d\left(\mathrm{e}^{\mathbf{K}(t-t_n)} \vec{\psi} \right) \sim \sigma \mathrm{e}^{\mathbf{K}\left[(t-t_n)+1-\mathrm{mod}_1\left(t-t_n+\mathcal{O}\left(\sigma,|\vec{\psi}|\right)\right)\right]} \mathbf{Y}_G^T \, d\vec{W}(t) \,. \tag{B.36}$$

Since the first return time to the Poincare section is $t_{n+1} = t_n + 1 + \mathcal{O}\left(\sigma, \left|\vec{\psi}\right|\right)$, we will integrate (B.36) over an interval $t \in \left(t_n, t_n + 1 + \mathcal{O}\left(\sigma, \left|\vec{\psi}\right|\right)\right)$. Note that $\mathrm{mod}_1\left(t - t_n + \mathcal{O}\left(\sigma, \left|\vec{\psi}\right|\right)\right)$ is simply $t - t_n + \mathcal{O}\left(\sigma, \left|\vec{\psi}\right|\right)$ for all but a subset of measure $\mathcal{O}\left(\sigma, \left|\vec{\psi}\right|\right)$ of that interval. Given that (B.36) is overall $\mathcal{O}\left(\sigma, \left|\vec{\psi}\right|\right)$, we may therefore replace the factor $[(t-t_n) + 1 - \mathrm{mod}_1(\ldots)]$ in the exponent on the right hand side of (B.36) with 1 to first order approximation. Integrating then yields

$$\mathrm{e}^{\mathbf{K}(t-t_n)} \vec{\psi}(t) \sim \vec{\psi}_n + \sigma \mathrm{e}^{\mathbf{K}} \mathbf{Y}_G^T \left(\vec{W}(t) - \vec{W}(t_n) \right),$$

for $t \in \left(t_n, t_n + 1 + \mathcal{O}\left(\sigma, \left|\vec{\psi}\right|\right)\right)$. Once we further evaluate at $t = t_{n+1}$, $\mathrm{e}^{\mathbf{K}(t-t_n)} = \mathrm{e}^{\mathbf{K}} + \mathcal{O}\left(\sigma, \left|\vec{\psi}\right|\right)$, and so the Poincare map takes $\vec{\psi}_n$ to

$$\vec{\psi}_{n+1} \equiv \vec{\psi}(t_{n+1}) \sim \mathrm{e}^{-\mathbf{K}} \vec{\psi}_n + \sigma \mathbf{Y}_G^T \Delta \vec{W}_n, \tag{B.37}$$

where $\Delta\vec{W}_n = \vec{W}(t_n+1) - \vec{W}(t_n)$ is a standard normal vector random variable. Also, from (B.35)

$$\phi(t_{n+1}) \sim \vec{m}_S^T \vec{\psi}_n + \sigma \vec{z}_G^T \Delta\vec{W}_n. \tag{B.38}$$

We now derive the first passage time to first order approximation. Because the Poincare section is defined via (3.22), the first return time is the first time t_{n+1} such that

$$\phi(t_{n+1}) = \vec{m}_S^T \vec{\psi}(t_{n+1}) \tag{B.39}$$

and $t_{n+1} = t_n + 1 + \mathcal{O}\left(\sigma, \left|\vec{\psi}\right|\right)$. Estimating the first passage time $\Delta T_n^S = t_{n+1} - t_n$ to first order in σ and $\vec{\psi}_n$, we write it as

$$\Delta T_n^S \sim 1 + \vec{\delta\mu}^T \vec{\psi}_n + \sigma\zeta_n. \tag{B.40}$$

Inserting $t_{n+1} = t_n + \Delta T_n^S$ into (B.39) and expanding it around $t_n + 1$ using (B.37) and (B.38) and the facts that $\dot{\phi} = 1 + \mathcal{O}\left(\sigma, \left|\vec{\psi}\right|\right)$ and $\dot{\vec{\psi}} = \mathcal{O}\left(\sigma, \left|\vec{\psi}\right|\right)$ then gives to first order

$$\begin{aligned} &\phi(t_n+1) + \dot{\phi}(t_n+1)\left(\vec{\delta\mu}^T\vec{\psi}_n + \sigma\zeta_n\right) \sim \vec{m}_S^T\left(\vec{\psi}(t_n+1) + \dot{\vec{\psi}}(t_n+1)\left(\vec{\delta\mu}^T\vec{\psi}_n + \sigma\zeta_n\right)\right) \\ &\left(\vec{m}_S^T\vec{\psi}_n + \sigma\vec{z}_G^T\Delta\vec{W}_n\right) + \left(\vec{\delta\mu}^T\vec{\psi}_n + \sigma\zeta_n\right) \sim \vec{m}_S^T\left[\mathrm{e}^{-\mathbf{K}}\vec{\psi}_n + \sigma\mathbf{Y}_G^T\Delta\vec{W}_n\right]. \end{aligned} \tag{B.41}$$

From (B.41) we see that

$$\vec{\delta\mu}^T \equiv -\vec{m}_S^T(\mathbf{I} - \mathbf{\Lambda}), \tag{B.42}$$

where $\mathbf{\Lambda} \equiv \mathrm{e}^{-\mathbf{K}}$, and

$$\begin{aligned} \zeta_n &\equiv \left(\vec{m}_S^T \mathbf{Y}_G^T - \vec{z}_G^T\right) \Delta\vec{W}_n \\ &= \left(-1, \;\; \vec{m}_S^T \; \mathbf{K}^{-\frac{1}{2}} \right) \mathbf{R}_G^T \Delta\vec{W}_n, \end{aligned} \tag{B.43}$$

where $\boldsymbol{R}_G = \left(\vec{z}_G, \;\; \vec{y}_{G,1}, \;\; \vec{y}_{G,2}, \;\; \cdots, \;\; \vec{y}_{G,d} \right)$, as defined in (B.28). Following (B.37), (B.40), and (B.43), the covariance matrix for $\left(\Delta T_n^S; \;\; \vec{\psi}_{n+1} \right)$ conditioned on $\vec{\psi}_n$ is given by

$$\begin{pmatrix} \alpha & \vec{\beta}^T \\ \vec{\beta} & \mathbf{\Gamma} \end{pmatrix} \equiv \begin{pmatrix} \left(\vec{m}_S^T \mathbf{Y}_G^T - \vec{z}_G^T\right) \left(\mathbf{Y}_G \vec{m}_S - \vec{z}_G\right) & \left(\vec{m}_S^T \mathbf{Y}_G^T - \vec{z}_G^T\right) \mathbf{Y}_G \\ \mathbf{Y}_G^T \left(\mathbf{Y}_G \vec{m}_S - \vec{z}_G\right) & \mathbf{Y}_G^T \mathbf{Y}_G \end{pmatrix} \tag{B.44}$$

$$= \begin{pmatrix} -1 & \vec{m}_S^T \\ \vec{0} & \mathbf{I} \end{pmatrix} \mathbf{\Xi} \begin{pmatrix} -1 & \vec{0}^T \\ \vec{m}_S & \mathbf{I} \end{pmatrix}, \tag{B.45}$$

where $\mathbf{\Xi} = \mathbf{R}_G^T \mathbf{R}_G$. It will be useful to rewrite α and $\vec{\beta}$ in terms of

$$\mathbf{\Gamma} \equiv \mathbf{Y}_G^T \mathbf{Y}_G \tag{B.46}$$

as

$$\begin{aligned} \alpha &= \vec{m}_S^T \mathbf{\Gamma} \vec{m}_S - 2\vec{m}_S^T \mathbf{Y}_G^T \vec{z}_G + \vec{z}_G^T \vec{z}_G \\ \vec{\beta} &= \mathbf{\Gamma} \vec{m}_S - \mathbf{Y}_G^T \vec{z}_G. \end{aligned} \tag{B.47}$$

B.3.2. Steady-State Probability Densities

Computation of the TVGR will require the steady-state distribution of $\vec{\psi}_n$ for n large, (3.24), the joint distribution of $\vec{\psi}_n$ and $\vec{\psi}_{n+m}$ for large n, (3.26), and the distribution of ΔT_n conditioned on $\vec{\psi}_n$ and $\vec{\psi}_{n+1}$, (B.52) below. These all follow from the linearized Poincare map and first passage time (3.23). Written in terms of a conditional distribution,

$$\text{(B.48)} \qquad \begin{pmatrix} \Delta T_n^S - 1 \\ \vec{\psi}_{n+1} \end{pmatrix} \underset{\text{given } \vec{\psi}_n}{\overset{dist.}{\sim}} \mathcal{N}\left(\begin{pmatrix} \vec{\delta\mu}^T \\ \mathbf{\Lambda} \end{pmatrix} \vec{\psi}_n, \sigma^2 \begin{pmatrix} \alpha & \vec{\beta}^T \\ \vec{\beta} & \mathbf{\Gamma} \end{pmatrix} \right),$$

where $\mathcal{N}(\vec{\mu}, \mathbf{\Sigma})$ is the multivariate normal distribution with mean $\vec{\mu}$ and covariance matrix $\mathbf{\Sigma}$ and α, $\vec{\beta}$ and $\mathbf{\Gamma}$ are as defined in Section B.3.1. The distribution for $\vec{\psi}_n$ is found in Section 3.5.2, and we outline the derivation of the others here.

To start, we cite elementary formulas for connecting conditional and joint multivariate normal distributions (see e.g. [**17**]). Generally, for two (potentially vector-valued) random variables y_1 and y_2 that are jointly normally distributed, i.e.

$$\text{(B.49)} \qquad \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \overset{dist.}{\sim} \mathcal{N}\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{12} \\ \mathbf{\Sigma}_{12}^T & \mathbf{\Sigma}_{22} \end{pmatrix} \right),$$

the distribution of y_1 conditioned on y_2 is again normally distributed:

$$\text{(B.50)} \qquad y_1 \underset{\text{given } y_2}{\overset{dist.}{\sim}} \mathcal{N}\left(\mu_1 + \mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}(y_2 - \mu_2),\ \mathbf{\Sigma}_{11} - \mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}\mathbf{\Sigma}_{12}^T \right).$$

Now, using (B.49) and (B.50) we will recover the joint distribution for ΔT_n^S, $\vec{\psi}_n$, and $\vec{\psi}_{n+1}$ from the conditional distribution for ΔT_n^S and $\vec{\psi}_{n+1}$. Taking $y_1 = \left(\Delta T_n^S - 1; \vec{\psi}_{n+1}\right)$ and

$y_2 = \vec{\psi}_n$ and comparing (B.48) and (B.50), we find that

$$\begin{aligned}
\mathbf{\Sigma}_{12} &= \begin{pmatrix} \vec{\delta\mu}^T \\ \mathbf{\Lambda} \end{pmatrix} \mathbf{\Sigma}_{22} \\
\mathbf{\Sigma}_{11} &= \sigma^2 \begin{pmatrix} \alpha & \vec{\beta}^T \\ \vec{\beta} & \mathbf{\Gamma} \end{pmatrix} + \begin{pmatrix} \vec{\delta\mu}^T \\ \mathbf{\Lambda} \end{pmatrix} \mathbf{\Sigma}_{22} \begin{pmatrix} \vec{\delta\mu} & \mathbf{\Lambda} \end{pmatrix} \\
\mu_1 &= \begin{pmatrix} \vec{\delta\mu}^T \\ \mathbf{\Lambda} \end{pmatrix} \mu_2,
\end{aligned}$$

where $\mathbf{\Lambda} \equiv \mathrm{e}^{-\mathbf{K}}$. From (3.24) we have $\mu_2 = \mathrm{E}\left\{\tilde{\psi}_{\mathrm{n+1}}\right\} = 0$ and $\mathbf{\Sigma}_{22} = \mathrm{var}\left\{\vec{\psi}_{n+1}\right\} = \sigma^2 \mathbf{\Gamma}_{ss}$ in steady-state, and (B.49) becomes

$$\text{(B.51)} \quad \begin{pmatrix} \Delta T_n^S - 1 \\ \vec{\psi}_{n+1} \\ \vec{\psi}_n \end{pmatrix} \overset{dist.}{\sim} \mathcal{N}\left(0, \sigma^2 \begin{pmatrix} \alpha + \vec{\delta\mu}^T \mathbf{\Gamma}_{ss} \vec{\delta\mu} & \vec{\beta}^T + \vec{\delta\mu}^T \mathbf{\Gamma}_{ss} \mathbf{\Lambda} & \vec{\delta\mu}^T \mathbf{\Gamma}_{ss} \\ \vec{\beta} + \mathbf{\Lambda}\mathbf{\Gamma}_{ss}\vec{\delta\mu} & \mathbf{\Gamma}_{ss} & \mathbf{\Lambda}\mathbf{\Gamma}_{ss} \\ \mathbf{\Gamma}_{ss}\vec{\delta\mu} & \mathbf{\Gamma}_{ss}\mathbf{\Lambda} & \mathbf{\Gamma}_{ss} \end{pmatrix}\right).$$

Similarly, taking $y_1 = \Delta T_n^S - 1$ and $y_2 = \vec{\psi}_{n+1}$ and matching (B.48) and (B.49), (B.50) gives the conditional distribution for ΔT_n^S,

$$\text{(B.52)} \quad \Delta T_n^S \underset{\text{given } \vec{\psi}_n \text{and} \vec{\psi}_{n+1}}{\overset{dist.}{\sim}} \mathcal{N}\left(1 + \vec{\delta\mu}^T \vec{\psi}_n + \vec{\beta}^T \mathbf{\Gamma}^{-1}\left(\vec{\psi}_{n+1} - \mathbf{\Lambda}\vec{\psi}_n\right), \sigma^2\left[\alpha - \vec{\beta}^T \mathbf{\Gamma}^{-1} \vec{\beta}\right]\right).$$

We now consider (3.26), the joint distribution of $\vec{\psi}_n$ and $\vec{\psi}_{n+m}$, for $m > 0$ and $n \to \infty$. Applying (3.23) recursively and neglecting higher-order terms, we have

$$\vec{\psi}_{n+m} = \mathbf{\Lambda}^m \vec{\psi}_n + \sum_{k=1}^{m} \mathbf{\Lambda}^{k-1} \vec{\eta}_{n+m-k}$$

Since the $\vec{\eta}_{n+m-k}$ are i.i.d. with mean 0 and covariance matrix Γ, $\text{var}\left\{\vec{\psi}_{n+m} \middle| \vec{\psi}_n\right\} = \sigma^2 \sum_{k=0}^{m-1} \mathbf{\Lambda}^k \mathbf{\Gamma} \mathbf{\Lambda}^k$ or, using (3.25), $\sigma^2 \left(\mathbf{\Gamma}_{ss} - \mathbf{\Lambda}^m \mathbf{\Gamma}_{ss} \mathbf{\Lambda}^m\right)$. So

$$\vec{\psi}_{n+m} \underset{\text{given } \vec{\psi}_n}{\overset{dist.}{\sim}} \mathcal{N}\left(\mathbf{\Lambda}^m \vec{\psi}_n,\, \sigma^2 \left(\mathbf{\Gamma}_{ss} - \mathbf{\Lambda}^m \mathbf{\Gamma}_{ss} \mathbf{\Lambda}^m\right)\right). \tag{B.53}$$

Then, as we did above for (B.51), we recover the joint density (3.26) using (B.49) and (B.50) and the facts that $E\left\{\vec{\psi}_n\right\} = \text{E}\left\{\tilde{\psi}_{\text{n+m}}\right\} = 0$ and $\text{var}\left\{\vec{\psi}_n\right\} = \text{var}\left\{\vec{\psi}_{n+m}\right\} = \sigma^2 \mathbf{\Gamma}_{ss}$.

B.3.3. The TVGR Term-by-Term

We compute the TVGR for the limit cycle oscillator via (3.15) and using the covariance matrix (B.44) as well as the other distributions cited in Sections 3.5.2 and B.3.2. Recall the state x of the Markov renewal process is the position $\vec{\psi}$ on the Poincare section. In the following, we use x and $\vec{\psi}$ interchangeably.

In order to evaluate $\mathcal{V}_E^{(t)}$, we find it useful to first compute $\mathcal{CV}^{(n)}_{(x \to x') \mapsto h(x), (x \to x') \mapsto x}$ and $\mathcal{CV}^{(n)}_{(x \to x') \mapsto h(x), (x \to x') \mapsto x'}$ for an arbitrary scalar or vector-valued, deterministic function h. Since $\text{E}\left\{x_n\right\} = \text{E}\left\{\vec{\psi}_n\right\} = 0$, (3.17) gives

$$\mathcal{CV}^{(n)}_{(x \to x') \mapsto h(x), (x \to x') \mapsto x} = \text{E}\left\{h\left(x_n\right) x_n^T\right\} + \sum_{k=1}^{\infty} \text{E}\left\{h\left(x_n\right) x_{n+k}^T\right\} + \sum_{k=1}^{\infty} \text{E}\left\{h\left(x_{n+k}\right) x_n^T\right\}. \tag{B.54}$$

Recall that we use $N(\cdot, \mu, \Sigma)$ to denote the probability density function for the normal distribution with mean μ and covariance matrix Σ. Note then that

$$\mathrm{E}\left\{h(x_n)\, x_{n+k}^T\right\} = \int_{S\times S} h(x_n)\, x_{n+k}^T N\left(\begin{pmatrix} x_{n+k} \\ x_n \end{pmatrix}; 0, \sigma^2 \mathbf{C}_k\right) dx_{n+k}\, dx_n,$$

where C_k is as defined in (3.26). We can integrate over x_{n+k} to find for $k \geq 0$

$$\begin{aligned} \mathrm{E}\left\{h(x_n)\, x_{n+k}^T\right\} &= \mathrm{E}\left\{h(x_n)\, \mathrm{E}\left\{x_{n+k}^T \,|x_n\right\}\right\} \\ &= \mathrm{E}\left\{h(x_n)\, \mathrm{x_n^T}\mathbf{\Lambda}^k\right\} = \mathcal{H}\mathbf{\Lambda}^k, \end{aligned}$$

where from (B.53) $\mathrm{E}\left\{x_{n+k}^T \,|x_n\right\} = \mathrm{x_n^T}\mathbf{\Lambda}^k$ and we define

$$\mathcal{H} \equiv \int_S h(x)\, x^T N\left(x; 0, \sigma^2 \mathbf{\Gamma}_{ss}\right) dx. \tag{B.55}$$

Analogously, using (3.26) and (B.50), we have $\mathrm{E}\left\{x_n^T \,|x_{n+k}\right\} = x_{n+k}^T \mathbf{\Gamma}_{ss}^{-1} \mathbf{\Lambda}^k \mathbf{\Gamma}_{ss}$, and, thus, for $k \geq 0$,

$$\begin{aligned} \mathrm{E}\left\{h(x_{n+k})\, x_n^T\right\} &= \mathrm{E}\left\{h(x_{n+k})\, \mathrm{E}\left\{x_n^T \,|x_{n+k}\right\}\right\} \\ &= \mathrm{E}\left\{h(x_{n+k})\, x_{n+k}^T \mathbf{\Gamma}_{ss}^{-1} \mathbf{\Lambda}^k \mathbf{\Gamma}_{ss}\right\} \\ &= \mathcal{H}\mathbf{\Gamma}_{ss}^{-1} \mathbf{\Lambda}^k \mathbf{\Gamma}_{ss}. \end{aligned}$$

So the sums in (B.54) can be written in closed-form as geometric series,

$$\begin{aligned} \mathcal{CV}^{(n)}_{(x\to x')\mapsto h(x),(x\to x')\mapsto x} &= \mathcal{H}\left[\mathbf{I} + \left((\mathbf{I}-\mathbf{\Lambda})^{-1} - \mathbf{I}\right) + \mathbf{\Gamma}_{ss}^{-1}\left((\mathbf{I}-\mathbf{\Lambda})^{-1} - \mathbf{I}\right)\mathbf{\Gamma}_{ss}\right] \\ &= \mathcal{H}\left[-\mathbf{I} + (\mathbf{I}-\mathbf{\Lambda})^{-1} + \mathbf{\Gamma}_{ss}^{-1}(\mathbf{I}-\mathbf{\Lambda})^{-1}\mathbf{\Gamma}_{ss}\right]. \end{aligned}$$

It will be useful to decompose the expression in brackets as

$$\mathcal{CV}^{(n)}_{(x\to x')\mapsto h(x),(x\to x')\mapsto x} = \mathcal{H}\boldsymbol{\Gamma}_{ss}^{-1}\left(\mathbf{I}-\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Gamma}\left(\mathbf{I}-\boldsymbol{\Lambda}\right)^{-1}, \tag{B.56}$$

which can be done via the fact that $\boldsymbol{\Gamma}_{ss} = \boldsymbol{\Lambda}\boldsymbol{\Gamma}_{ss}\boldsymbol{\Lambda}+\boldsymbol{\Gamma}$ (see the discussion preceding (3.25)). It can similarly be shown that

$$\begin{aligned}\mathcal{CV}^{(n)}_{(x\to x')\mapsto h(x),(x\to x')\mapsto x'} &= \mathcal{H}\left[\boldsymbol{\Lambda}+\left(\left(\mathbf{I}-\boldsymbol{\Lambda}\right)^{-1}-\left(\mathbf{I}+\boldsymbol{\Lambda}\right)\right)+\boldsymbol{\Gamma}_{ss}^{-1}\left(\mathbf{I}-\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Gamma}_{ss}\right] \\ &= \mathcal{H}\boldsymbol{\Gamma}_{ss}^{-1}\left(\mathbf{I}-\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Gamma}\left(\mathbf{I}-\boldsymbol{\Lambda}\right)^{-1} \qquad \text{(B.57)}\end{aligned}$$

is precisely the same as $\mathcal{CV}^{(n)}_{(x\to x')\mapsto h(x),(x\to x')\mapsto x}$.

It is also useful to consider the expression $\Delta t\left(x,x'\right)$ that appears in the TVGR. Recall from the discussion in Section 3.3.2 and Appendix B.1 that $\Delta t\left(x,x'\right)$ is a random function of x and x': even when conditioned on x and x', there may be uncertainty in its value. For the linearized Poincare map dynamics, $\Delta t\left(\vec{\psi}_n,\vec{\psi}_{n+1}\right)$ has the same distribution as ΔT_n^S conditioned on $\vec{\psi}_n$ and $\vec{\psi}_{n+1}$ and can be decomposed as

$$\Delta t\left(\vec{\psi}_n,\vec{\psi}_{n+1}\right) = \overline{\Delta t}\left(\vec{\psi}_n,\vec{\psi}_{n+1}\right)+\delta t_n, \tag{B.58}$$

where $\overline{\Delta t}\left(x,x'\right)$ is the mean of $\Delta t\left(x,x'\right)$ conditioned on x and x' and the δt_n are zero mean i.i.d. Gaussian random variables independent of the $\vec{\psi}_j$ for any j (see (B.52)). Since each of the terms constituting the mixed component are covariances of $\Delta t\left(\vec{\psi}_n,\vec{\psi}_{n+1}\right)$ with $1_E\left(\vec{\psi}_j\right)$ (see (3.17)), the mixed component does not depend on the δt_n, i.e. it can

be rewritten as

$$\mathcal{CV}^{(n)}_{(x\to x')\mapsto 1_E(x),(x\to x')\mapsto \Delta t(x,x')} = \mathcal{CV}^{(n)}_{(x\to x')\mapsto 1_E(x),(x\to x')\mapsto \overline{\Delta t}(x,x')}.$$

While δt_n likewise does not contribute to the covariances in the temporal component, $\mathcal{V}^{(n)}_{(x\to x')\mapsto \Delta t(x,x')}$, it does affect the first, variance term in (3.9) $\text{var}\left\{\Delta t\left(\vec{\psi}_0, \vec{\psi}_1\right)\right\}$. By the law of total variance

$$\begin{aligned} \text{var}\left\{\Delta t\left(\vec{\psi}_0, \vec{\psi}_1\right)\right\} &= \text{var}\left\{\text{E}\left\{\Delta t\left(\vec{\psi}_0, \vec{\psi}_1\right)\middle|\vec{\psi}_0, \vec{\psi}_1\right\}\right\} + \text{E}\left\{\text{var}\left\{\Delta t\left(\vec{\psi}_0, \vec{\psi}_1\right)\middle|\vec{\psi}_0, \vec{\psi}_1\right\}\right\} \\ &= \text{var}\left\{\overline{\Delta t}\left(\vec{\psi}_0, \vec{\psi}_1\right)\right\} + \text{var}\left\{\delta t_0\right\}. \end{aligned} \tag{B.59}$$

The first term of (B.59) is the first term of $\mathcal{V}^{(n)}_{(x\to x')\mapsto \overline{\Delta t}(x,x')}$. So, beyond those terms included in $\mathcal{V}^{(n)}_{(x\to x')\mapsto \overline{\Delta t}(x,x')}$, $\mathcal{V}^{(n)}_{(x\to x')\mapsto \Delta t(x,x')}$ has the additional contribution $\text{var}\left\{\delta t_0\right\}$. Thus,

$$\mathcal{V}^{(t)}_E = \mathcal{V}^{(n)}_{x\mapsto 1_E(x)} - 2\,\mathcal{E}\,\mathcal{CV}^{(n)}_{(x\to x')\mapsto 1_E(x),(x\to x')\mapsto \overline{\Delta t}(x,x')} + \mathcal{E}^2\mathcal{V}^{(n)}_{(x\to x')\mapsto \overline{\Delta t}(x,x')} + \text{var}\left\{\delta t_0\right\}.$$

We now evaluate $\mathcal{V}^{(t)}_E$ term-by-term.

(1) $\mathcal{V}^{(n)}_{x\mapsto 1_E(x)}$

This term is as written in (3.28).

(2) $\mathcal{CV}^{(n)}_{(x\to x')\mapsto 1_E(x),(x\to x')\mapsto \overline{\Delta t}(x,x')}$

From (B.52) and (B.58) we see that $\overline{\Delta t}\left(\vec{\psi}_n, \vec{\psi}_{n+1}\right)$ is linear in $x_n = \vec{\psi}_n$ and $x_{n+1} = \vec{\psi}_{n+1}$, we write

$$\overline{\Delta t}\left(\vec{\psi}_n, \vec{\psi}_{n+1}\right) = 1 + \vec{b}_0^T\vec{\psi}_n + \vec{b}_1^T\vec{\psi}_{n+1}, \tag{B.60}$$

where, using (B.42), (B.47), and (B.52),

$$\begin{aligned}\vec{b}_1 &= \mathbf{\Gamma}^{-1}\vec{\beta} = \vec{m}_S - \mathbf{\Gamma}^{-1}\mathbf{Y}_G^T\vec{z}_G \\ \vec{b}_0 &= \vec{\delta\mu} - \mathbf{\Lambda}\vec{b}_1 = -\vec{m}_S + \mathbf{\Lambda}\mathbf{\Gamma}^{-1}\mathbf{Y}_G^T\vec{z}_G.\end{aligned}$$

Because it will be useful, we note that their sum is simply

$$\vec{b}_0 + \vec{b}_1 = -\left(\mathbf{I} - \mathbf{\Lambda}\right)\mathbf{\Gamma}^{-1}\mathbf{Y}_G^T\vec{z}_G. \tag{B.61}$$

Then, since $\mathcal{CV}$ is a bilinear functional (see (3.16)),

$$\mathcal{CV}^{(n)}_{(x\to x')\mapsto 1_E(x),(x\to x')\mapsto\overline{\Delta t}(x,x')} = \mathcal{CV}_{(x\to x')\mapsto 1_E(x),(x\to x')\mapsto x}\vec{b}_0 + \mathcal{CV}_{(x\to x')\mapsto 1_E(x),(x\to x')\mapsto x'}\vec{b}_1,$$

where $x = \vec{\psi}$. Now we can apply (B.56) and (B.57) with $h\left(x\right) = 1_E\left(x\right)$, yielding

$$\mathcal{CV}^{(n)}_{(x\to x')\mapsto 1_E(x),(x\to x')\mapsto\overline{\Delta t}(x,x')} = \mathcal{X}^T\mathbf{\Gamma}_{ss}^{-1}\left(\mathbf{I} - \mathbf{\Lambda}\right)^{-1}\mathbf{\Gamma}\left(\mathbf{I} - \mathbf{\Lambda}\right)^{-1}\left(\vec{b}_0 + \vec{b}_1\right),$$

where, using (B.55),

$$\mathcal{X}^T = \mathrm{E}\left\{1_E\left(x_n\right)x_n\right\} = \int_E x^T N\left(x; 0, \sigma^2\mathbf{\Gamma}_{ss}\right)dx.$$

Note that $\mathcal{X} = x_E\mathcal{E}$, where x_E and $\mathcal{E}$ are as defined in Section 3.5.3. We rewrite $\mathcal{CV}_{(x\to x')\mapsto 1_E(x),(x\to x')\mapsto\overline{\Delta t}(x,x')}$ as $-\frac{1}{2}\mathcal{X}^T\vec{b}$, where

$$\vec{b} = 2\mathbf{\Gamma}_{ss}^{-1}\left(\mathbf{I} - \mathbf{\Lambda}\right)^{-1}\mathbf{Y}_G^T\vec{z}_G \tag{B.62}$$

follows from (B.61) and contains the factors that are independent of σ and E.

(3) $\mathcal{V}^{(n)}_{(x\to x')\mapsto\overline{\Delta t}(x,x')}$

Again decomposing $\overline{\Delta t}$ via (B.60), we have from the linearity of $\mathcal{V}$ and $\mathcal{CV}$ (see (3.9) and (3.17)),

(B.63)
$$\mathcal{V}^{(n)}_{(x\to x')\mapsto\overline{\Delta t}(x,x')} = \begin{pmatrix} \vec{b}_0^{\,T}, & \vec{b}_1^{\,T} \end{pmatrix} \begin{pmatrix} \mathcal{V}^{(n)}_{(x\to x')\mapsto x} & \mathcal{CV}^{(n)}_{(x\to x')\mapsto x,(x\to x')\mapsto x'} \\ \mathcal{CV}^{(n)}_{(x\to x')\mapsto x',(x\to x')\mapsto x} & \mathcal{V}^{(n)}_{(x\to x')\mapsto x'} \end{pmatrix} \begin{pmatrix} \vec{b}_0 \\ \vec{b}_1 \end{pmatrix}.$$

Note that $\mathcal{V}^{(n)}_{(x\to x')\mapsto x} = \mathcal{CV}^{(n)}_{(x\to x')\mapsto x,(x\to x')\mapsto x}$, and so we can apply (B.56) and (B.57) for each of the matrix-valued entries in the matrix appearing in (B.63). But since (B.56) and (B.57) give the same result, all of those entries are the same: $\mathcal{V}^{(n)}_{(x\to x')\mapsto x}$. To compute $\mathcal{V}^{(n)}_{(x\to x')\mapsto x}$, we take h equal to the identity function. (B.55) gives $\mathcal{H} = \sigma^2\mathbf{\Gamma}_{ss}$, and thus

$$\mathcal{V}^{(n)}_{(x\to x')\mapsto x} = \sigma^2 \left(\mathbf{I}-\mathbf{\Lambda}\right)^{-1}\mathbf{\Gamma}\left(\mathbf{I}-\mathbf{\Lambda}\right)^{-1}.$$

Then we have from (B.61) and (B.63)

$$\begin{aligned} \mathcal{V}^{(n)}_{(x\to x')\mapsto\overline{\Delta t}(x,x')} &= \left(\vec{b}_0+\vec{b}_1\right)^T \mathcal{V}^{(n)}_{(x\to x')\mapsto x}\left(\vec{b}_0+\vec{b}_1\right) \\ &= \sigma^2 \vec{z}_G^{\,T}\mathbf{Y}_G\mathbf{\Gamma}^{-1}\mathbf{Y}_G^T\vec{z}_G \end{aligned} \tag{B.64}$$

(4) $\mathrm{var}\left\{\delta t_0\right\}$

(B.52) reveals that the conditional variance of $\Delta t\left(x,x'\right)$ is given by

$\sigma^2 \left[\alpha - \vec{\beta}^T \mathbf{\Gamma}^{-1} \vec{\beta}\right]$. Then following (B.47),

$$\begin{aligned}
\text{var}\left\{\delta t_0\right\} &= \sigma^2 \left[\alpha - \vec{\beta}^T \mathbf{\Gamma}^{-1} \vec{\beta}\right] \\
&= \sigma^2 \left[\vec{m}_S^T \mathbf{\Gamma} \vec{m}_S - 2\vec{m}_S^T \mathbf{Y}_G^T \vec{z}_G + \vec{z}_G^T \vec{z}_G - \left(\vec{m}_S^T \mathbf{\Gamma} - \vec{z}_G^T \mathbf{Y}_G\right) \mathbf{\Gamma}^{-1} \left(\mathbf{\Gamma} \vec{m}_S - \mathbf{Y}_G^T \vec{z}_G\right)\right] \\
&= \sigma^2 \left[\vec{z}_G^T \vec{z}_G - \vec{z}_G^T \mathbf{Y}_G \mathbf{\Gamma}^{-1} \mathbf{Y}_G^T \vec{z}_G\right]. \qquad \text{(B.65)}
\end{aligned}$$

(5) $\mathcal{V}_E^{(t)}$

Putting all of the above pieces together, we arrive at (3.31),

$$\mathcal{V}_E^{(t)} = \left(\mathcal{E} - \mathcal{E}^2\right) + 2\sum_{m=1}^{\infty} \left(\mathcal{E}_m - \mathcal{E}^2\right) + \left(\vec{b} \cdot x_E\right) \mathcal{E}^2 + c\frac{\sigma^2}{2}\mathcal{E}^2,$$

where $\vec{b}$ is given by (B.62) and c follows from (B.64) and (B.65):

$$\begin{aligned}
c &= \frac{2}{\sigma^2} \left(\mathcal{V}_{(x \to x') \mapsto \overline{\Delta t}(x,x')} + \text{var}\left\{\delta t_0\right\}\right) \\
&= 2\vec{z}_G^T \vec{z}_G. \qquad \text{(B.66)}
\end{aligned}$$

B.4. Bounds on the Markov-only Component

In this section we derive (3.51): upper and lower bounds on $\mathcal{V}^{(n)}_{x \mapsto 1_E(x)}$, the Markov-only component of the temporal variance growth rate, for planar oscillators. For the lower bound, we show that $\mathcal{E}_m - \mathcal{E}^2 > 0$ for all finite $m \geq 0$. For the upper bound, we bound each term $(\mathcal{E}_m - \mathcal{E}^2)$ individually in such a way that the infinite sum can be carried out as a geometric series. Specifically, we find a $\mathcal{B}$ that is independent of Λ and m such that

$$\mathcal{E}_m - \mathcal{E}^2 \leq \Lambda^m \mathcal{B}, \qquad \text{(B.67)}$$

and therefore

$$\begin{aligned}\mathcal{V}^{(n)}_{x\mapsto 1_E(x)} &= \mathcal{E}\left(1-\mathcal{E}\right)+2\sum_{m=1}^{\infty}\left(\mathcal{E}_m-\mathcal{E}^2\right)\\ &\leq \mathcal{B}+2\frac{\Lambda}{1-\Lambda}\mathcal{B}=\frac{1+\Lambda}{1-\Lambda}\mathcal{B}.\end{aligned}$$

Both the lower and upper bounds involve a comparison of

$$\mathcal{E}_m=\int_{E\times E}N\left(z;0,\sigma^2\mathbf{C}_m\right)dz$$

with $\mathcal{E}^2$, where

$$\mathbf{C}_m=\begin{pmatrix}1 & \Lambda^m\\ \Lambda^m & 1\end{pmatrix}.$$

Note that, since $0\leq\Lambda<1$, $C_\infty=I$, and so $\mathcal{E}^2=\mathcal{E}_\infty$. We start our analysis with a transformation of variables z to ζ such that the resulting integrands are independent of m,

$$\zeta=\frac{1}{\sqrt{2}}\begin{pmatrix}\frac{1}{\sqrt{1+\mu}} & \frac{1}{\sqrt{1+\mu}}\\ -\frac{1}{\sqrt{1-\mu}} & \frac{1}{\sqrt{1-\mu}}\end{pmatrix}z,$$

where $\mu\equiv\Lambda^m$. This amounts to a rotation by $-\frac{\pi}{4}$ and a scaling of the two resulting components by $\frac{1}{\sqrt{1+\mu}}$ and $\frac{1}{\sqrt{1-\mu}}$, respectively. Under this transformation, the integrand becomes the probability density function for a 2D normal random variable ζ with uncorrelated components: $N\left(\zeta;0,\sigma^2I\right)$. The cost of this step, of course, is that the domain of integration $z\in E\times E=\left[\delta-\frac{1}{2},\delta+\frac{1}{2}\right]\times\left[\delta-\frac{1}{2},\delta+\frac{1}{2}\right]$ is no longer rectangular and now depends on m. On the other hand, $\mathcal{E}_m$ and $\mathcal{E}_\infty$ can now be directly compared, since they share the same probability density $N\left(\zeta;0,\sigma^2I\right)$. Each term in the infinite sum, $\mathcal{E}_m-\mathcal{E}_\infty$,

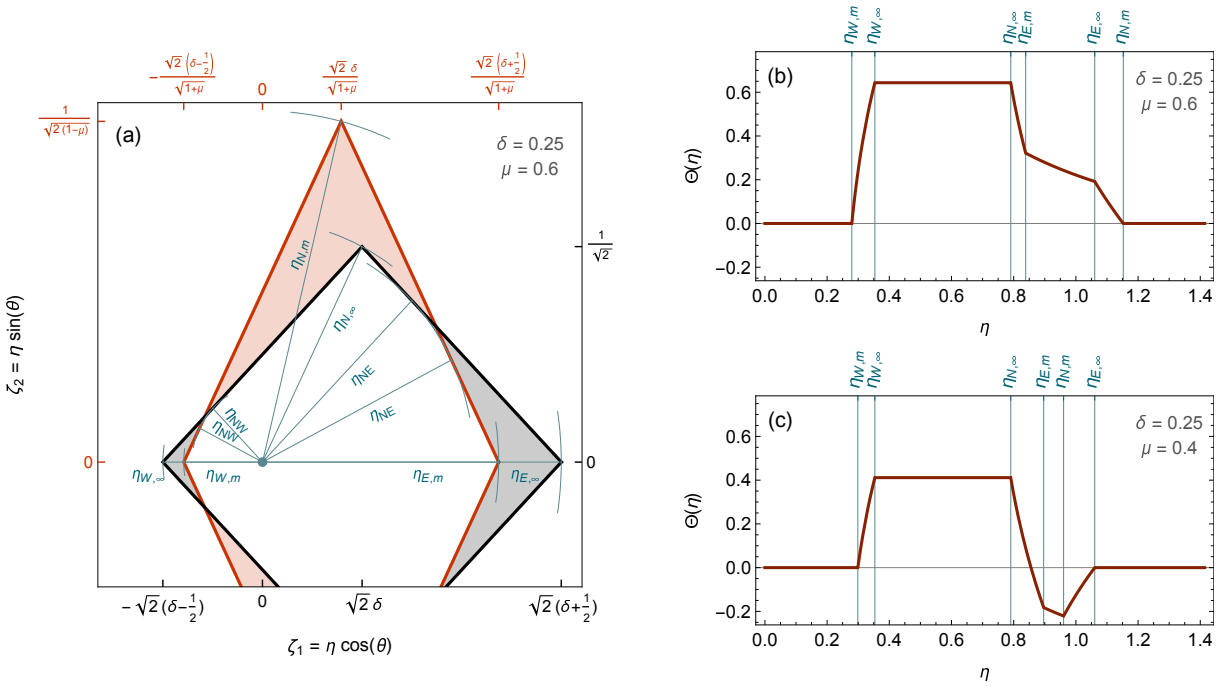

Figure B.2. (a): Regions of integration for $\mathcal{E}_m$ (bounded by the red lines) and $\mathcal{E}_\infty$ (bounded by the black lines) in the ζ-plane. The red regions count positively towards $\mathcal{E}_m - \mathcal{E}_\infty$ and the black regions negatively. This picture is to scale for $\delta = 0.25$ and $\mu = 0.6$, but is qualitatively representative for all $\delta \in \left(0, \frac{1}{2}\right)$ and $\mu \in (0, 1)$. (b) and (c): Radial representation $\Theta(\eta)$ of the exact difference in the areas of the regions. In (a) and (b), $\delta = 0.25$ and $\mu = 0.6$. In (c), where $\delta = 0.25$ and $\mu = 0.4$, $\Theta(\eta)$ becomes negative for an interval of η.

is given by the difference in the total probability mass in a diamond-shaped region and in a square region in the ζ-plane (red and black regions, respectively, Figure B.2a).

Since the probability density has circular symmetry, we analyze the difference in probability mass in polar coordinates. Progress can be made by integrating over the angular coordinate, which yields

$$\mathcal{E}_m - \mathcal{E}_\infty = \frac{1}{2\pi\sigma^2} \int_0^\infty \eta\, \Theta(\eta)\, \mathrm{e}^{-\frac{\eta^2}{2\sigma^2}}\, d\eta, \tag{B.68}$$

where η is the radial coordinate and $\Theta(\eta)$ is the net angular content of the diamond-shaped region less the square region at a distance of η from the origin. Note that for $\mu < 1$ (i.e. $\Lambda < 1$ and $m > 0$) and for η sufficiently small, $\Theta(\eta) = 0$, since the diamond and square overlap at small η. $\Theta(\eta)$ is also 0 for sufficiently large η since the diamond and square are finite for $\mu < 1$. These facts will act as "boundary conditions" in our analysis, since we characterize $\Theta(\eta)$ in terms of its derivative in the following.

Between its initial and final values, $\Theta(\eta)$ is a piecewise smooth function that depends parametrically on δ and μ. We expect it might become non-smooth at radii corresponding to 1) the distance between the origin and the corners of the square and diamond and 2) the minimum distance between the origin and the sides of the square and diamonds (Figure B.2a). As it turns out, though, the distance to the sides of the square and to the sides of the diamond are equal; those radii / distances are invariant in μ. At those values of η, $\Theta(\eta)$ is smooth, since the contributions from the square and diamond cancel. There are 6 remaining, unique "non-smooth radii", taking the symmetry about the ζ_1-axis into account. We label them via cardinal directions (north, east, and west; south is equivalent to north) and an index of m if that radius is specific to the diamond and of ∞ if it is specific to the square (Figure B.2a):

(B.69)
$$\eta_{N,m} = \sqrt{\frac{1}{2(1-\mu)} + \frac{2\delta^2}{1+\mu}};\ \eta_{E,m} = \sqrt{\frac{2}{1+\mu}}\left(\frac{1}{2} + \delta\right);\ \eta_{W,m} = \sqrt{\frac{2}{1+\mu}}\left(\frac{1}{2} - \delta\right).$$

Here the $m = \infty$ counterparts are given by replacing $\mu = \Lambda^m$ with 0. Note that $\eta_{E,m}$ and $\eta_{W,m}$ are strictly decreasing as functions of μ, while $\eta_{N,m}$ is strictly increasing for $\mu \in [0, 1)$ when $\delta \in \left[0, \frac{1}{2}\right)$.

The behavior of the function $\Theta(\eta)$ between those 6 radii can be fully specified by the ordering of the radii. For example, if $\eta_{E,m} > \eta_{N,\infty}$ (as in the two cases shown in Figure B.2), then Θ is constant between $\eta = \eta_{W,\infty}$ and $\eta = \eta_{N,\infty}$. 5 different orderings of the non-smooth radii and therefore 5 different qualitative behaviors of $\Theta(\eta)$ appear non-degenerately for $\delta \in \left(0, \frac{1}{2}\right)$ and $\mu \in (0,1)$ (Figure B.3a,b). Between subsequent pairs of the non-smooth radii, $\Theta(\eta)$ takes on one of only six potential functional forms up to an additive constant. In particular, each piecewise segment of the derivative $\Theta'(\eta)$ is one of 0, $\Theta'_{NE}(\eta)$, $\pm\Theta'_{NW}(\eta)$, or $\pm\left(\Theta'_{NE}(\eta) + \Theta'_{NW}(\eta)\right)$, with

$$\Theta_{Nd}(\eta;\delta) \equiv 2\,\mathrm{arcsec}\left(\frac{\eta}{\eta_{Nd}(\delta)}\right) \tag{B.70}$$

for $d = E, W$, where η_{NE} and η_{NW} are the distances from the origin to the sides of diamond/square (Figure B.2a),

$$\eta_{NE} = \frac{1}{2} + \delta;\ \eta_{NW} = \frac{1}{2} - \delta. \tag{B.71}$$

The functional form of $\Theta_{Nd}(\eta;\delta)$ follows from elementary geometry along with the "initial condition" of $\Theta = 0$ for small η. Since Θ is 0 initially, we see the parts of the circle of radius η that protrude *outside* of the square (diamond) as contributing positively (negatively) towards Θ. For $d = E$, for example, $\frac{\Theta_{NE}(\eta)}{2}$ is the angle subtended by each the northeast and southeast side of the square or diamond when restricted to the *exterior* of a circle of radius η (Figure B.3c,d). Note that while $\eta < \eta_{E,m} < \eta_{E,\infty}$, i.e. before the circle passes through the eastern corner of the diamond, the net contribution from this eastern quadrant is 0 (in Figure B.3c, the dashed and solid arcs are equal in length). It is only when $\eta_{E,m} < \eta < \eta_{E,\infty}$, i.e. when the circle intersects the square and not the diamond on

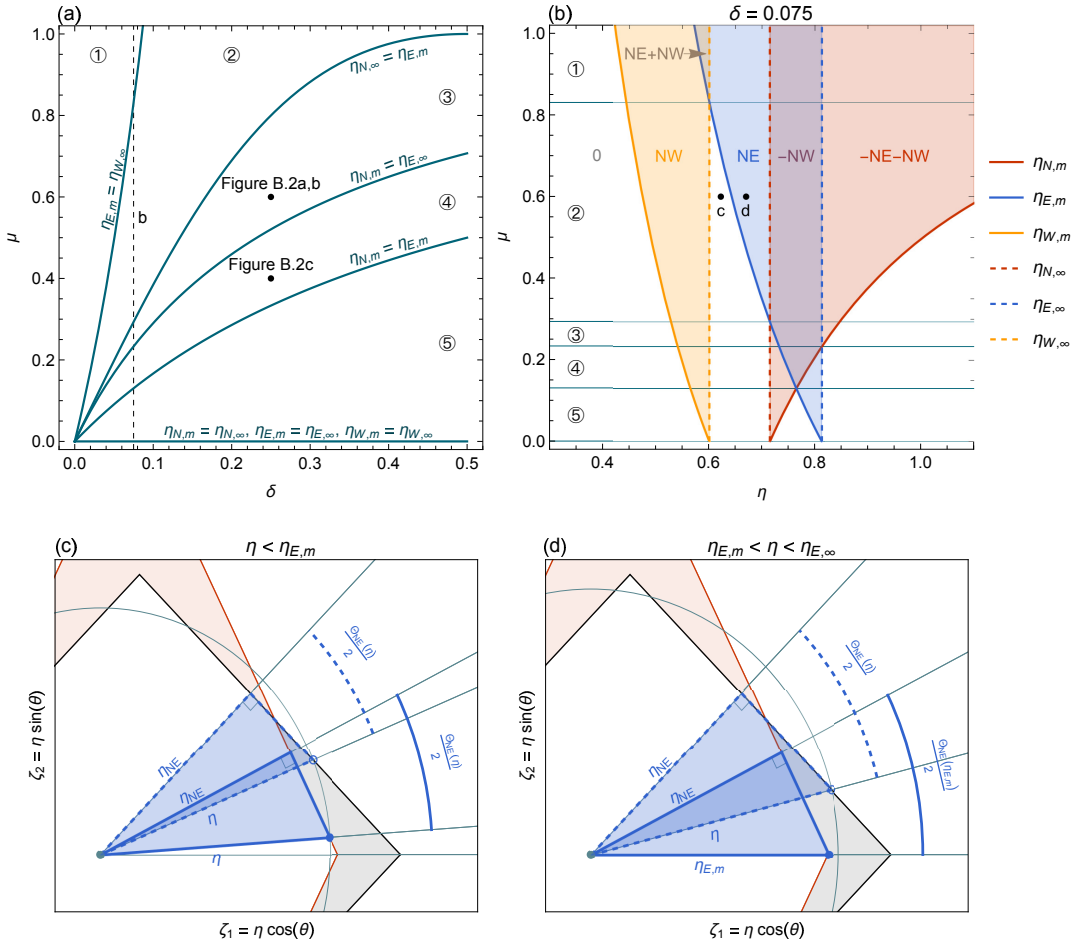

Figure B.3. (a): Curves in the δ-μ plane on which two of the non-smooth radii are degenerately equal. In the numbered regions between the curves, the radii have a particular ordering and $\Theta(\eta)$ a particular piecewise behavior. (b): Curves in the η-μ plane that separate the different smooth segments of $\Theta(\eta)$. While (b) is to-scale only for $\delta = 0.075$ and corresponds to the vertical dashed line in (a), each of the numbered horizontal strips is representative of the corresponding region in (a). They indicate the ordering of non-smooth radii and smooth segments, which is invariant within each region. The shading in (b) indicates the smooth behavior that $\Theta(\eta)$ takes on in that interval of η. 0 means Θ is constant, NE stands in to mean $\Theta'(\eta) = \Theta'_{NE}(\eta)$, $-$NE $-$ NW for $\Theta'(\eta) = -\Theta'_{NE}(\eta) - \Theta'_{NW}(\eta)$, etc. The points marked c and d correspond to the situations shown in subfigures (c) and (d). (c) and (d) give a geometric interpretation of $\Theta_{NE}(\eta)$ in the ζ-plane, as an example: $\frac{\Theta_{NE}(\eta)}{2}$ is one of the acute angles in right triangles with adjacent side length η_{NE} and hypotenuse η.

the northeast and southeast sides, that the net angle accumulates at a rate of $\Theta'_{NE}(\eta)$ (in Figure B.3d, η-dependent dashed arc is longer than the fixed solid arc). The accumulation is due to the increasing part of the circle falling outside the square and stops once $\eta > \eta_{E,\infty}$ (note how the blue shading ends at $\eta = \eta_{E,\infty}$ in Figure B.3b). Note that $\Theta'_{Nd}(\eta) > 0$ and therefore the appearance of a minus sign in some of preceding functional forms indicates that $\Theta(\eta)$ is decreasing, as it does in the northern quadrant. Generally, Θ'_{Nd} makes a positive contribution to Θ when $\eta_{d,m} < \eta < \eta_{d,\infty}$ and a negative contribution when $\eta_{N,\infty} < \eta < \eta_{N,m}$ (Figure B.3b).

With the above, common analysis set-up, we now treat the lower and upper bounds separately.

B.4.1. The Lower Bound

We aim to show that (B.68) can be bounded as

$$\mathcal{E}_m - \mathcal{E}_\infty \propto \int_0^\infty \eta\,\Theta(\eta)\,g(\eta;\sigma)\,d\eta \geq 0, \tag{B.72}$$

where $g(\eta;\sigma) \equiv \mathrm{e}^{-\frac{\eta^2}{2\sigma^2}}$. If $\Theta(\eta)$ is non-negative, the above inequality will hold trivially. This is the case for values of δ and μ corresponding to regions 1-3 of Figure B.3a. For $\mu < 1$ (i.e. $\Lambda < 1$ and $m > 0$), $\Theta(\eta)$ must eventually be 0 at large η, since both the diamond and square extend only to finite η. In regions 1-3, as η increases, $\Theta(\eta)$ is first (non-strictly) increasing from 0, and then decreases monotonically to 0, remaining 0 for $\eta > \eta_{N,m}$ (consider the "boundary conditions" $\Theta(0) = \Theta(\infty) = 0$ and the signs of $\Theta'(\eta)$ associated with the shading in Figure B.3b). This implies that $\Theta(\eta)$ remains non-negative for all η.

In regions 4 and 5, the argument is more subtle. Here $\Theta(\eta)$ initially increases (non-strictly) for $\eta < \eta_{N,\infty}$, then decreases (again, non-strictly) for $\eta \in (\eta_{N,\infty}, \eta_{N,m})$, and then again increases until $\eta = \eta_{E,\infty}$, where $\Theta = 0$. Therefore $\Theta(\eta)$ is initially 0, becomes positive for some interval of η, and then crosses 0, becoming negative before returning to 0 (as in e.g. Figure B.2c). The same is true of $\eta\,\Theta(\eta)$. Most importantly, the interval where $\Theta(\eta)$ and $\eta\,\Theta(\eta)$ are negative is on the rightmost end of their support. At the same time, for any $0 < \eta_1 < \eta_2$, ${}^{g(\eta_2;\sigma)}/_{g(\eta_1;\sigma)}$ is strictly increasing as a function σ. This means that, in the integral in (B.72), the negative part of $\eta\,\Theta(\eta)$ has the greatest relative contribution in the limit $\sigma \to \infty$. It is therefore sufficient that $\int_0^\infty \eta\,\Theta(\eta)\,d\eta \geq 0$ in order to show that (B.72) holds. But the integral $\int_0^\infty \eta\,\Theta(\eta)\,d\eta$ is just the difference in area between the diamond and the square, $\frac{1}{\sqrt{1-\mu^2}} - 1$, which is positive for $0 < \mu < 1$.

We conclude that $\mathcal{E}_m - \mathcal{E}_\infty > 0$ for any $\Lambda \in (0,1)$, $m > 0$, and $\delta \in \left(0, \frac{1}{2}\right)$.

B.4.2. The Upper Bound

We recall our goal, a bound $\mathcal{B}$ independent of $\mu = \Lambda^m$ such that (cf. (B.68) and (B.67))

$$\frac{\mathcal{E}_m - \mathcal{E}_\infty}{\mu} = \frac{1}{2\pi\sigma^2\mu}\int_0^\infty \eta\,\Theta(\eta;\delta,\mu)\,g(\eta;\sigma)\,d\eta \leq \mathcal{B}(\sigma;\delta),$$

and we therefore aim to select an appropriate upper bound of $\frac{1}{\mu}\Theta(\eta;\delta,\mu)$ uniform in μ. However, many of the obvious (e.g. piecewise constant) upper bounds on Θ result in divergences in the above integral as $\mu \to 1$. The difficulty is that $\eta_{N,m} \to \infty$ as $\mu \to 1$. Indeed, the limiting shape of the diamond region is the infinite strip $\left[\delta - \frac{1}{2}, \delta + \frac{1}{2}\right] \times (-\infty, \infty)$, and so Θ remains nonzero for arbitrarily large values of η. We find it necessary to incorporate the behavior in the limit as $\mu \to 1$ into our bound. In particular, we will

demonstrate in the following that

$$\frac{1}{\mu}\Theta\left(\eta;\delta,\mu\right)\leq\Theta\left(\eta;\delta,\mu=1\right). \tag{B.73}$$

(B.73) in turn gives $\mathcal{B}=\mathcal{E}\left(1-\mathcal{E}\right)$, since $\mathcal{E}_{\infty}=\mathcal{E}^{2}$ and $\mathcal{E}_{m}\rightarrow\mathcal{E}$ as $\mu\rightarrow 1$.

To demonstrate that our bound (B.73) holds true, we first recall that $\Theta\left(\eta;\delta,\mu\right)$ is either 0 or negative for $\eta>\eta_{N,m}\left(\delta,\mu\right)$ and $\eta_{N,m}$ increases as a function of μ (Figure B.3b). Since $\Theta\left(\eta;\delta,\mu=1\right)$ is non-negative for all η, the bound holds trivially for $\eta>\eta_{N,m}\left(\delta,\mu\right)$. For $\eta\leq\eta_{N,m}$, consider the decomposition $\Theta=\Theta_{ENE}+\Theta_{NNE}+\Theta_{NNW}+\Theta_{WNW}$, where for $d=E,W$

$$\begin{aligned}\Theta_{NNd}\left(\eta;\delta,\mu\right) &\equiv \begin{cases}0, & \eta\leq\eta_{N,\infty}\\ -\int_{\eta_{N,\infty}}^{\eta}\Theta'_{Nd}\left(\eta\right)d\eta, & \eta_{N,\infty}<\eta\leq\eta_{N,m}\\ \Theta_{Nd}\left(\eta_{N,\infty}\right)-\Theta_{Nd}\left(\eta_{N,m}\right), & \eta>\eta_{N,m}\end{cases}\\ &= \Theta_{Nd}\left(\min\left(\eta,\eta_{N,\infty}\left(\delta\right)\right);\delta\right)-\Theta_{Nd}\left(\min\left(\eta,\eta_{N,m}\left(\delta,\mu\right)\right);\delta\right) \qquad \text{(B.74)}\end{aligned}$$

and $\Theta_{NNE}+\Theta_{NNW}$ reflects the "$-$NE$-$NW" contribution that, as a function of η, monotonically decreases in the red region in Figure B.3b and otherwise remains constant. Similarly,

$$\Theta_{dNd}\left(\eta;\delta,\mu\right)\equiv\Theta_{Nd}\left(\min\left(\eta,\eta_{d,\infty}\left(\delta\right)\right);\delta\right)-\Theta_{Nd}\left(\min\left(\eta,\eta_{d,m}\left(\delta,\mu\right)\right);\delta\right) \tag{B.75}$$

reflects for $d=E$ the blue "NE" and for $d=W$ the yellow "NW" contributions. It is sufficient to show that $\frac{1}{\mu}\Theta_{NNd}\left(\eta;\delta,\mu\right)\leq\Theta_{NNd}\left(\eta;\delta,1\right)$ and $\frac{1}{\mu}\Theta_{dNd}\left(\eta;\delta,\mu\right)\leq\Theta_{dNd}\left(\eta;\delta,1\right)$ for $d=E,W$ and $\eta\leq\eta_{N,m}\left(\delta,\mu\right)$ to validate (B.73). The first of these is trivial,

since $\Theta_{NNd}(\eta;\delta,\mu)$ is independent of μ and non-positive for $\eta \leq \eta_{N,m}(\delta,\mu)$, and thus $\frac{1}{\mu}\Theta_{NNd}(\eta;\delta,\mu) \leq \Theta_{NNd}(\eta;\delta,\mu) = \Theta_{NNd}(\eta;\delta,1)$. The latter follows if $h(x;\mu) \leq h(x;1)$ for all $x \geq 0$ and $\mu \in (0,1)$, where (cf. (B.69), (B.70), (B.71), and (B.75))

$$h(x;\mu) \equiv \frac{1}{\mu}\left[\operatorname{arcsec}\left(\min\left(x,\sqrt{2}\right)\right) - \operatorname{arcsec}\left(\min\left(x,\sqrt{\frac{2}{1+\mu}}\right)\right)\right].$$

Note that $h'(x;\mu)$ is non-negative and only depends on μ by a factor of $\frac{1}{\mu}$. So, for any given value of x, it is minimized as $\mu \to 1$. It is therefore sufficient to verify the inequality $h(x;\mu) \leq h(x;1)$ in the limit as $x \to \infty$; if $h(\infty;\mu) \leq h(\infty;1)$,

$$\begin{aligned} h(x;\mu) &= h(\infty;\mu) - \int_x^\infty h'(u;\mu)\,du \\ &\leq h(\infty;1) - \int_{x_0}^\infty h'(u;1)\,du \\ &= h(x;1) \end{aligned}$$

for all $x \geq 0$. So we verify that

$$h(\infty;\mu) = \frac{1}{\mu}\left[\frac{\pi}{4} - \operatorname{arcsec}\left(\sqrt{\frac{2}{1+\mu}}\right)\right] \leq \frac{\pi}{4} = h(\infty;1).$$

This is equivalent to

$$\sin\left(\frac{\pi}{4}(1-\mu)\right)^2 \leq \frac{1-\mu}{2},$$

which is satisfied for $\mu \in (0,1)$, since equality is obtained at $\mu = 0, 1$, $\frac{1-\mu}{2}$ is linear, and $\sin\left(\frac{\pi}{4}(1-\mu)\right)^2$ is a concave-up function of μ on the open interval.

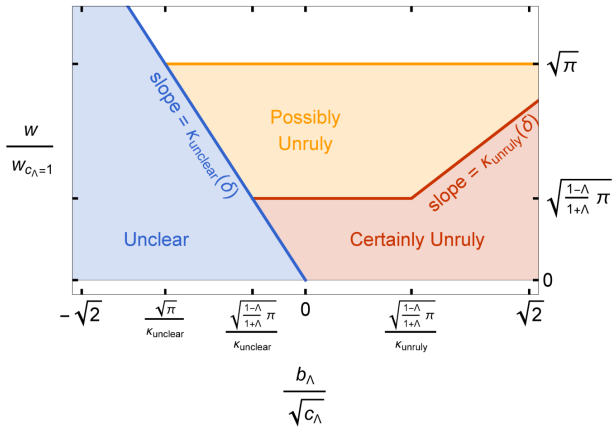

Figure B.4. An alternative representation of the phase diagram, Figure 3.8a, showing the dependence of unruliness on the event interval width w.

B.5. Dependence on Event Interval Width

In this section, we derive the relationship between the event interval width $|E| = w$ and unruliness in planar oscillators. From Figure 3.8a, we see that changing w with a fixed value of $\frac{b_\Lambda}{\sqrt{c_\Lambda}}$ moves the oscillator between the regions of possible unruliness, certain unruliness and uncertainty. We therefore first translate the b_Λ-c_Λ phase diagram (Figure 3.8a) into a phase diagram over $\frac{b_\Lambda}{\sqrt{c_\Lambda}}$ and w (Figure B.4). It is useful to normalize w by a characteristic width, the value of w that is required to fix $c_\Lambda = 1$: from (3.33) and (3.59),

$$w_{c_\Lambda=1} = \sqrt{\frac{1+\Lambda}{1-\Lambda}\frac{\Gamma_{ss}}{c}} = \frac{1}{\sqrt{2}\,(1-\Lambda)}\frac{\|\vec{y}_G\|}{\|\vec{z}_G\|}.$$

It then turns out that the linear boundaries from the b_Λ-c_Λ phase diagram remain linear and with the same slope in the $\frac{b_\Lambda}{\sqrt{c_\Lambda}}$-$\frac{w}{w_{c_\Lambda=1}}$ phase diagram. Referring to Figure B.4, we see the following:

i) If $\frac{b_\Lambda}{\sqrt{c_\Lambda}} > 0$, the TVGR is not unruly for $w > w^*_{possible}$, is possibly unruly for $w \in \left(w^*_{certain}, w^*_{possible}\right)$, and is certainly unruly for $w \in (0, w^*_{certain})$, where

$$w^*_{possible} = \sqrt{\pi} w_{c_\Lambda=1} = \sqrt{\frac{\pi}{2}} \frac{1}{(1-\Lambda)} \frac{\|\vec{y}_G\|}{\|\vec{z}_G\|}$$

ands

$$\begin{aligned} w^*_{certain} &= \max\left(\sqrt{\frac{1-\Lambda}{1+\Lambda}}\pi, \kappa_{unruly}(\delta) \frac{b_\Lambda}{\sqrt{c_\Lambda}}\right) w_{c_\Lambda=1} \\ &= \max\left(\sqrt{\frac{\pi}{2(1-\Lambda^2)}} \frac{\|\vec{y}_G\|}{\|\vec{z}_G\|}, \frac{\kappa_{unruly}(\delta)}{1-\Lambda} \frac{\vec{z}_G^T \vec{y}_G}{\|\vec{z}_G\|^2}\right). \end{aligned}$$

ii) If $\frac{1}{\kappa_{unclear}(\delta)}\sqrt{\frac{1-\Lambda}{1+\Lambda}\pi} < \frac{b_\Lambda}{\sqrt{c_\Lambda}} < 0$, the TVGR is possibly unruly for $w \in \left(w^*_{certain}, w^*_{possible}\right)$ and is certainly unruly for $w \in (w^*_{unclear}, w^*_{certain})$, where

$$w^*_{unclear} = \kappa_{unclear}(\delta) \frac{b_\Lambda}{\sqrt{c_\Lambda}} w_{c_\Lambda=1} = \frac{\kappa_{unclear}(\delta)}{1-\Lambda} \frac{\vec{z}_G^T \vec{y}_G}{\|\vec{z}_G\|^2}.$$

iii) If $\frac{1}{\kappa_{unclear}(\delta)}\sqrt{\pi} < \frac{b_\Lambda}{\sqrt{c_\Lambda}} < \frac{1}{\kappa_{unclear}(\delta)}\sqrt{\frac{1-\Lambda}{1+\Lambda}\pi}$, the TVGR is possibly unruly for $w \in \left(w^*_{unclear}, w^*_{possible}\right)$, but the situation is unclear for $w < w^*_{unclear}$.

iv) If $\frac{b_\Lambda}{\sqrt{c_\Lambda}} < \frac{1}{\kappa_{unclear}(\delta)}\sqrt{\pi}$, the TVGR is not unruly for $w > w^*_{unclear}$ and the situation is unclear for $w < w^*_{unclear}$.