Counterdiabatic driving in the classical β-Fermi-Pasta-Ulam-Tsingou chain

Nik O. Gjonbalaj⬤, David K. Campbell⬤, and Anatoli Polkovnikov⬤
Department of Physics, Boston University, Boston, Massachusetts 02215, USA

Shortcuts to adiabaticity (STAs) have been used to make rapid changes to a system while eliminating or minimizing excitations in the system's state. In quantum systems, these shortcuts allow us to minimize inefficiencies and heating in experiments and quantum computing protocols, but the theory of STAs can also be generalized to classical systems. We focus on one such STA, approximate counterdiabatic (ACD) driving, and numerically compare its performance in two classical systems: a quartic anharmonic oscillator and the β Fermi-Pasta-Ulam-Tsingou lattice. In particular, we modify an existing variational technique to optimize the approximate driving and then develop classical figures of merit to quantify the performance of the driving. We find that relatively simple forms for the ACD driving can dramatically suppress excitations regardless of system size. ACD driving in classical nonlinear oscillators could have many applications, from minimizing heating in bosonic gases to finding optimal local dressing protocols in interacting field theories.

I. INTRODUCTION

Losses and noise represent a ubiquitous challenge in the control and use of any physical system. Whether we are trying to prepare a high-fidelity quantum state for some calculation, maximizing the efficiency of some thermodynamic process, or tuning some parameter of a sensitive Hamiltonian, transitions away from the desired outcome are inevitable. Adiabatic evolution represents the most direct method for combating these inefficiencies, as asymptotically slow evolution removes all diabatic excitations and allows us to perform arbitrary physical processes without losses. However, during the resulting long time scales, the system suffers from environmental noise and decoherence, making true adiabatic transport not only costly but impossible to implement. This tension has led to the development of many "shortcuts to adiabaticity" (STAs) [1–4] that achieve exactly or nearly the same effect as adiabatic transport in a fraction of the time, thereby limiting the effect of outside forces and increasing the speed of the process. These methods have shown great promise theoretically and experimentally and apply to nearly any system where high-fidelity control is desired.

One such STA that has found application in many different systems is counterdiabatic (CD) or transitionless driving [5–8]. This method adds to the bare Hamiltonian a driving counterterm constructed out of the "adiabatic gauge potential" (AGP). In quantum geometry, this object reveals how eigenstates are connected to each other as the control parameter in the Hamiltonian defining the adiabatic transformation is tuned. This counterterm kills off all diabatic transitions exactly for any tuning speed. In simple systems, CD driving offers a powerful method for exactly transporting eigenstates as the Hamiltonian is altered, but in most many-body systems this approach breaks down, and the AGP generically does not exist as a local continuous operator. Recently, however,

Ref. [9] found that in many cases an approximate local variational AGP exists, which allows one to significantly suppress dissipative effects even in chaotic systems. This method has been extensively applied to quantum spin systems with many promising results [10–15]. Moreover, Ref. [14] showed that this ACD protocol can be realized using Floquet engineering with only the Hamiltonian and the control field.

Although ACD driving can easily be generalized to classical spins, the method outlined in the original paper [9] does not apply to systems with unbounded local Hilbert spaces. As such, the original procedure is limited in scope to spin chains and fermionic systems, leaving out coupled oscillators and collections of bosons. While exact CD driving (and other related STAs like fast-forward driving) has been implemented in potential wells before [16–18], the variational method naively breaks down. In this work, we fill in the gaps and generalize ACD driving to systems with unbounded Hilbert spaces of any dimension. In particular, we focus on coupled nonlinear oscillators, where the classical phase space and local quantum Hilbert space are unbounded. Our formalism and the developed protocols apply equally to quantum and classical systems. However, in order to verify the performance of the ACD protocol, we have to rely on numerical simulations, which are currently out of reach for quantum systems with many degrees of freedom. In small systems, we are able to show that quantum and classical ACD protocols perform very similarly. We begin by working with a single anharmonic oscillator to explain the procedure before moving to a more complicated system: the β Fermi-Pasta-Ulam-Tsingou (FPUT) lattice. This nonlinear chain not only provides a simple many-body Hamiltonian to study but also has many deep connections to statistical mechanics and thermalization. Originally, Fermi, Pasta, Ulam, and Tsingou investigated this system numerically with the expectation that the nonlinearity would cause the chain to thermalize quickly

and reach equipartition [19]. However, when they initialized the system with all the energy in the first normal mode, they famously discovered that almost all the energy returned to this initial mode in what are now called FPUT "recurrences." In contradiction with the expectation of equipartition, these recurrences continue as the simulation progresses, keeping the system in a "metastable state" before finally thermalizing after very long time scales [20–22]. Recently, it was discovered that recurrences themselves oscillate in a periodic fashion as well, a feature dubbed "higher-order recurrences" [23,24].

This intimate connection between the FPUT system, thermalization, and chaotic dynamics also motivates an application of ACD driving. Indeed, it was recently shown that the AGP can act as an extremely sensitive probe of chaos in certain quantum systems [25]. Developing approximate AGPs for the β FPUT lattice in future work could therefore reveal why the system remains quasiperiodic for certain initial conditions. In addition, FPUT-like recurrences come up in many theoretical [26–29] and experimental [30,31] models. ACD driving can be used to efficiently prepare such systems in different nearly stationary initial conditions, where these recurrences are expected to be strongly suppressed, to study the effects of long-time thermalization.

The paper is organized as follows. We first introduce the concept of ACD driving in Sec. II and discuss its subtleties before explaining our modifications to the variational method. We then apply the procedure to a simple anharmonic oscillator in Sec. III as an example implementation. In Sec. IV, we apply the same machinery to the β-FPUT lattice and show how our results survive in the thermodynamic limit. Finally, we discuss instabilities that arise when one considers long-wavelength initial states like those in the original FPUT problem. We argue that these instabilities arise because the corresponding initial distribution is far from any statistical equilibrium and is very fragile against adding nonlinearity to the system. Nevertheless, even in this situation we provide a method for combating this instability by adding fluctuations to the initial state and variationally optimizing the AGP over a broader range of phase space than the actual initial distribution. We show that in this way we can stabilize the variational method, albeit at the expense of reducing its accuracy somewhat.

II. APPROXIMATE COUNTER-DIABATIC DRIVING

The concept of CD driving is most easily explained in a quantum sense, so we will begin there and then move to the classical regime. Consider a Hamiltonian $H_0(\beta(t))$ dependent on some parameter $\beta(t)$ that varies in time, and let the system start in some stationary state $|\psi_i\rangle$ at time $t = 0$. We now want to ramp β from some initial value $\beta(t = 0) = \beta_i$ to some final value $\beta(t = \tau) = \beta_f$ over some "turn-on time" τ. Assuming there are no degeneracies on the way, if we take an infinite time to ramp up the perturbation ($\tau \to \infty$), the system will undergo adiabatic evolution, finishing in a new stationary state $|\psi_{\text{target}}\rangle$ of the final Hamiltonian $H_0(\beta(\tau))$. However, for any finite value of τ, the system will inevitably experience diabatic transitions away from this state. The origin of these transitions is easy to understand by going to the moving frame defined by the time-dependent instantaneous basis of $H_0(\beta)$. Here, the effective moving Hamiltonian picks up an extra term

familiar, e.g., from doing a Galilean transformation or going to a rotating frame:

$$H_0^{\text{eff}} = H_0 - \dot{\beta}\mathcal{A}_\beta, \tag{1}$$

where $\mathcal{A}_\beta$ is the adiabatic gauge potential (AGP), the generator of perturbations in β. Because the Hamiltonian H_0 is diagonal in its own instantaneous basis, all transitions from our initial stationary state occur due to the presence of the second term. Exact CD driving provides a method for exactly reaching our target state $|\psi_{\text{target}}\rangle$ by evolving under a CD Hamiltonian:

$$H_{\text{CD}}(t) = H_0(\beta(t)) + \dot{\beta}(t)\mathcal{A}_\beta(\beta(t)). \tag{2}$$

In the moving frame, we can see that the two terms with the AGP will cancel, making the effective quantum Hamiltonian fully diagonal and eliminating transitions. It is straightforward to show that, as the generator conjugate to changes in β, the AGP satisfies [10]

$$\langle m|\mathcal{A}_\beta|n\rangle = i\hbar\frac{\langle m|\partial_\beta H_0|n\rangle}{E_n - E_m}, \tag{3}$$

or, equivalently,

$$[H_0, i\hbar\partial_\beta H_0 - [\mathcal{A}_\beta, H_0]] = 0. \tag{4}$$

In simple systems, we can solve Eq. (4) to find the exact AGP and implement transitionless driving. However, one can show that in general, the AGP blows up exponentially with the system size in chaotic systems [25]. More specifically, quantum chaotic systems suffer from small denominators in Eq. (3) as a result of the eigenstate thermalization hypothesis (ETH) [10]. Intuitively, this blowup occurs because we are trying to follow eigenstates that, according to the ETH, are essentially random vectors and therefore require fine-tuning. Instead of exactly canceling off transitions for this exponentially difficult cost, we want to suppress transitions as much as possible, i.e., maximize the fidelity of our evolved state with the target state.

Reference [9] outlines a method for finding an approximate form of the AGP in what is known as approximate CD (ACD) driving. First, we choose some *Ansatz* for the form of the AGP, ideally one that does not suffer from the nonlocal terms typically present in chaotic systems [10,25]. Once we have our *Ansatz*, we can compute an action that will optimize the free parameters of the AGP *Ansatz* as functions of β. Consider the object

$$G_\beta(\mathcal{A}_\beta^*) \equiv \partial_\beta H_0 + \frac{i}{\hbar}[\mathcal{A}_\beta^*, H_0], \tag{5}$$

where $\mathcal{A}_\beta^*$ is our *Ansatz*. It can be shown that Eq. (4) follows from minimizing the following action $\mathcal{S}$ with respect to the all possible operators $\mathcal{A}_\beta^*$ [10]:

$$\mathcal{S}(\mathcal{A}_\beta^*) = \langle G_\beta^2(\mathcal{A}_\beta^*)\rangle - \langle G_\beta(\mathcal{A}_\beta^*)\rangle^2, \tag{6}$$

where the angular brackets denote an average with respect to some equilibrium density matrix ρ commuting with the Hamiltonian H_0. If the variational *Ansatz* is complete, i.e., if it spans the operator basis, then this minimization leads to the exact AGP. If the variational manifold is restricted to, for example, a class of local operators, the minimization leads to the best variational AGP within this class. One can use flexibility in choosing the equilibrium density matrix ρ to

target specific states of interest. For example, $\rho = |\psi_0\rangle\langle\psi_0|$ optimizes the AGP with respect to the ground state, which could be beneficial for quantum annealing protocols. A different choice of $\rho = I = \frac{1}{D}\sum_n |n\rangle\langle n|$, where D is the Hilbert space size, optimizes the AGP with respect to all states. Other choices of ρ can include, e.g., finite temperature Gibbs ensembles or microcanonical ensembles targeting states in a specific energy range. Once we fix the form of ρ, we find our optimal approximate AGP by extremizing the action in the usual sense:

$$\frac{\delta\mathcal{S}(\mathcal{A}_\beta^*)}{\delta\mathcal{A}_\beta^*} = 0. \tag{7}$$

If our operator $\mathcal{A}_\beta^*$ is expanded in a set of some fixed basis operators X_j: $\mathcal{A}_\beta^* = \sum_j \gamma_j X_j$, then variational optimization with respect to $\mathcal{A}_\beta^*$ is equivalent to minimization with respect to the set of parameters γ_j. Note that by construction, the action is quadratic in γ_j, so this minimization is simply a minimization of a quadratic form. After finding the variational AGP, we can then construct the approximate CD Hamiltonian

$$H_{\mathrm{ACD}}(t) = H_0 + \dot{\beta}\mathcal{A}_\beta^*, \tag{8}$$

which reduces transitions compared to the bare protocol.

From quantum spins to classical oscillators

Previous work on ACD driving has focused on the infinite-temperature norm, where the average in Eq. (6) is taken with respect to the identity matrix and focused mostly on quantum spin systems [9,10,14,32]. Such an infinite temperature optimization is ill-defined in systems like the FPUT model, where the local Hilbert space is unbounded and the infinite temperature state corresponds to infinite energy density. To proceed, we therefore have to choose a different ρ. As we are dealing with classical systems, let us also briefly discuss how the variational approach extends to them from the quantum regime.

In classical chaotic systems, exact CD driving is also impossible due to formal divergence of the AGP [33]. As such, we will turn to ACD driving for our classical systems. To make this shift, most of the standard relations between quantum and classical are involved: operators become functions and commutators become Poisson brackets, so now Eq. (5) becomes

$$G_\beta(\mathcal{A}_\beta^*) \equiv \partial_\beta H_0 - \{\mathcal{A}_\beta^*, H_0\}. \tag{9}$$

The classical analog of the equilibrium density matrix ρ is a stationary probability distribution $P(\vec{q}, \vec{p})$, where $\vec{q}$, $\vec{p}$ are canonically conjugate phase space variables. The approximate AGP can again be found from the minimization of the action (6), where the average is now taken with respect to the probability distribution P_β, i.e.,

$$\langle G_\beta \rangle = \int P(\vec{q}, \vec{p}, \beta)\, G_\beta(\vec{q}, \vec{p})\, D\vec{q}D\vec{p}, \tag{10}$$

and similarly for $\langle G_\beta^2 \rangle$, where $D\vec{q}D\vec{p}$ denotes a differential volume of phase space. There is the same flexibility in choosing the stationary probability distribution $P(\vec{q}, \vec{p})$ in classical systems as in choosing a stationary density matrix ρ in quantum systems. In particular, a single quantum eigenstate averaging

in a chaotic system corresponds to averaging over a microcanonical ensemble: $P_\beta(\vec{q}, \vec{p}) \propto \delta(E - H(\vec{q}, \vec{p}, \beta))$; similarly a finite-temperature Gibbs density matrix corresponds to a finite-temperature classical Gibbs probability distribution $P_\beta(\vec{q}, \vec{p}) \propto \exp[-H(\vec{q}, \vec{p}, \beta)/(k_b T)]$.

It is intuitively clear that the optimal choice of the probability distribution $P(\vec{q}, \vec{p})$ should be closest to the one adiabatically connected to the initial state. That is, if we initialize the system in some microcanonical ensemble, then P_β should remain a microcanonical distribution with respect to the instantaneous Hamiltonian $H(\vec{q}, \vec{p}, \beta)$ at the energy corresponding to constant initial entropy. The easiest way to generate such a distribution, which we use in this work, is by slowly evolving the initial distribution in time to some coupling β between initial and final values. We then sample from this distribution to minimize the action (6) and find the approximate AGP. This procedure does not require finding the adiabatically connected distribution P_β, and it has another clear advantage in that it applies both to integrable and nonintegrable systems, as it does not require the system to thermalize. In principle, one can repeat this procedure iteratively: evolve the initial state with the ACD protocol to find a better adiabatically connected state and reminimize the action to find a better AGP. We find that this reoptimization is not necessary, and even the first iteration gives excellent results.

III. SINGLE ANHARMONIC OSCILLATOR

To illustrate the approach, we first start with a single anharmonic oscillator with a Hamiltonian given by

$$H_0 = \frac{p^2}{2} + \frac{x^2}{2} + \beta\frac{x^4}{4}, \tag{11}$$

where units have been chosen to set the mass and frequency to 1. Because this one-dimensional (1D) system is integrable, it is amenable to exact CD driving, and progress has already been made on this front for both scale-invariant systems [18] and using the Lax pair formalism familiar from classical nonlinear integrable systems [34]. Other STAs like fast-forward driving have seen success in anharmonic traps as well [16,17]. The work here differs in that variational ACD driving allows us to drive any trapping potential, target any manifold of states without knowing exact wave functions, and also drive under any sort of ramp protocol $\beta(t)$. However, even beyond these differences, the following section simply functions as an illustrative example before we dive into the complex dynamics of FPUT.

We initialize our system with a value of $\beta(t = 0) = 0$ in a microcanonical distribution of energy $E_0 = 1$. We then ramp up to $\beta(t = \tau) = 1$ using the protocol

$$\beta(t) = \sin^2\left[\frac{\pi}{2}\sin^2\left(\frac{\pi t}{2\tau}\right)\right], \quad t \in [0, \tau] \tag{12}$$

from Ref. [9]. This protocol is smooth at the boundaries, suppressing spurious nonadiabatic transitions related to turning on and turning off the drive, which appear if $\dot{\beta}$ jumps at the protocol boundaries. Under perfect adiabatic evolution ($\tau \to \infty$), the system would end in a microcanonical distribution of the new Hamiltonian. In the quantum language, this new state would have the same "eigennumber" as the original.

Any finite turn-on time τ will cause transitions away from this transport. Our goal is to minimize these transitions for all τ. In passing, let us note that for single-particle systems an alternative approximate approach based on classical flow field was recently developed in Refs. [17,35].

A. Ansatz

First, we must choose an *Ansatz* for the form of our driving term. Following Ref. [14], we can take a series expansion of the AGP, truncate the series at a finite order, and then insert this into our variational procedure. In the classical limit, this *Ansatz* takes the form

$$\mathcal{A}_\beta^{(\ell)} = \sum_{k=1}^{\ell} (-1)^k \gamma_k \underbrace{\{H_0, \{H_0, \dots \{H_0}_{2k-1}, \partial_\beta H_0\}\}\}, \qquad (13)$$

where ℓ is the order at which we truncate the expansion, γ_k are the variational parameters that we will optimize as functions of β, and the factor $(-1)^k$ is introduced for convenience. Here we will only consider *Ansätzes* with one and two driving parameters, so we have

$$\mathcal{A}_\beta^{(1)} = \gamma_1 x^3 p,$$
$$\mathcal{A}_\beta^{(2)} = \gamma_1 x^3 p + \gamma_2 (-6p^3 x + 10px^3 + 12\beta px^5). \qquad (14)$$

Notice that the last term in the expression for $\mathcal{A}_\beta^{(2)}$ is proportional to β and vanishes in the limit $\beta \to 0$. Therefore, it is not very important at least at small values β, and we will exclude it from the variational *Ansatz*. The px^3 term in $\mathcal{A}_\beta^{(2)}$ is identical to the one from $\mathcal{A}_\beta^{(1)}$, so they can be merged together. Therefore for the remainder of the paper we will refer to our first- and second-order *Ansätzes* as

$$\mathcal{A}_1^* \equiv \tilde{\gamma}_1 x^3 p,$$
$$\mathcal{A}_2^* \equiv \tilde{\gamma}_1 x^3 p + \tilde{\gamma}_2 x p^3, \qquad (15)$$

where $\tilde{\gamma}_1$ and $\tilde{\gamma}_2$ are our new variational parameters.

Let us point out that there is an alternative way to justify this variational *Ansatz* by finding the AGP in the harmonic limit $\beta \to 0$. In Refs. [33,36] it is shown that AGP can be represented through the following time integral:

$$\mathcal{A}_\beta = -\lim_{\epsilon \to 0^+} \frac{1}{2} \int_{-\infty}^{\infty} dt \, \mathrm{sgn}(t) e^{-\epsilon|t|} (\partial_\beta H_0)(t), \qquad (16)$$

where in quantum language $(\partial_\beta H_0)(t)$ is the Heisenberg representation of the operator $\partial_\beta H_0$, which in the classical language translates to the function $\partial_\beta H_0$ evaluated on time-dependent trajectories, which are solutions to the equations of motion (EOM) $(x(t), p(t))$ for the full nonlinear Hamiltonian. In the leading order in perturbation theory, we can use the solutions of the noninteracting problem $\beta = 0$:

$$x(t) = x \cos t + p \sin t,$$
$$p(t) = p \cos t - x \sin t, \qquad (17)$$

where x and p correspond to time-independent Schrödinger operators in the quantum language. In this limit, Eq. (16) yields the following form for the AGP:

$$\mathcal{A}_\beta \to -\frac{5}{32} x^3 p - \frac{3}{32} x p^3. \qquad (18)$$

We thus see that the exact AGP at $\beta \to 0$ is described by the *Ansatz* (15) with a particular choice of the coefficients $\tilde{\gamma}_1$ and $\tilde{\gamma}_2$. At $\beta > 0$ this *Ansatz* is no longer exact, but we can still variationally optimize $\tilde{\gamma}_1$ and $\tilde{\gamma}_2$ hoping that the resulting AGP will accurately describe the nonlinear system. We checked that both *Ansätzes* (14) and (15) lead to a similar performance, so the extra px^5 appearing in Eq. (14) has a very small effect on the ACD protocol.

B. Optimization

As we discussed above, in order to perform optimization of the variational parameters, we slowly evolve the initial distribution corresponding to a microcanonical ensemble at $\beta = 0$ to a finite value of nonlinearity and then minimize the action (6), sampling from this probability distribution. In quantum language, this procedure corresponds to sampling from an instantaneous energy state adiabatically connected to the initial state of the harmonic oscillator. We discuss details of this procedure and show the obtained functional forms of the variational coefficients $\gamma_i(\beta)$ in Appendix A.

C. Figures of merit

Once we simulate the driven system, we need a method for evaluating the performance of the given *Ansatz*. In quantum systems, the golden standard is the fidelity of the final state with the target state. This figure of merit (FOM) tells us exactly how close we are to achieving adiabatic transport, as it ensures not only that we approach a stationary state but also that it is the *correct* stationary state. In the classical regime, this task becomes much harder. In principle, we could evaluate some distance such as the KL-divergence between the obtained and target probability distributions. But finding this distance is not easy, especially when we extend this analysis to systems with many degrees of freedom. We find that there is a much simpler measure, which is the energy variance of the final Hamiltonian in the obtained distribution. It works both in quantum and classical setups and is easy to evaluate. If we start in an exact microcanonical state and evolve it adiabatically, we should also end up in a microcanonical state such that the energy variance is zero. Similarly, in quantum mechanics if we start from a single eigenstate and adiabatically evolve in time we should remain in a single eigenstate and the energy variance is again zero. In both cases, nonzero energy variance indicates that the evolution was not adiabatic. For the remainder of this paper, we will focus on this FOM. We also checked other FOMs like fidelity or temporal fluctuations of observables and found that they contain similar information (see Appendix B for more details.)

D. Results

In Fig. 1, we show the results of implementing the ACD protocols with the optimized AGPs and compare their performance against a bare, unassisted ramp for various values of the turn-on time τ. In these simulations, we initialize our system in a microcanonical distribution of $E_0 = 1$ and evolve while β grows according to the protocol (12).

Note that, as expected, all turn-on curves interpolate between some asymptotic value at small τ and zero at large τ.
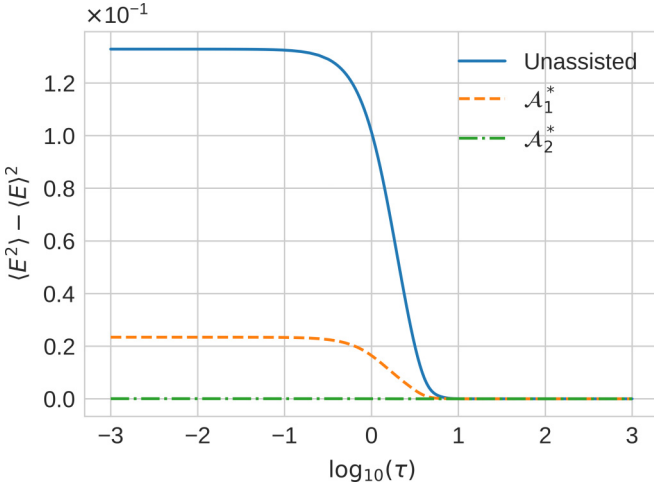
FIG. 1. Final energy variance after ramps of β of various durations τ for a single anharmonic oscillator. Orange (green) lines correspond to first- (second-) order CD driving protocols. In the limit of $\tau \to 0$, which corresponds to driving only with the AGP, the first-order CD protocol gives the energy variance suppression of the order of 5.7 compared to the unassisted (quench) protocol; the second-order CD protocol increases this suppression to a factor of ~ 2700.

All protocols in the limit $\tau \gg 1$ give zero energy variance in agreement with the adiabatic theorem. In the opposite $\tau \to 0$ limit, which corresponds to the instantaneous (quenched) protocol for unassisted driving and the protocols where one drives the system only with the AGP for ACD driving, we see the largest difference between the three protocols. In particular, in this limit the unassisted protocol gives the largest energy variance, and the first-order ACD suppresses the energy variance by an approximate factor of 5.7, which pales in comparison to the performance of the second-order ACD, which suppresses the energy variance by a factor of 2700. We thus see that the addition of the rather unusual xp^3 to the Hamiltonian yields a dramatic improvement in its performance. The difference becomes even larger if we analyze the quantum fidelity of the state preparation or the temporal variance of the harmonic energy as FOMs (see Appendix B). Incidentally, this xp^3 term also appears in the exact CD driving of a potential in the shape of a KdV soliton moving in a gas of noninteracting quantum particles [10,37], and Ref. [34] also encounters a CD term like p^3. Clearly, the xp^3-term plays a crucial role in creating the efficient CD driving protocol. However, this presents a unique challenge when implementing this protocol in experiments, as such terms do not naturally appear in typical Hamiltonians. These terms can be, however, designed using periodic lattice potentials, where the dispersion relation generally contains $\sin(p)$ type terms, where p is the lattice momentum.

E. Floquet implementation

Generally the expansion (13) can be implemented through the experimentally accessible Floquet protocols [14]. Such protocols only involve H_0 and $\partial_\beta H_0$ as driving terms. Specifically choosing

$$H_{FE}(t) = f(t)H_0 + g(t)\partial_\beta H_0, \quad (19)$$

where $f(t)$ and $g(t)$ are periodic functions in time, one can always realize a CD protocol with the AGP given by Eq. (13) with arbitrary time-dependent coefficients γ_k. If the function $f(t)$ is positive at any time, then one can further simplify the protocol by rescaling time, i.e., dividing the Schrödinger equation $i\hbar\partial_t\psi = H_{FE}(t)\psi$ by $f(t)$. One can define a new "physical" time as

$$\tau(t) = \int_0^t f(t')dt'. \quad (20)$$

Then the Schrödinger equation becomes

$$i\hbar\partial_\tau\psi = \tilde{H}_{FE}(\tau)\psi, \quad \tilde{H}_{FE}(\tau) = H_0 + \kappa(\tau)\partial_\beta H_0, \quad (21)$$

where t is now a parameter, which depends on the physical protocol time τ through inverting Eq. (20) and $\kappa(\tau) = g(t(\tau))/f(t(\tau))$. Thus, we can implement the Floquet protocol by only varying the coupling $\partial_\beta H_0$ in time.

Note that generally the Floquet protocols found in Ref. [14] require nonpositive $f(t)$ and the mapping (20) is not possible without inverting the sign of H_0. Finding a class of general protocols with $f(t) > 0$ remains an open problem. In Ref. [38] it was solved for a particular setup leading to a finite speed limit.

IV. β-FPUT CHAIN

Now that we have a method for implementing ACD driving in classical oscillators, we want to apply it to the more complex β-FPUT chain. The Hamiltonian for this system is given by

$$H_0 = \sum_{n=1}^{N} \frac{p_n^2}{2} + \sum_{n=0}^{N} \frac{1}{2}(q_{n+1} - q_n)^2 + \frac{\beta}{4}(q_{n+1} - q_n)^4, \quad (22)$$

where N is the number of active oscillators in the system. We work with fixed boundary conditions, such that $p_0 = p_{N+1} = 0$ and $q_0 = q_{N+1} = 0$. We initialize our state in a single normal mode with a fixed momentum k keeping all other modes empty. The normal modes are defined through a canonical transformation

$$q_n = \sqrt{\frac{2}{N+1}} \sum_{k=1}^{N} Q_k \sin\left(\frac{nk\pi}{N+1}\right), \quad (23)$$

and similarly for p_n and P_k. As long as $\beta = 0$, each of these modes corresponds to a stationary orbit in phase space with a conserved energy $E_0^{(k)} = P_k^2/2 + \omega_k^2 Q_k^2/2$, where $\omega_k = 2\sin(\frac{\pi k}{2(N+1)})$. Before proceeding to a general case, we will analyze a simpler two-oscillator system with $N = 2$, which is already chaotic in the presence of nonlinearity and hence does not have a closed-form AGP. We will choose a normal mode with $k = 1$ initially populated with the energy $E_0^{(k)} \equiv E_0 = 1$.

If we compute the AGP in the noninteracting point, we will get four independent terms. This number grows rapidly if we add more oscillators. Keeping all of these terms in the variational *Ansatz* is costly. For this reason, we will use the full *Ansatz* (13) truncating the expansion at second order and optimizing for the two variational coefficients γ_1 and γ_2.

Following the same process outlined in Sec. III, we optimize and simulate our ACD driving in the $N = 2$ FPUT
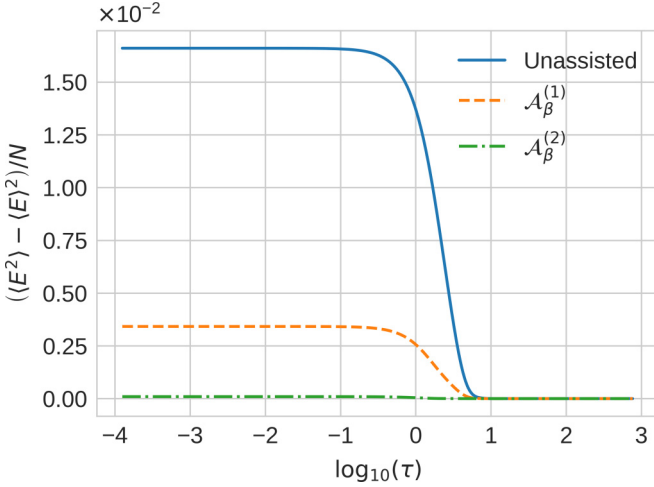
FIG. 2. Final energy variance (over system size) after turn-ons of various durations τ for the β-FPUT chain with $N = 2$ active sites. $\mathcal{A}_\beta^{(1)}$ corresponds to first-order driving using the Poisson bracket expansion and suppresses energy variance by a factor of ∼4.9 relative to no driving. $\mathcal{A}_\beta^{(2)}$ corresponds to second-order driving and suppresses energy variance by a factor of ∼176 relative to no driving.

FIG. 3. Final energy variance (over system size) after turn-ons of various durations τ for the β-FPUT chain with $N = 50$ active sites. $\mathcal{A}_\beta^{(i)}$ still correspond to ith order driving using the Poisson bracket expansion, and they have the same factors of suppression as Fig. 2.

chain. We plot the results in Fig. 2. We can see that essentially the same results are produced: first-order driving suppresses energy variance over the unassisted protocol by about a factor of 5 in the $\tau \to 0$ limit. The second-order driving gives a significantly more substantial improvement, reducing the energy variance by a factor of ∼180. This improvement is somewhat less than the one for the single-oscillator case, where the suppression factor was ∼2700, but it is still fairly impressive and almost completely eliminates nonadiabatic transitions.

We can now extend the analysis to the thermodynamic limit $N \to \infty$. This limit is well defined if we choose initial conditions with a fixed wavelength of the excited mode and the fixed amplitude of the normal mode in the initial state, which also corresponds to fixed momentum and energy densities in the system. Specifically, we choose the initial state where the initial momenta of particles are zero ($p_n = 0$) and the initial displacements of particles are given by $q_n = A \sin(nk\pi/(N + 1))$. A stationary probability distribution, where we want to initialize the system and which corresponds to a quantum stationary state, is obtained by time-averaging the initial probability distribution with the initial noninteracting Hamiltonian $H_0(\beta = 0)$. In practice, this means that before starting the ramp of the nonlinearity, one has to evolve the system with $H_0(\beta = 0)$ for a random time within one period of oscillation $T_k = 2\pi/\omega_k$ (for our simulations, we implemented this by choosing a set of times uniformly distributed between 0 and T_k and averaging over them). It is easy to see that the wavelength and the energy corresponding to these states are given by

$$\lambda = \frac{2(N + 1)}{k},$$

$$E_0 = A^2(N + 1) \sin^2\left(\frac{k\pi}{2(N + 1)}\right). \tag{24}$$

Now consider our two-site problem. Plugging in our values of $N = 2$, $E_0 = 1$, and $k = 1$, we see that our wavelength and
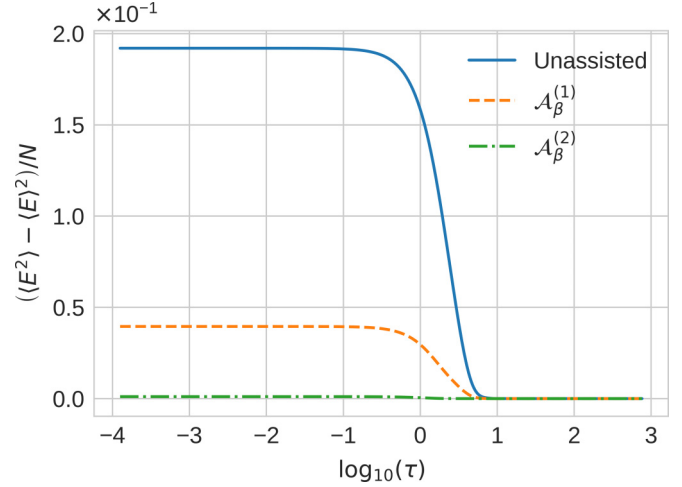
amplitude are $\lambda = 6$ and $A = 2/\sqrt{3}$. To approach our desired thermodynamic limit, we have to find a way to increase N while keeping λ and A fixed. This can be done by fixing the amplitude at $A = 2/\sqrt{3}$ and the wavelength at $\lambda = 6$ and choosing $N = k\lambda/2 - 1 \leftrightarrow k = 2(N + 1)/\lambda$, where k is an integer. It is easy to see that this choice corresponds to the system size-independent energy density in the system $E_0 = k \propto N$. It is expected that the protocol performance quickly becomes N-independent as we are dealing with a local Hamiltonian, a local *Ansatz* for the AGP, and local initial conditions. The only way the system knows about its size is through the boundary conditions, which are expected to play a negligible role at large N. For concrete calculations we choose $k = 17$, corresponding to $N = 50$. We checked that the results indeed do not change if we further increase N and k. We show the results of such simulations in Fig. 3. Aside from the different scale on the vertical axis, they look identical to Fig. 2. The relative suppression between each curve is also preserved: increasing N does not sacrifice the performance of the ACD driving. We also note that the optimized driving coefficients $\gamma_i(\beta)$ show essentially no N-dependence, testifying to the stability and applicability of the method.

V. LONG-WAVELENGTH INSTABILITY

Traditionally, the FPUT problem was analyzed for a very special initial condition corresponding excitation only in the first mode, $k = 1$. Note that as we fix the total energy in the system given by Eq. (24), the amplitude A scales as $\sqrt{N}$ as one increases the system size. Such a state is not very stable thermodynamically. This instability is precisely reflected in the very large amplitude oscillations in the mode occupancy, which emerge after quenching the nonlinearity [19,24,39]. In other words, studying the FPUT problem is equivalent to turning on β instantaneously in our setup. Note that should we fix the energy density while keeping $k = 1$, A would scale even faster, linearly with N, leading to even stronger instability. For $N = 2$ this initial condition is identical to the one we
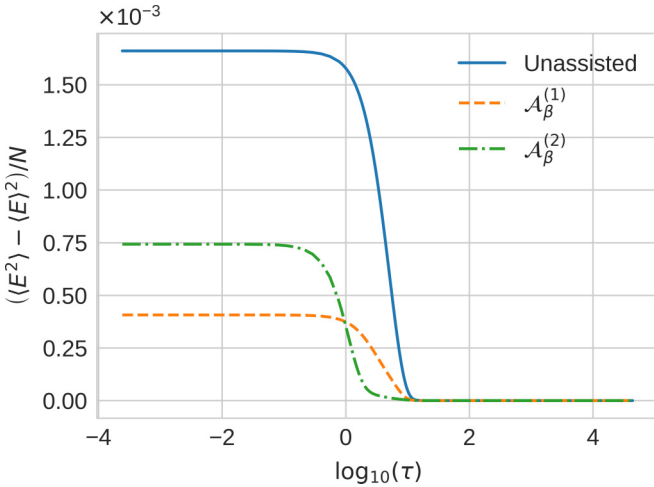
FIG. 4. Final energy variance (over system size) after turn-ons of various durations τ for the β-FPUT chain with $N = 5$ active sites. Rather than hold wavelength and amplitude fixed as N increases, we keep our initial mode number and energy fixed at 1. We clearly run into an instability at second-order driving in this limit.

FIG. 5. Final energy variance (over system size) after turn-ons of various durations τ for the β-FPUT chain with $N = 5$ active sites. We fight the instability in Fig. 4 by evaluating the action along a distribution of adiabatic trajectories over a broadened distribution of initial energies (see text for details).

used earlier, but as N increases we start to see a difference. In particular, we find that the second-order driving performs slightly worse and worse in terms of suppression of the energy variance until we reach $N = 5$, where the variational protocol essentially breaks down as illustrated in Fig. 4. Our optimization method, which previously led to dramatic suppression of excitations, has now gone awry and made our second-order *Ansatz* perform *worse* than the first-order *Ansatz* in the quench limit $\tau \to 0$. Note, however, that at intermediate speeds $\tau \geqslant 2$, the second-order ACD protocol strongly outperforms the first-order one, which indicates that some instability develops as the protocol duration gets shorter.

This failure of the second-order ACD protocol looks somewhat mysterious at first sight, so we checked and confirmed that it is not due to some artifact of the numerical integration. We also note that Ref. [8] found long-wavelength states difficult to drive, so this issue is not without precedent. We found that the origin of the failure of the variational method in this case is high sensitivity of the dynamical response of the system to the energy of the state. In particular, the ACD protocol tries to keep the system close to the adiabatically connected state. However, because the variational AGP does this only approximately, the actual state of the system under fast ACD protocol deviates from the adiabatically connected state. Normally this deviation is not important, as the variational AGP is very stable with respect to small changes in the system. But for our thermodynamically unstable state, this is not the case, and the evolution is highly sensitive to the precise form of the AGP, so any deviation of the probability distribution from the target one is amplified. We found that even in this situation, one can fix the ACD protocol by choosing the initial distribution with fluctuating energy. Specifically, we choose an initial Gaussian ensemble, where we populated the first $k = 1$ mode with the mean energy $\bar{E}_0 = 1$ and the standard deviation $d_E = 0.4$. As the energy must always be positive, we suppressed the tail of the distribution corresponding to $E_0 < 0$. The total action is then defined as the weighted
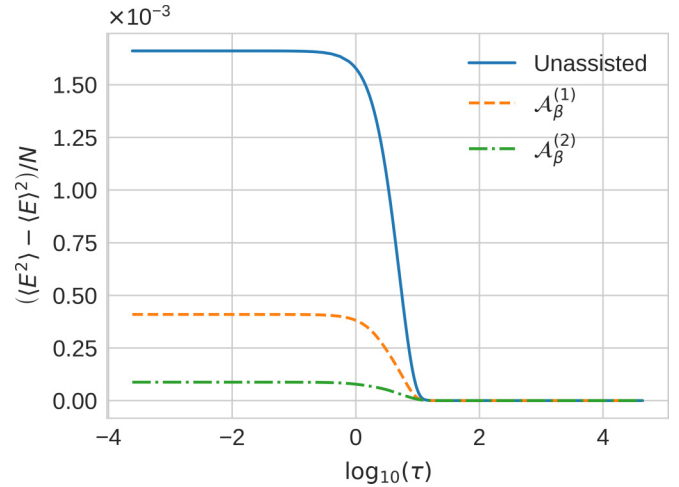
average of the actions corresponding to different values of E_0 weighted with this distribution. For calculations we used a set of 40 different realizations of E_0. As illustrated in Fig. 5, the outlined procedure stabilizes the performance of the ACD driving, although the enhancement of the second-order AGP is not as pronounced as in the previous thermodynamically stable case of initial constant λ and energy density. These results suggest that in nontrivial situations where ACD is used to drive the system through some unstable regime, one needs to carefully choose the distribution $P_\beta(\vec{q}, \vec{p})$ defining the action: too broad a distribution might lead to poor performance of the ACD protocol, while too narrow a distribution might lead to developing instabilities at short protocol times. In Appendix C we provide a more detailed analysis of the origin of this instability.

VI. DISCUSSION AND CONCLUSIONS

Counterdiabatic driving can find applications in a broad range of systems as a way to control the evolution of a state under changes to the Hamiltonian. With ACD driving, one can find protocols suppressing excitations even when the protocol durations are short, and unassisted protocols lead to large irreversible losses. In this work, we have successfully generalized ACD driving from quantum spin systems to classical nonlinear oscillator systems. In particular, we have shown that ACD driving can dramatically suppress transitions in both a single anharmonic oscillator and the β-FPUT chain for the cost of a couple of driving parameters. We showed that our results survive in the thermodynamic limit and do not require fine-tuning for generic initial states corresponding to populating a normal mode with fixed system size-independent wavelength and amplitude.

We found that the dramatic suppression of dissipation at short protocol durations requires adding counterterms to the Hamiltonian, which are cubic in momentum. Such terms are rather unusual and do not naturally appear in standard setups.

However, they can be engineered either in periodic potentials having nonquadratic dispersion or using Floquet protocols. Precise implementation of such ACD protocols might depend on the details of the system and available controls. Our work shows that ACD can be potentially extended to both quantum and classical systems of interacting particles and fields. Because the AGP can also be used to find effective degrees of freedom [40], our work shows considerable potential for finding effective nonperturbative low-energy descriptions of such systems as well as efficient computational methods of simulations of these systems.

ACKNOWLEDGMENTS

APPENDIX A: NUMERICAL OPTIMIZATION

Here we outline how we optimize $\gamma_i(\beta)$ for $\mathcal{A}_2^*$ in Eq. (15). First, we can analytically expand G_β into three terms:

$$G_\beta = \partial_\beta H_0 - \gamma_1\{x^3 p, H_0\} - \gamma_2\{x p^3, H_0\}$$
$$= G_{00} + \gamma_1 G_{10} + \gamma_2 G_{01}. \tag{A1}$$

When we square G_β to find the action, we get

$$G_\beta^2 = G_{00}^2 + \gamma_1^2 G_{10}^2 + \gamma_2^2 G_{01}^2 + 2\gamma_1 G_{00}G_{10}$$
$$+ 2\gamma_2 G_{00}G_{01} + 2\gamma_1\gamma_2 G_{10}G_{01}. \tag{A2}$$

It is also easy to see that

$$\langle G_\beta\rangle^2 = \langle G_{00}\rangle^2 + \gamma_1^2\langle G_{10}\rangle^2 + \gamma_2^2\langle G_{01}\rangle^2 + 2\gamma_1\langle G_{00}\rangle\langle G_{10}\rangle$$
$$+ 2\gamma_2\langle G_{00}\rangle\langle G_{01}\rangle + 2\gamma_1\gamma_2\langle G_{10}\rangle\langle G_{01}\rangle. \tag{A3}$$

Taking the expectation value of Eq. (A2) and subtracting Eq. (A3), we get the action in Eq. (6). We can now write this object in the form

$$\mathcal{S}(\mathcal{A}_2^*) = a_{00} + a_{10}\gamma_1 + a_{20}\gamma_1^2$$
$$+ a_{01}\gamma_2 + a_{02}\gamma_2^2 + a_{11}\gamma_1\gamma_2. \tag{A4}$$

If we now adiabatically transport the system in our simulation, we can numerically compute $\{a_{ij}\}$ at each separate value of $\beta \in [0, 1]$. For each of these values, we can independently optimize the action by requiring $\partial\mathcal{S}/\partial\gamma_i = 0$, which yields

$$\gamma_1 = \frac{2a_{02}a_{10} - a_{01}a_{11}}{a_{11}^2 - 4a_{02}a_{20}},$$
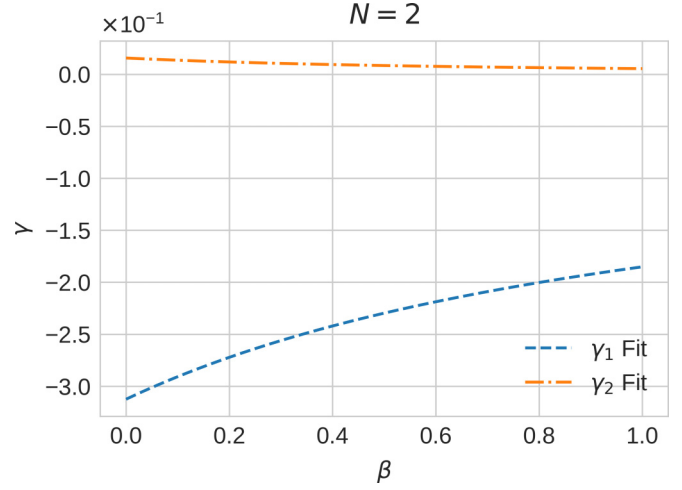$$\gamma_2 = \frac{2a_{01}a_{20} - a_{10}a_{11}}{a_{11}^2 - 4a_{02}a_{20}}. \tag{A5}$$

FIG. 6. Driving coefficients γ_i as functions of β for the $N = 2$ FPUT system discussed in the text. These are the functions spit out by the variational procedure after fitting the numerical data.

At this point, our program will spit out an array of optimal values for each γ_i, with each element corresponding to the optimal value for a certain value of β on our interval. We can now fit these arrays to functional forms, and we will have our optimal driving protocol. For the entirety of this paper, we used the fitting *Ansatz*

$$\gamma(\beta) = \frac{b_0 + b_1\beta + b_2\beta^2 + b_3\beta^3 + \cdots}{1 + c_1\beta + c_2\beta^2 + c_3\beta^3 + \cdots}. \tag{A6}$$

We restrict $c_i \geq 0$ to improve the stability of the fit and bypass possible singularities.

In Fig. 6, we plot the fitted driving coefficients $\gamma_i(\beta)$ as functions of the nonlinearity for the $N = 2$ β-FPUT system discussed in Sec. IV. Note that, although the magnitude of γ_1 is always significantly greater than γ_2, the presence of this second term allows us to achieve dramatic suppression with ACD driving. In Fig. 7, we plot the same optimized coefficients but
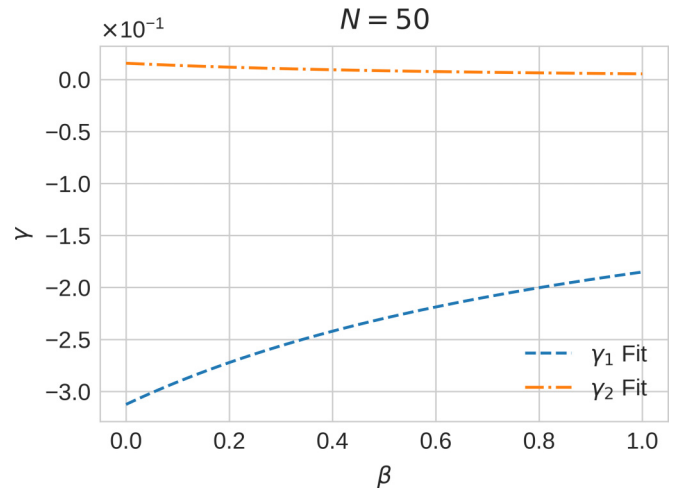

FIG. 7. Driving coefficients γ_i as functions of β for the $N = 50$ FPUT system discussed in the text. Note the essentially identical results compared to Fig. 6.

for the $N = 50$ FPUT system discussed in the text. Note that these optimized functions are essentially identical to those in the $N = 2$ case, indicating that our thermodynamic limit is stable and that the same driving parameters can be applied to systems with drastically different N.

APPENDIX B: FIGURES OF MERIT

Here we will discuss other figures of merit we considered as ways to quantify nonadiabatic effects. To begin, we will start in the quantum version of the single anharmonic oscillator. Here, we can compare our classical FOMs to the gold standard: the overlap of the final state with the target state. More specifically, we look into $1 - F^2$, where $F = |\langle \psi_{\text{target}} | \psi_f \rangle|$ is the fidelity. Once this is calculated, we can compare it to $\langle \psi_f | H_0^2 | \psi_f \rangle - \langle \psi_f | H_0 | \psi_f \rangle^2$, the quantum version of the FOM used in the body of the paper, to determine whether or not the two measures give similar information about the performance of the ACD protocols.

In addition to these two FOMs, we also considered a third candidate: temporal variance. If our transport is not adiabatic and $|\psi_f\rangle$ is not a steady state, a general observable O will have time-dependent matrix elements. As long as O is not conserved (like total energy), we can take its expectation value $\langle O \rangle(t)$ and find its infinite-time variance: $\overline{\langle O \rangle^2} - \overline{\langle O \rangle}^2$, where the overline denotes time averaging over $t \in [\tau, \infty)$. For our single anharmonic oscillator, we chose the observable to be the harmonic part of the Hamiltonian, such that our temporal variance became

$$\delta_{\text{QM}} \equiv \overline{\langle H_{\text{lin}} \rangle^2} - \overline{\langle H_{\text{lin}} \rangle}^2, \quad t \in [\tau, \infty), \quad \text{(B1)}$$

where $H_{\text{lin}} = (p^2 + x^2)/2$. With these definitions, we can now simulate them in the quantum-mechanical version of our single anharmonic oscillator. Standard substitutions between quantum and classical apply: functions become operators, averages become expectation values, and Poisson brackets become commutators up to a factor of $i\hbar$. Because in quantum systems position and momentum operators do not commute, we use symmetrized (Weyl) ordering of the quantum operators in the variational *Ansatz* (see, e.g., Ref. [41] for details):

$$\mathcal{A}_\beta^{(1)} = \gamma_1 x^3 p \rightarrow \hat{\mathcal{A}}_\beta^{(1)} = \frac{\gamma_1}{2}(\hat{x}^3 \hat{p} + \hat{p}\hat{x}^3), \quad \text{(B2)}$$

where we have added hats to emphasize the difference between functions and operators. Similarly for the second-order *Ansatz* we use

$$\hat{\mathcal{A}}_\beta^{(2)} = \frac{\gamma_1}{2}(\hat{x}^3 \hat{p} + \hat{p}\hat{x}^3) + \frac{\gamma_2}{4}(\hat{x}\hat{p}^3 + \hat{p}\hat{x}\hat{p}^2 + \hat{p}^2\hat{x}\hat{p} + \hat{p}^3\hat{x}). \quad \text{(B3)}$$

We can now optimize these *Ansätzes* and test our candidate FOMs. The results are shown in Fig. 8. First, consider unassisted driving: all three FOMs interpolate between a quench plateau and zero in the adiabatic limit. However, both fidelity and temporal variance develop peaks at intermediate turn-on times, which likely emerge from some resonant process in the system. Next, look to first-order ACD driving. With the fidelity FOM, this *Ansatz* only suppresses transitions at intermediate times relative to no driving; it has essentially no effect in the $\tau \to 0$ limit. Temporal energy fluctuations
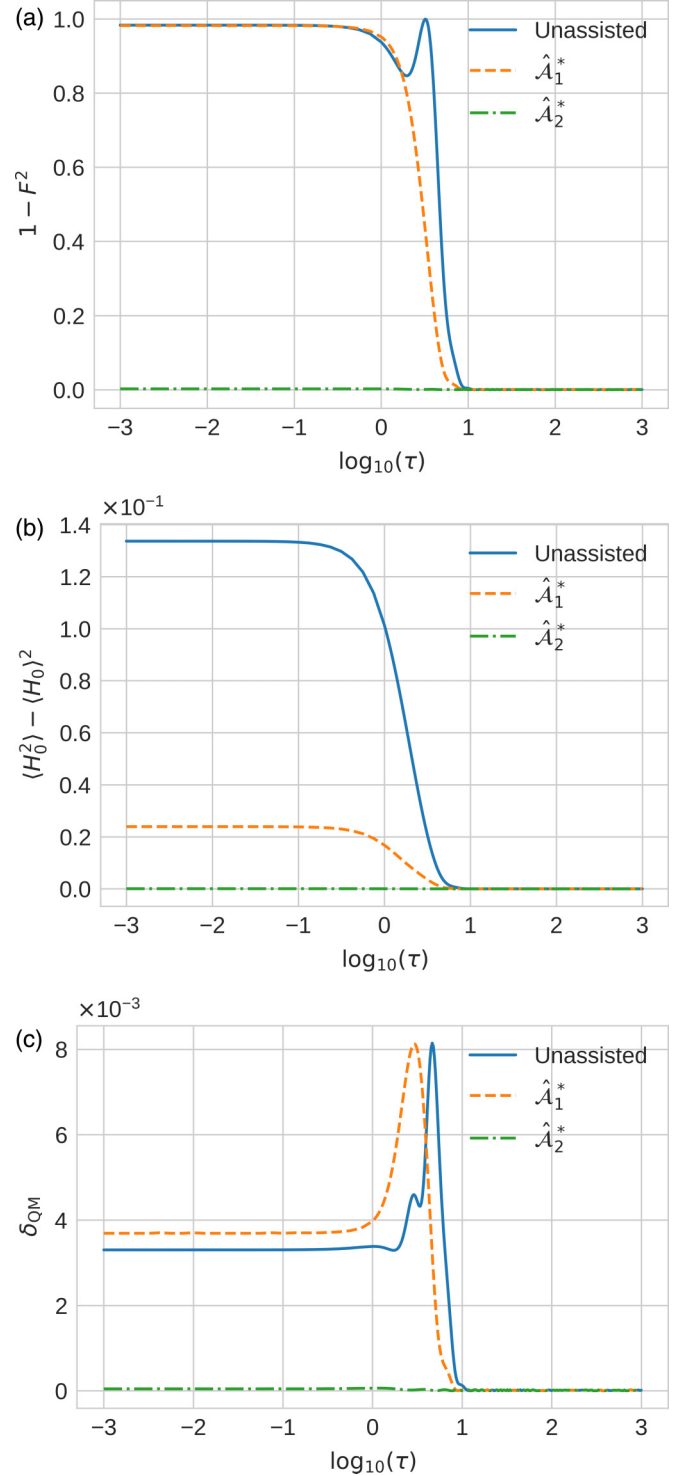

FIG. 8. Candidate FOMs in a quantum anharmonic oscillator. Note that all three FOMs agree only when the suppression is dramatic (second-order driving). The FOMs do not agree on the performance of first-order driving.

even slightly increase at short first-order ACD protocols compared to unassisted driving. Overall this analysis suggests that while first-order ACD suppresses energy fluctuations, it does not really bring the system closer to the adiabatic limit at short protocol times. The situation changes dramatically
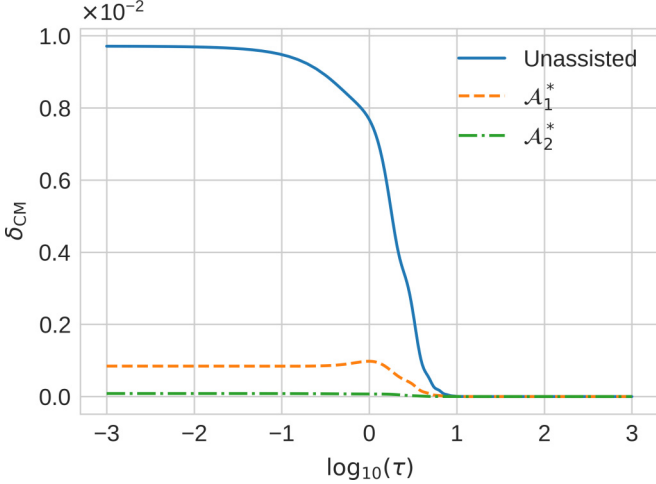
FIG. 9. Classical definition of temporal variance vs turn-on time for the single harmonic oscillator in Sec. III. Note that these curves closely match those in Fig. 1.

FIG. 10. Functional overlap of driving coefficients $\gamma_i(\beta)$ at width d_E with the same coefficients at width $d_E = 0$. The values interpolate smoothly and remain close to 1 by the time $d_E = 0.4$, indicating that the instability does not appear during optimization.

if we analyze the second-order ACD driving where now all three FOMs show dramatic improvement over the unassisted protocol. We conclude that both in quantum and in classical systems, significant suppression of diabatic transitions comes from the presence of the cubic in momentum counterterm.

We can now move to the classical anharmonic oscillator. In this limit, the fidelity FOM obviously will not survive, but we also have to modify our temporal variance FOM. In the quantum case, the discrete spectrum allowed us to take an infinite-time variance and have a nonzero, sensible answer. However, when we move to the classical limit, our spectrum becomes continuous. This fact means that for a distribution with nonzero spread in energy, the infinite-time average of temporal fluctuations will vanish. Although at short times observables will oscillate, the continuum of frequencies will cause the system to dephase and relax to a steady state. To construct a sensible FOM, we must focus our attention on this period of transient oscillations before relaxation, and as such, we need some characteristic time to average over.

Because this behavior arises from energy spread, we take our characteristic time to be

$$T_{\text{char}} \equiv \frac{E_0}{\omega \sigma_E}, \tag{B4}$$

where σ_E is the standard deviation in energy over the distribution, measured immediately after a quench turn-on. Because $E_0 = 1$ and $\omega = 1$ for our system, the characteristic time is just the inverse of this deviation. We can now redefine our temporal variance FOM as

$$\delta_{\text{CM}} \equiv \overline{\langle H_{\text{lin}} \rangle^2} - \overline{\langle H_{\text{lin}} \rangle}^2, \quad t \in [\tau, \tau + T_{\text{char}}], \tag{B5}$$

where H_{lin} is the classical version of the harmonic Hamiltonian, and the angular brackets now indicate a classical average over our probability distribution.

In Fig. 9 we plot our classical temporal variance as a function of turn-on time for the same simulation as in Fig. 1. Comparing the two figures, we see that both classical FOMs essentially agree on performance, although their turn-on curves have slightly different shapes.
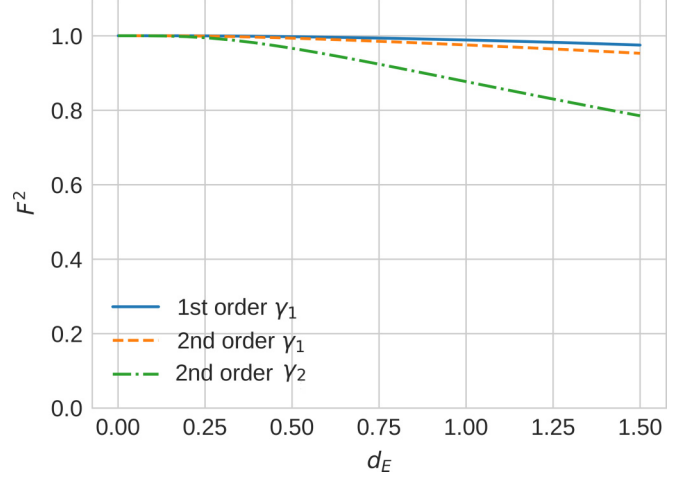
We find that for the analyzed β-FPUT problem, energy variance is a preferable FOM because it is less ambiguous than δ_{CM} and does not require choosing some rather arbitrary energy window. However, we point out that in other situations, temporal fluctuations might be preferable. For example, it is easy to check that for our initial conditions in the α-FPUT model, the energy variance remains zero even in the quench limit because the cubic nonlinear term identically vanishes. Even in the β model, temporal fluctuations might also be easier to detect experimentally, so the corresponding FOM is easier to measure.

APPENDIX C: LONG-WAVELENGTH SENSITIVITY

In this Appendix, we present evidence for why the long-wavelength instability in the text seems to arise because our system (and FOM) is very sensitive to changes in γ_i rather than the γ_i's being sensitive to changes in energy or phase space manifold. Although this analysis is by no means exhaustive, it does function as a preliminary treatment to be improved upon in future work.

1. Optimization sensitivity

Our first goal is to investigate the sensitivity of $\gamma_i(\beta)$ to changes in our chosen region of phase space. We will quantify this "region" by choosing different values of d_E corresponding to different widths of our microcanonical distribution. To quantify the sensitivity of the coefficients, we will compute the functional overlap of $\gamma_i(\beta)$ for some d_E in question with $\gamma_i(\beta)$ for $d_E = 0$. More specifically, we use

$$F_i^2(d_E) = \frac{\left(\int_0^1 \gamma_i(\beta, d_E) \gamma_i(\beta, d_E = 0) d\beta \right)^2}{\left(\int_0^1 \gamma_i^2(\beta, d_E) d\beta \right) \left(\int_0^1 \gamma_i^2(\beta, d_E = 0) d\beta \right)}. \tag{C1}$$

If the optimization process is not sensitive to small changes in d_E, we would expect this overlap to smoothly interpolate between 1 and some value $1 - \epsilon$ as our width increases from 0, indicating that the coefficients are not very sensitive to d_E.
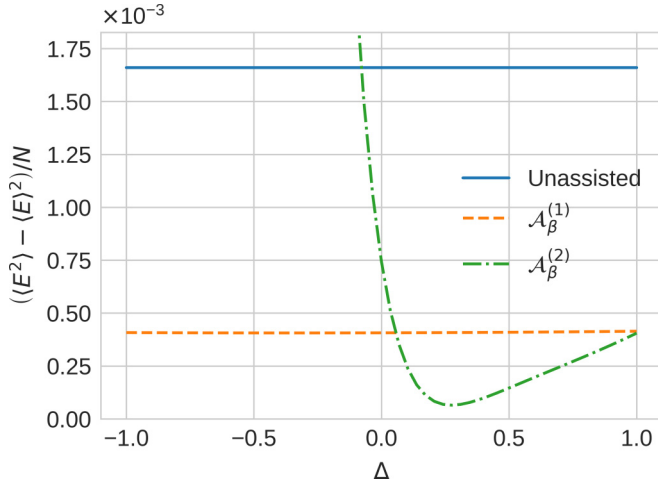
FIG. 11. Final energy variance (over system size) after quench turn-on ($\tau \approx 3 \times 10^{-4}$) as a function of perturbation Δ. We perform the simulations in the $N = 5$ β-FPUT chain with $E_0 = 1$ and $d_E = 0$.

FIG. 12. Final energy variance (over system size) after quench turn-on ($\tau \approx 3 \times 10^{-4}$) as a function of perturbation Δ. We perform the simulations in the $N = 5$ β-FPUT chain with $E_0 = 1$ and $d_E = 0.4$.

However, if they are sensitive, we would expect a sharp drop in the overlap to some value near zero.

In Fig. 10, we plot this metric for $d_E \in [0, 1.5]$ in our long-wavelength $N = 5$ setup. As is evident from the plot, F^2 interpolates smoothly as we vary the width, and by the time $d_E = 0.4$, the lowest overlap is still around 0.97. As such, we can conclude that the sensitivity of the ACD protocol is not coming from an unstable variational minimization procedure.

2. System sensitivity

Our next goal is to investigate the sensitivity of our energy variance FOM to small changes in our driving parameters $\gamma_i(\beta)$. To simplify the process of adding perturbations to these functional forms, we will use the following method: first, choose a perturbation strength Δ. Then, to each of the optimized coefficients, perturb the rational fit from Eq. (A6) such that it becomes

$$\gamma_i'(\beta) = \frac{b_0 + b_1(1 + \Delta)\beta + b_2\beta^2 + b_3\beta^3 + \cdots}{1 + c_1\beta + c_2\beta^2 + c_3\beta^3 + \cdots}. \quad \text{(C2)}$$
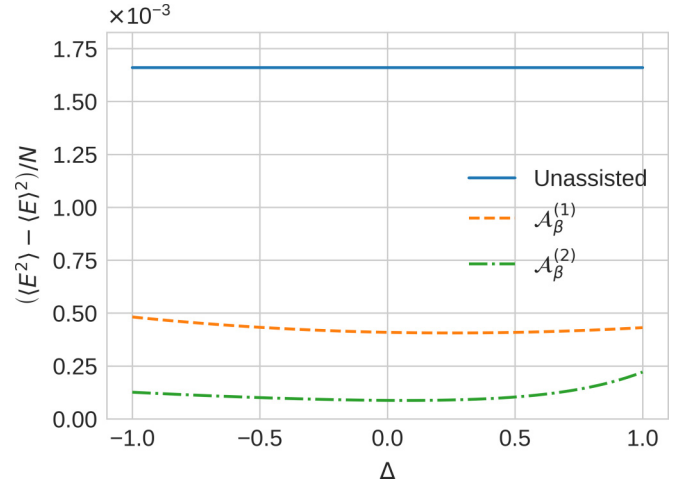
Now, simulate a quench turn-on (sufficiently small τ) and measure the energy variance FOM. Repeat for many different values of Δ. The dependence of the FOM on this perturbation will reveal how sensitive the system is to small changes in the driving.

In Fig. 11, we plot the FOM as a function of Δ for the case in which our width is $d_E = 0$. We use a quench turn-on time of $\tau \approx 3 \times 10^{-4}$ and the same $E_0 = 1$, so just as expected the $\Delta = 0$ line corresponds to the far left of Fig. 4. We see that near $\Delta = 0$, the first-order curve is essentially flat, indicating that the system is not very sensitive to perturbations in this driving. Meanwhile, the curve for second-order has a large negative slope when $\Delta = 0$, indicating a strong dependence on the driving coefficients. It is here that we observe the sensitivity necessary to explain the long-wavelength instability, not in the optimization process.

We can also perform the same analysis when optimizing for width $d_E = 0.4$, the results of which are plotted in Fig. 12. Now, both curves are essentially flat in the neighborhood of $\Delta = 0$. These attributes indicate that our method of fighting the instability actually worked: exploring a larger neighborhood of phase space suppressed the system's sensitivity to driving coefficients.

[1] X. Chen, A. Ruschhaupt, S. Schmidt, A. del Campo, D. Guéry-Odelin, and J. G. Muga, Fast Optimal Frictionless Atom Cooling in Harmonic Traps: Shortcut to Adiabaticity, Phys. Rev. Lett. **104**, 063002 (2010).

[2] D. Guéry-Odelin, A. Ruschhaupt, A. Kiely, E. Torrontegui, S. Martínez-Garaot, and J. Muga, Shortcuts to adiabaticity: Concepts, methods, and applications, Rev. Mod. Phys. **91**, 045001 (2019).

[3] E. Torrontegui, S. Ibáñez, S. Martínez-Garaot, M. Modugno, A. del Campo, D. Guéry-Odelin, A. Ruschhaupt, X. Chen, and J. G. Muga, Shortcuts to adiabaticity, in *Advances in Atomic,*

Molecular, and Optical Physics, edited by E. Arimondo, P. R. Berman, and C. C. Lin (Academic Press, 2013), Vol. 62, pp. 117–169.

[4] A. del Campo and K. Kim, Focus on shortcuts to adiabaticity, New J. Phys. **21**, 050201 (2019).

[5] M. Demirplak and S. A. Rice, Adiabatic population transfer with control fields, J. Phys. Chem. A **107**, 9937 (2003).

[6] M. Demirplak and S. A. Rice, Assisted adiabatic passage revisited, J. Phys. Chem. B **109**, 6838 (2005).

[7] M. V. Berry, Transitionless quantum driving, J. Phys. A **42**, 365303 (2009).

[8] A. del Campo, M. M. Rams, and W. H. Zurek, Assisted Finite-Rate Adiabatic Passage Across a Quantum Critical Point: Exact Solution for the Quantum Ising Model, Phys. Rev. Lett. **109**, 115703 (2012).

[9] D. Sels and A. Polkovnikov, Minimizing irreversible losses in quantum systems by local counterdiabatic driving, Proc. Natl. Acad. Sci. (USA) **114**, E3909 (2017).

[10] M. Kolodrubetz, D. Sels, P. Mehta, and A. Polkovnikov, Geometry and non-adiabatic response in quantum and classical systems, Phys. Rep. **697**, 1 (2017).

[11] A. Hartmann, V. Mukherjee, W. Niedenzu, and W. Lechner, Many-body quantum heat engines with shortcuts to adiabaticity, Phys. Rev. Research **2**, 023145 (2020).

[12] A. Hartmann, V. Mukherjee, G. B. Mbeng, W. Niedenzu, and W. Lechner, Multi-spin counter-diabatic driving in many-body quantum Otto refrigerators, Quantum **4**, 377 (2020).

[13] L. Prielinger, A. Hartmann, Y. Yamashiro, K. Nishimura, W. Lechner, and H. Nishimori, Two-parameter counter-diabatic driving in quantum annealing, Phys. Rev. Research **3**, 013227 (2021).

[14] P. W. Claeys, M. Pandey, D. Sels, and A. Polkovnikov, Floquet-Engineering Counterdiabatic Protocols in Quantum Many-Body Systems, Phys. Rev. Lett. **123**, 090602 (2019).

[15] S. Kumar, S. Sharma, and V. Tripathi, Counterdiabatic route for preparation of state with long-range topological order, Phys. Rev. B **104**, 245113 (2021).

[16] S. Masuda and K. Nakamura, Fast-forward of adiabatic dynamics in quantum mechanics, Proc. R. Soc. London A **466**, 1135 (2009).

[17] A. Patra and C. Jarzynski, Shortcuts to adiabaticity using flow fields, New J. Phys. **19**, 125009 (2017).

[18] S. Deffner, C. Jarzynski, and A. del Campo, Classical and Quantum Shortcuts to Adiabaticity for Scale-Invariant Driving, Phys. Rev. X **4**, 021013 (2014).

[19] E. Fermi, J. Pasta, S. Ulam, and M. Tsingou, Studies of nonlinear problems, I, Los Alamos Report LA-1940 (1955).

[20] G. Benettin, A. Carati, L. Galgani, and A. Giorgilli, The Fermi-Pasta-Ulam problem and the metastability perspective, *The Fermi-Pasta-Ulam Problem: A Status Report*, edited by G. Gallavotti (Springer, Berlin, Heidelberg, 2008), pp. 151–189.

[21] D. Bambusi and A. Ponno, Resonance, metastability and blow up in fpu, in *The Fermi-Pasta-Ulam Problem: A Status Report*, edited by G. Gallavotti (Springer, Berlin, 2008), pp. 191–205.

[22] C. Danieli, D. K. Campbell, and S. Flach, Intermittent many-body dynamics at equilibrium, Phys. Rev. E **95**, 060202(R) (2017).

[23] S. D. Pace and D. K. Campbell, Behavior and breakdown of higher-order Fermi-Pasta-Ulam-Tsingou recurrences, Chaos **29**, 023132 (2019).

[24] S. D. Pace, K. A. Reiss, and D. K. Campbell, The β Fermi-Pasta-Ulam-Tsingou recurrence problem, Chaos **29**, 113107 (2019).

[25] M. Pandey, P. W. Claeys, D. K. Campbell, A. Polkovnikov, and D. Sels, Adiabatic Eigenstate Deformations as a Sensitive Probe for Quantum Chaos, Phys. Rev. X **10**, 041017 (2020).

[26] H. C. Yuen and W. E. Ferguson, Relationship between benjamin-feir instability and recurrence in the nonlinear Schrödinger equation, Phys. Fluids **21**, 1275 (1978).

[27] V. P. Ruban, Numerical study of Fermi-Pasta-Ulam recurrence for water waves over finite depth, JETP Lett. **93**, 195 (2011).

[28] C.-H. Hsueh, M. Tsubota, and W.-C. Wu, Thermalization of a weakly interacting bose gas in a disordered trap, arXiv:1807.11047 [cond-mat.quant-gas].

[29] V. Balasubramanian, A. Buchel, S. R. Green, L. Lehner, and S. L. Liebling, Holographic Thermalization, Stability of Anti–de Sitter Space, and the Fermi-Pasta-Ulam Paradox, Phys. Rev. Lett. **113**, 071601 (2014).

[30] S. Coen and M. Haelterman, Domain Wall Solitons in Binary Mixtures of Bose-Einstein Condensates, Phys. Rev. Lett. **87**, 140401 (2001).

[31] M. Wu and C. E. Patton, Experimental Observation of Fermi-Pasta-Ulam Recurrence in a Nonlinear Feedback Ring System, Phys. Rev. Lett. **98**, 047202 (2007).

[32] G. Passarelli, V. Cataudella, R. Fazio, and P. Lucignano, Counterdiabatic driving in the quantum annealing of the *p*-spin model: A variational approach, Phys. Rev. Research **2**, 013283 (2020).

[33] C. Jarzynski, Geometric Phases and Anholonomy for a Class of Chaotic Classical Systems, Phys. Rev. Lett. **74**, 1732 (1995).

[34] M. Okuyama and K. Takahashi, From Classical Nonlinear Integrable Systems to Quantum Shortcuts to Adiabaticity, Phys. Rev. Lett. **117**, 070401 (2016).

[35] A. Patra and C. Jarzynski, Semiclassical fast-forward shortcuts to adiabaticity, Phys. Rev. Research **3**, 013087 (2021).

[36] S. Sugiura, P. W. Claeys, A. Dymarsky, and A. Polkovnikov, Adiabatic landscape and optimal paths in ergodic systems, Phys. Rev. Research **3**, 013102 (2021).

[37] M. Okuyama and K. Takahashi, Quantum-classical correspondence of shortcuts to adiabaticity, J. Phys. Soc. Jpn. **86**, 043002 (2017).

[38] T. Villazon, A. Polkovnikov, and A. Chandran, Swift heat transfer by fast-forward driving in open quantum systems, Phys. Rev. A **100**, 012126 (2019).

[39] J. Tuck and M. Menzel, The superperiod of the nonlinear weighted string (FPU) problem, Adv. Math. **9**, 399 (1972).

[40] J. Wurtz, P. W. Claeys, and A. Polkovnikov, Variational Schrieffer-Wolff transformations for quantum many-body dynamics, Phys. Rev. B **101**, 014302 (2020).

[41] P. W. Claeys and A. Polkovnikov, Quantum eigenstates from classical gibbs distributions, SciPost Phys. **10**, 014 (2021).