

Tail-constraining stochastic linear–quadratic control: a large deviation and statistical physics approach

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

J. Stat. Mech. (2012) P08007

(<http://iopscience.iop.org/1742-5468/2012/08/P08007>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 192.12.184.6

The article was downloaded on 13/03/2013 at 15:00

Please note that [terms and conditions apply](#).

Tail-constraining stochastic linear–quadratic control: a large deviation and statistical physics approach

Michael Chertkov¹, Igor Kolokolov^{1,2} and Vladimir Lebedev^{1,2}

¹ Center for Nonlinear Studies and Theoretical Division at LANL, and New Mexico Consortium, Los Alamos, NM, USA

² Landau Institute for Theoretical Physics, Moscow, Russia

E-mail: chertkov@lanl.gov, kolokol@itp.ac.ru and lebede@itp.ac.ru

Received 6 April 2012

Accepted 16 July 2012

Published 10 August 2012

Online at stacks.iop.org/JSTAT/2012/P08007

doi:10.1088/1742-5468/2012/08/P08007

Abstract. The standard definition of the stochastic risk-sensitive linear–quadratic (RS-LQ) control depends on the risk parameter, which is normally left to be set exogenously. We reconsider the classical approach and suggest two alternatives, resolving the spurious freedom naturally. One approach consists in seeking for the minimum of the tail of the probability distribution function (PDF) of the cost functional at some large fixed value. Another option suggests minimizing the expectation value of the cost functional under a constraint on the value of the PDF tail. Under the assumption of resulting control stability, both problems are reduced to static optimizations over a stationary control matrix. The solutions are illustrated using the examples of scalar and 1D chain (string) systems. The large deviation self-similar asymptotic of the cost functional PDF is analyzed.

Keywords: robust and stochastic optimization, large deviations in non-equilibrium systems

ArXiv ePrint: [1204.0820](https://arxiv.org/abs/1204.0820)

Contents

1. Introduction	2
2. The deterministic case and LQ optimal control	6
3. The generating function	7
4. Optimal asymptotic controls	10
4.1. The scalar case	11
5. The example of a string	12
6. Conclusions and the path forward	13
Acknowledgments	15
References	15

1. Introduction

Stochastic differential equations are used both in control [1]–[7] and in statistical physics [8]–[12] to state the problems. The two fields also use similar mathematical methods to analyze these equations. However, and in spite of the commonalities, there has been relatively little overlap between the disciplines in the past, even though the communications between the two communities have improved in recent years. Some new areas in control, for example stochastic path integral control [13]–[16], have emerged influenced by analogies, intuition and advances in statistical/theoretical physics. Vice versa, many practical experimental problems in physics, chemistry and biology dealing with relatively small systems (polymers, membranes, etc), which are driven and experience significant thermal fluctuations, can now be analyzed and manipulated/controlled with accuracy and quality unheard of in the past; see for example [17, 18]. Also, approaches from both control theory and statistical physics have started to be applied to large natural and engineered networks, like chemical, biochemical and queuing networks [19]–[22]. The dynamics over these networks is described by stochastic differential equations; the networks have enough control knobs, and they function under significant fluctuations which need to be controlled to prevent rare but potentially devastating failures. The related setting of stochastic optimization, i.e. optimization posed under uncertainty, has also come recently into the spotlight as regards statistical physics inspired algorithms and approaches [23]. The convergence of these and related ideas motivated this paper, where we discuss the analysis and control of rare events in the simplest possible, but practically rather widespread, universal and general, linear setting. We realize that the general topic of linear control is well studied and many (if not all) possible questions, e.g. related to the proper way of accounting for risk (rare events), have been discussed in the field in the past. In spite of that, we still hope that this paper may also be useful not only to

physicists, who may wish to explore new and largely unusual (in physics) formulations, but also to control theorists.

Consider the first-order (in time derivatives) stochastic linear dynamics of a vector $x = (x_i|i = 1, \dots, N)$ over a time interval $t' \in [t; T]$:

$$\frac{d}{dt'}x = Ax + Bu + \xi(t'), \quad (1.1)$$

where: A and B are constant matrices; $u(t')$ is the control vector applied at the moment of time t' ; and $\{\xi\} = (\xi(t')|t' \in [t; T])$ is the zero-mean, short-range-correlated noise with covariance V :

$$\langle \xi_i(t') \rangle = 0, \quad \langle \xi_i(t') \xi_j(t'') \rangle = \delta(t' - t'') V_{ij}, \quad i, j = 1, \dots, N \quad (1.2)$$

where one utilizes ‘statistical physics’ notation for the expectation value (average) over noise, $\langle \dots \rangle$. Here in equation (1.2) and below, the averages are over multiple possible realizations of the noise, each generating a new trajectory of the system, $\{x\} = (x(t')|t' \in [t; T])$, under a given control $\{u\} = (u(t')|t' \in [t; T])$. equation (1.1) is causal, thus assuming retarded (Stratonovich) regularization of the noise on the right-hand side of the discrete version of equation (1.1). The physical meaning of the vectors and matrices in equation (1.1) is as follows. A is the matrix explaining the stretching, shearing and rotation of the system trajectory in the N -dimensional space if control and external noise were not applied. Matrix B describes possible limitations on the degrees of freedom in the system that one can control. To simplify the notation we consider signal, control and noise vectors having the same dimension, N , where thus B is quadratic. The setting of equations (1.1) and (1.2) is the classic one in control theory. It describes the so-called linear–quadratic (LQ) stochastic control problem, which was introduced in [2, 3, 5, 6] and became foundational for control theory as a field; see e.g. [24, 25] and references therein. In the classical formulation one seeks to solve the following optimization, for $t \in [0; T]$:

$$\min_{\{u\}} \langle J(t; T; \{u\}, \{x\}) \rangle, \quad (1.3)$$

$$J(t; T; \{u\}, \{x\}) \equiv \frac{1}{2} x^*(T) F x(T) + \frac{1}{2} \int_t^T dt' (x^*(t') Q x(t') + u^*(t') R u(t')),$$

where Q, R and F are pre-defined stationary (time-independent) symmetric positive matrices and one uses an asterisk, $*$, to mark transposition. $J(t; T; \{u\}, \{x\})$ (later on, and when it is not confusing, we will use the abbreviated notation J) is a scalar quadratic cost functional of the state vector, $\{x\} = (x(t')|t' \in [t; T])$, and the control vector, $\{u\} = (u(t')|t' \in [t; T])$, evaluated for all intermediate times t' from the $[0; T]$ interval. Here in equation (1.3) (and everywhere below in the paper) the average over noise $\{\xi\}$ includes conditioning on equation (1.3), i.e. $\{x\}$ is dependent on the realization of the noise, $\{\xi\}$, and on the control, $\{u\}$, according to equation (1.1). It is assumed that the stochastic LQ control is evaluated offline, i.e. the optimal solution $u_*(t; x(t))$ of equation (1.3) is computed and saved prior to the executing actual experiment for any initial condition $x(t)$ at any t . Then in the course of the actual experiment (execution of the dynamics), $x(t)$ is measured at any time t and the respective $u_*(t; x(t))$ is applied. (When the observation of $x(t)$ is partial and noisy one needs to generalize the stochastic LQ control, for example considering the stochastic linear–quadratic–Gaussian (LQG) control; see e.g. [25] for details.) We also assume (and the details will be clarified below) that the

optimal control succeeds, i.e. the system stabilizes and J does not grow with T faster than linearly.

An unfortunate caveat as regards the LQ setting (1.3) is the lack of fluctuation control: even though the LQ solution is optimal in terms of minimizing the expectation value of the cost functional, it may generate very significant fluctuations when it comes to analysis of the $\mathcal{J} \gg \langle J \rangle$ tail of the probability distribution function, $\mathcal{P}(\mathcal{J})$, of the cumulative cost $\mathcal{J}(T; x(0)) \equiv J(0; T; \{u_*\}, \{x\})$. A stochastic risk-sensitive LQ (RS-LQ) scheme [26]–[28] was introduced to improve control of the abnormal fluctuations of J . RS-LQ constitutes the following generalization of the LQ scheme (1.3):

$$\max_{\{u\}} \langle \exp(-\theta J) \rangle, \quad (1.4)$$

where θ is a pre-defined parameter. Intuitively one relates the case of positive θ to a risk-averse optimum. It is assumed within the standard RS-LQ scheme that θ is fine-tuned by means of some additional considerations. Note that, as shown in [29], the RS-LQ control is also ultimately related to the so-called H_∞ -norm robust control. (See also [30] for further discussion of the relation.)

In this paper we analyze two natural modifications of the stochastic RS-LQ control. The two schemes can both be interpreted in terms of the RS-LQ approach supplemented by an additional optimization over θ . Our first, *tail-optimum* (TO) scheme consists in the following modification of the LQ (1.3) and RS-LQ (1.4) ones:

$$\min_{\{u\}} \mathcal{P}(J = j \cdot (T - t) | \{u\}). \quad (1.5)$$

In words, the TO-LQ control minimizes (at any time t and given the current observation $x(t)$) the probability of the current value of the cost functional $J(t; T; \{u\}, \{x\})$ evaluated at a pre-defined value, $j \cdot (T - t)$, where thus j is the only external parameter left in the formulation. Another strategy, which we call *chance-constrained* LQ (CC-LQ), in reference to similar formulations in optimization theory [31]–[33], consists in minimizing the mean of the cost functional under the condition that the tail probability evaluated at $j \cdot (T - t)$ does not exceed the prescribed threshold value $\varepsilon(t; T)$:

$$\min_{\{u\}} \langle J \rangle \quad \text{s.t.} \quad \mathcal{P}(J = j \cdot (T - t) | \{u\}) \leq \varepsilon(t; T). \quad (1.6)$$

The main objectives, and consequently results of this study, are:

- To extend the asymptotic, $T \rightarrow \infty$, approach developed in the past for LQ (1.3) and RS-LQ (1.4) optimal controls to the new TO-LQ (1.5) and CC-LQ (1.6) optimal settings. At $T \rightarrow \infty$ the optimal control takes the following universal form, linear in x :

$$u_*(t; x) = -Kx, \quad (1.7)$$

where K is a t -independent but model-dependent matrix. The condition of the system stability, intuitively translating into the expectation that \mathcal{J} grows not faster than linearly with T , naturally requires that all the eigenvalues of the stability matrix, $\mu = BK - A$, have positive real part. The linearity of the optimal control (1.7) in x is a direct consequence of the linearity of the initial dynamical system. The time independence and initial condition independence of the optimal control (1.7) are asymptotic: they are achieved at $T \gg \tau_*$, where τ_* can be estimated as the inverse of

the absolute value of the eigenvalue of $BK - A$ with the smallest real part. Systems of algebraic equations defining K implicitly for the TO-LQ and CC-LQ cases are presented and then juxtaposed against the previously analyzed cases of the LS and RS-LQ controls. (See equations (4.1), (4.3) and (4.5).) Finding optimal control is reduced to optimization over time-independent K . The resulting dependences are homogeneous in time, with t and T always entering in the $T - t$ combination. (This also simplifies the analysis, allowing us to set $t = 0$.)

- To analyze the statistics of the optimal cost functional, \mathcal{J} , in the large deviation (LD) regime, i.e. at large but finite T . We show that in the stable regime the PDF of \mathcal{J} attains the following universal LD form:

$$\log \mathcal{P}(\mathcal{J}) \sim -T\mathcal{S}(\mathcal{J}/T), \quad (1.8)$$

where the LD function, $\mathcal{S}(j)$, is a convex function of its argument found implicitly (in a closed algebraic form, which may or may not yield an efficient algorithm) for the four cases (of LS, RS-LQ, TO-LQ and CC-LQ controls) considered. The LD function shows a universal, $\mathcal{S}(j) \rightarrow aj$, tail at large (i.e. larger than typical) j , where the value of positive a depends on the model. This suggests, in particular, that it is natural to choose in the CC optimization (1.6) $\varepsilon(t; T) = \exp(-c(T - t))$, for the threshold, with c being a constant. To derive compact algebraic expressions for the LD function we, first, analyze the generating function of \mathcal{J} evaluated at linear u parameterized by K as in equation (1.7),

$$\mathcal{Z}(\theta; K) \equiv \langle \exp(-\theta\mathcal{J}) \rangle_*, \quad (1.9)$$

and then express the optimization/control objective as a convolution of the integral or differential operator/kernel in θ (the choice will depend on the model) and $\mathcal{Z}(\theta; K)$, and finally evaluate the optimization over K in the asymptotic LD approximation. Here in equation (1.9), the subscript asterisk sign $*$ in the expectation/average (over noise and constrained to equation (1.1) indicates that the control vector is taken in the form of the linear ansatz $u \rightarrow -Kx$, where K is left as yet undetermined.

The remainder of the paper is organized as follows. We start by discussing the deterministic case (of zero noise) in section 2. This regime is of interest for two reasons. First, in the asymptotics for zero noise the four, generally different, control schemes become equivalent. Besides, and as is well known from the classical papers [2, 24, 25], optimal control in the bare LQ case (corresponding to minimization of the cost function average) is not sensitive (and thus independent of) the level of the noise. Section 3 is devoted to analysis of the generating function (1.9), the average value of the cost function and the tail of the cost function distribution restricted to an as yet unspecified value of K . Optimization over K , resulting in the known RS-LQ optimal relations and also derivation of the new optimal relations for K in the TO-LQ and CC-LQ cases, is discussed in section 4. We describe and compare asymptotic large deviation forms of the cost function PDF, $\mathcal{P}(\mathcal{J})$, in the optimal regimes. In this part and section 2 we also discuss many times the illustrative ‘scalar’ example where x and u are scalars. An infinite system example, of a ‘string’ formed from a linear 1D chain, is discussed in section 5. We conclude and discuss related future challenges in section 6.

2. The deterministic case and LQ optimal control

We start this section with a disclaimer: all results reported here are classical, described in [2], [24]–[29] and later papers and books; see e.g. [34]–[36]. We present them here only to make the whole story of the paper self-explanatory and coherent.

When the noise is ignored, equation (1.1) should be considered as a deterministic constraint, reducing any of the optimal control schemes (1.3), (1.4), (1.5) and (1.6) to a simple variation of the cost functional (1.3) over u . Using the standard variational technique with a time-dependent Lagrangian multiplier for the constraint, and then excluding the multiplier, one derives the equation

$$\frac{d}{dt'} u^* + u^* R B^{-1} A B R^{-1} = x^* Q B R^{-1}, \quad (2.1)$$

which should be supplied by the boundary condition (also following from the variation), $u^*(T) + x^*(T) F B R^{-1} = 0$. (Let us recall that we choose the notation where the dimensionality of u coincides with the dimensionality of x . We also assume that the inverses of all the matrices involved in the formulation are well defined. This assumption is not critical and is made here only to simplify the notation. In the general case when some of the matrices, and in particular R , are not full rank, one can generalize the formulas properly, using a proper notion of a pseudo-inverse.) Substituting $u = -R^{-1} B^* \Pi x$ in equation (2.1) one arrives at the following equation for Π :

$$\frac{d}{dt'} \Pi + \Pi A + A^* \Pi - \Pi B R^{-1} B^* \Pi + Q = 0. \quad (2.2)$$

with the boundary condition $\Pi(T) = F$. equation (2.2), solved backwards in time, results in $\Pi(t)$ and, then, $u_*(t; x) = -R^{-1} B^* \Pi(t) x = -K x$.

To gain a qualitative understanding of the backwards in time dynamics of Π , let us briefly discuss the simplest possible case with all the matrices entering equation (2.2) replaced by scalars, then yielding the following analytic solution for the optimal K :

$$K = \frac{A - \sqrt{A^2 + Q B^2 / R} \left(\tanh \left(\sqrt{A^2 + Q B^2 / R} (t - T_0) \right) \right)^{\pm 1}}{B}, \quad (2.3)$$

where T_0 and ± 1 are chosen to satisfy the boundary condition, $K(T) = B F / R$. When $T \gg \tau = 1 / \sqrt{A^2 + Q B^2 / R}$, the backwards in time dynamics saturates (after a short $\sim \tau$ transient) to an F -independent constant, resulting from replacing \tanh in equation (2.3) by -1 . Therefore, in the stationary regime, $T \rightarrow \infty$, the optimal control is with the constant in time, frozen K . One also finds that the optimal control in the one-dimensional deterministic case is always stable, $\mu = K B - A > 0$.

Returning to the general (finite vector) case one concludes that when T is sufficiently large the optimal control is of the form described by equation (1.7), i.e. it is linear in x and asymptotically time independent, with $K = R^{-1} B^* \Pi_0$ where Π_0 solves equation (2.2) with the first term replaced by zero. It is well known in control theory that (under some standard common-sense assumptions on B and R matrices) the stable solution of the system of the algebraic Riccati equations is unique and moreover it can be found efficiently. (See e.g. chapter 12 of [37] and references therein.)

Let us now discuss the bare LQ control, now in the presence of noise. Since equation (1.1) is linear, one can naturally split the full solution into a sum, $x = x_1 + x_2$,

where x_1 satisfies equation (1.1) without noise and it is equivalent to the noiseless solution, just discussed in this section. Then, the second term satisfies $dx_2/dt' = Ax_2 + \xi$, with $x_2(t) = 0$. However, since the noise is zero mean, $\langle \xi \rangle = 0$, x_2 is zero mean too, i.e. $\langle x_2 \rangle = 0$. Next, let us analyze the splitting of the term in $\langle J \rangle$, which is the optimization objective of the LQ scheme. Since x_1 and x_2 are independent (by construction) and because x_2 is zero mean, $\langle J \rangle$ splits into two terms, $\langle J_1 \rangle + \langle J_2 \rangle$, each dependent on vectors x_1 and x_2 only. $\langle J_1 \rangle$ is simply equivalent to J analyzed above in the deterministic case, while $\langle J_2 \rangle$ is independent of u , thus not contributing to the optimization at all. To summarize, the LQ optimal control is not sensitive to the noise and it is thus equivalent to the deterministic (noiseless) case described above.

3. The generating function

Consider the generating function (GF), $\mathcal{Z}(\theta; K)$, defined by equation (1.9). $\mathcal{Z}(\theta; K)$ is of obvious relevance to the RS-LQ scheme, but it is also useful for the analysis of other schemes as well, because of the following (Laplace transform) relation to the PDF of \mathcal{J} :

$$\mathcal{Z}(\theta; K) = \int_0^\infty d\mathcal{J} \exp(-\theta\mathcal{J})\mathcal{P}_*(\mathcal{J}), \quad (3.1)$$

where (as before) the subscript asterisk indicates that the PDF was evaluated at $u = Kx$, with K being an as yet undefined constant matrix. The inverse of equation (3.1) is

$$\mathcal{P}_*(\mathcal{J}) = \int_{c-i\infty}^{c+i\infty} \frac{d\theta}{2\pi i} \exp(\theta\mathcal{J})\mathcal{Z}(\theta; K), \quad (3.2)$$

where it is assumed that the integration contour, considered in the complex plane of θ , goes on the right from all the singularities (poles and cuts) of $\mathcal{Z}(\theta; K)$. In the path integral representation, GF takes the following form:

$$\mathcal{Z}(\theta; K) \sim \int \mathcal{D}x \mathcal{D}p \exp\left(\int_0^T dt \left(-\frac{\theta}{2}x^*\tilde{Q}x + p^*(\partial_t x + \mu x) + \frac{1}{2}p^*Vp\right)\right), \quad (3.3)$$

$$\tilde{Q} = Q + K^*RK, \quad \mu = BK - A, \quad (3.4)$$

where p is an auxiliary vector variable (momentum). Here and everywhere below we assume that, even if the dynamics was not stable before application of the control, control stabilizes it. Formally, this means that μ , defined by equation (3.4), has no eigenvalues with negative real values. The ‘boundary’ (F -dependent term) in equation (3.3) was ignored, assuming that (like in the one-dimensional LQ case discussed above) it may only influence how the optimum is approached (backwards in time) and remains inessential for describing asymptotic behavior of the optimal control. This path integral is (most conveniently) evaluated by changing to the Fourier (frequency) domain, expressing the pair correlation function as the frequency integral, and then relating it to the derivative of log GF over θ ,

$$\langle x_i x_j \rangle = \int_{-\infty}^{+\infty} \frac{d\omega}{2\pi} (\omega^2 V^{-1} + \mu^* V^{-1} \mu + \theta \tilde{Q})_{ij}^{-1}, \quad (3.5)$$

$$\frac{\partial \log \mathcal{Z}(\theta; K)}{\partial \theta} = -\frac{T}{2} \langle x^* \tilde{Q} x \rangle. \quad (3.6)$$

Here in equations (3.5) and (3.6) the averaging is over the path integral measure described by equation (3.3). Further, evaluating the integral over θ , fixing the normalization, $\mathcal{Z}(0; K) = 1$, and using the standard formula of matrix calculus, $d/d\theta \log \det(\theta \cdot 1 + D) = \text{tr}((\theta \cdot 1 + D)^{-1})$, where 1 stands for the unit matrix, one arrives at the following expression:

$$\log(\mathcal{Z}(\theta; K)) = -\frac{T}{2} \int_{-\infty}^{+\infty} \frac{d\omega}{2\pi} \log \frac{\det(\omega^2 V^{-1} + \mu^* V^{-1} \mu + \theta \tilde{Q})}{\det(\omega^2 V^{-1} + \mu^* V^{-1} \mu)}, \quad (3.7)$$

which is asymptotically exact at $T \rightarrow \infty$. Moreover, one can show that for any (spatially) finite system, the next order corrections to the rhs of equation (3.7) are $O(1)$. Note that this representation (3.7) of log GF, as an integral over frequency of the log det, is similar to the relation discussed in section 3 of [29] in the context of linking the RS-LQG control to the maximum entropy formulation of the H_∞ control. The log det has also appeared in [38], where statistics of currents were analyzed for a general non-equilibrium (off-detailed-balance) linear system.

To gain intuition, let us first analyze equation (3.7) in the simple scalar case where the integral on the rhs can be evaluated analytically:

$$\log(\mathcal{Z}(\theta; K)) = \frac{T}{2} \left(\mu - \sqrt{\mu^2 + \theta \tilde{Q} V} \right). \quad (3.8)$$

Substituting this expression into equation (3.2) and estimating the integral over θ in a saddle-point approximation (justified when T is large) one arrives at the LD expression (1.8) where

$$\mathcal{S}_*(j) = \frac{V(Q + RK^2)}{16j} + \frac{(BK - A)^2 j}{V(Q + RK^2)} - \frac{BK - A}{2}. \quad (3.9)$$

The LD function is obviously convex and it is defined only for positive j . (The asterisk marks, as before, that the average and the probability are computed conditioned on an as yet unspecified K .) $\mathcal{S}_*(j)$ achieves its minimum at $\langle j \rangle_* = -T^{-1} \partial \log \mathcal{Z} / \partial \theta |_{\theta=0} = V\tilde{Q}/(4\mu)$, and shows a linear asymptotic, $\mathcal{S}_*(j) \approx j\mu^2/(V\tilde{Q})$, at $j \gg \langle j \rangle$. Note, that the aforementioned asymptotic is associated with the cut singularity in the complex θ plane of the GF expression (3.8). Indeed, substituting equation (3.8) into (3.2) and shifting the integration contour to the left, thus forcing it to surround counter-clockwise the $]-\infty; \theta_* = -\mu^2/(V\tilde{Q})]$ cut, and then estimating the integral by a small part of the contour surrounding the vicinity of the cut tip at θ_* , we arrive at the aforementioned $j \gg \langle j \rangle$ asymptotic, $\mathcal{S}(j) \approx -j\theta_*$.

Returning to the analysis of the general formulas (3.7) and (3.2), one observes that even though reconstructing $\mathcal{S}_*(j)$ in its full integrity explicitly as a function of K does not look feasible, we can still, motivated by the scalar case analysis, make some useful general statements about both the average, $\langle j \rangle_*$, and the $j \gg \langle j \rangle_*$ asymptotic of $\mathcal{S}_*(j)$. We will start from the latter problem.

For analysis of the tail, the key object of interest is the det in equation (3.7) considered at zero frequency, $\omega = 0$. Specifically, one aims to find the zero of the determinant with the largest real value:

$$\theta_* = \max_{\theta} \text{Re}(\theta)_{\det(\mu^* V^{-1} \mu + \theta \tilde{Q})=0}. \quad (3.10)$$

Indeed, any zero (there might be many of these in the general matrix case) marks the tip of the respective cut singularity of $\mathcal{Z}(\theta; K)$ in the complex θ plane. Then, the tail, $j \gg \langle j \rangle_*$, asymptotic of the LD function becomes $\mathcal{S}_*(j) = -j\theta_*$. Note that this linear in j estimation is valid only in the case of a finite system, when the set (spectrum) of zeros (defined by the condition in equation (3.10)) is discrete. In the case of an infinite system, when the spectrum of zeros becomes quasi-continuous, one needs to account for the multiple zeros, as illustrated in the ‘string’ example of section 5.

To evaluate $\langle j \rangle_*$ (as a function of K) in the general case, one first analyzes it in the time representation. Substituting the $u = -Kx$ ansatz with constant K in equation (1.1), expressing $x(t)$ formally as an integral over time (for a given realization of the noise), substituting the result into equation (1.3), averaging over noise, and then taking the $T \rightarrow \infty$ limit, one arrives at

$$\langle j \rangle_* = \frac{1}{2} \int_0^\infty \text{tr}(V e^{-\mu^* t} \tilde{Q} e^{-\mu t}) dt = \frac{1}{2} \text{tr}(V \Pi) |_{\mu^* \Pi + \Pi \mu = \tilde{Q}}, \quad (3.11)$$

where the latter expression is an implicit (as the condition is a matrix one, and thus not resolvable explicitly in general) function of K . It is straightforward but tedious to check (introducing a matrix Lagrangian multiplier for the condition in equation (3.11) and carrying out a variation over K and Π) that optimization of equation (3.11) over K results in the algebraic Riccati equation equivalent to equation (2.2) with the first term ignored. Note that the fact that the optimal control derived from the optimization of the average cost function in the stochastic case coincides with the result of the deterministic optimization (ignoring stochasticity) is the fact very well known in control theory³. The optimal value of the functional in the deterministic case saturates to a constant at $T \rightarrow \infty$, while in the stochastic case the average optimal cost grows with time linearly. Asymptotic convergence of the two seemingly different schemes to the same optimal control is thus an indication of the asymptotic self-consistency of the linear ansatz (1.7).

Differentiating equation (3.7) over θ and then setting θ to zero, one derives an alternative (to equation (3.11)) representation for the average rate of the cost function conditioned on K :

$$\langle j \rangle_* = \int_{-\infty}^{+\infty} \frac{d\omega}{4\pi} \text{tr}((\omega^2 V^{-1} + \mu^* V^{-1} \mu)^{-1} \tilde{Q}). \quad (3.12)$$

Note that comparison of equations (3.7), (3.11) and (3.12) also allows us to derive an expression for the derivative of log GF as a time integral, and then present it in an implicit algebraic form:

$$\begin{aligned} -T^{-1} \partial_\theta \log \mathcal{Z}(\theta; K) &= \int_{-\infty}^{+\infty} \frac{d\omega}{4\pi} \text{tr}((\omega^2 V^{-1} + \mu^* V^{-1} \mu + \theta \tilde{Q})^{-1} \tilde{Q}) \\ &= \frac{1}{2} \text{tr}(\tilde{V} \Pi) |_{\mu^* \Pi + \Pi \mu = \tilde{Q}} \end{aligned} \quad (3.13)$$

$$= \frac{1}{2} \int_0^\infty \text{tr}(\tilde{V} e^{-\mu^* t} \tilde{Q} e^{-\mu t}) dt, \quad (3.14)$$

where $\tilde{V} = V(1 + \theta V(\mu^*)^{-1} \tilde{Q} \mu^{-1})^{-1}$.

³ This fact is akin to a similar statistical physics statement naturally emerging in the analysis of the linear stochastic systems driven by white Gaussian noise and characterized by flux/current (see e.g. [38]): the average flux corresponds to the minimum of a functional quadratic in the flux.

4. Optimal asymptotic controls

In this section we formulate the RS-LQ, TO-LQ and CC-LQ asymptotic schemes in the general vector/matrix form as an optimization over K . (Note that the asymptotic LQ scheme was already stated as a minimization of equation (3.11), or equivalently of equation (3.12) in section 3.) Then we illustrate these formulations on the scalar example.

K_{RS} , which is asymptotically optimal for the RS-LQ control considered at $\theta > 0$, is found by maximizing $\mathcal{Z}(\theta; K)$. Using equation (3.7) one derives

$$\min_K \int_{-\infty}^{+\infty} d\omega \log \frac{\det(\omega^2 V^{-1} + (BK - A)^* V^{-1} (BK - A) + \theta(Q + K^* RK))}{\det(\omega^2 V^{-1} + (BK - A)^* V^{-1} (BK - A))} \Big|_{\text{Re}(\lambda(BK - A)) > 0}, \quad (4.1)$$

where $\text{Re}(\lambda(BK - A)) > 0$ denotes the stability condition ensuring that the real values of all the eigenvalues of $BK - A$ are positive. Note that constancy of the stationary RS-LQ optimal control was proven in [26], therefore making our approach self-consistent. An alternative, but obviously equivalent, formulation of the RS-LQ optimal control consists in minimizing $-T^{-1} \partial_\theta \log \mathcal{Z}(\theta; K)$. Going along this path and utilizing equation (3.13) one arrives at

$$\min_{K, \Pi} \frac{1}{2} \text{tr}(V(1 + \theta V(\mu^*)^{-1} \tilde{Q} \mu^{-1})^{-1} \Pi)_{\mu^* \Pi + \Pi \mu = Q + K^* RK}, \quad (4.2)$$

generalizing the LQ formulation stated in section 3 as the minimization of equation (3.11). Solving equation (4.2) is reduced to analyzing the respective generalization of the Riccati equations, which can then be turned into a linear eigenvalue problem described within the so-called Hamiltonian approach to the RS-LQ problem discussed in [28].

From equation (1.5), and assuming time independence of the control, one can state the general asymptotic TO-LQ optimum, utilizing equations (3.2) and (3.7) as an optimization of a double integral over frequency and θ . However, in practice one is interested in discussing the TO-LQ optimization only at sufficiently large values of the cost, jT . Using the analysis of section 3 one derives the desired double-asymptotic (valid at large T and large j) and simpler to state expression describing K_{TO} :

$$\min_K \max_\theta \text{Re}(\theta) \quad \begin{aligned} &\text{Re}(\lambda(BK - A)) > 0 \\ &\det(\mu^* V^{-1} \mu + \theta \tilde{Q}) = 0, \end{aligned} \quad (4.3)$$

where max is over complex θ and the optimal LD value of the PDF tail is exponential,

$$\log \mathcal{P}_{\text{TO}}(\mathcal{J}) \approx -\text{Re}(\theta_{\text{TO}}) \mathcal{J}, \quad (4.4)$$

with θ_{TO} solving equation (3.10). Note that the $\det = 0$ condition in equation (4.3) is reminiscent of the μ -measure which is the key element of the robust control approach; see [35, 37] and references therein.

In the same double-asymptotic (large T and large j) regime the optimal CC-LQ control (1.6) is given by

$$\min_K \operatorname{tr}(V\Pi) \left| \begin{array}{l} \operatorname{Re}(\lambda(BK - A)) > 0 \\ \mu^* \Pi + \Pi \mu = \tilde{Q}, \\ \max_{\theta} (\operatorname{Re}(\theta))_{\det(\mu^* V^{-1} \mu + \theta \tilde{Q})=0} \geq \frac{\log(1/\varepsilon)}{jT}. \end{array} \right. \quad (4.5)$$

Note that unlike equations (4.1) and (4.3), equation (4.5) does not have valid solutions for any value of the $\log(1/\varepsilon)/\mathcal{J}$ ratio. In fact, it is clear from equation (4.4) that to have a nonempty solution of equation (4.5) one needs to require that $\operatorname{Re}(\theta_{\text{TO}})jT \leq \log(1/\varepsilon)$. Once the optimum solution is found, one estimates the LD asymptotic of the cost function PDF by an expression similar to the one given by equation (4.4), with the TO subscript replaced by the CC one.

4.1. The scalar case

In the remainder of this section we illustrate all of the aforementioned formulas on the scalar example. In this simple case, the integral on the rhs of equation ((4.1)) is equal to

$$2\pi \left(\sqrt{(BK - A)^2 + V\theta(Q + RK^2)} - (BK - A) \right), \quad (4.6)$$

resulting in the following optimal value:

$$K_{\theta} = \frac{A/B + \sqrt{A^2/B^2 + Q/R + QV\theta/B^2}}{1 + VR\theta/B^2}. \quad (4.7)$$

The large deviation tail of the PDF of j at a given K can be extracted from equation (3.9):

$$j \gg \langle j \rangle : \quad \mathcal{S}_*(j) \approx \frac{(BK - A)^2}{V(Q + RK^2)} j + o(j). \quad (4.8)$$

Optimizing the PDF over K , we find two different cases depending on the sign of A . At $A > 0$ the coefficient in front of the linear in j term on the rhs of equation (4.8) grows monotonically with K from the $(A/B, +\infty)$ interval. To find the optimal value of K in this case one has to consider the $o(j)$ term, thus deriving

$$A > 0 : \quad K_{\text{TO}} = \sqrt{4AjRV}, \quad \log \mathcal{P}_{\text{TO}} \approx -\frac{B^2 jT}{RV}. \quad (4.9)$$

In the other case of $A = -|A| < 0$, the linear coefficient in equation (4.8) reaches its maximum at $K = BQ/(R|A|)$, thus resulting in

$$A < 0 : \quad \log \mathcal{P}_{\text{TO}} \approx -\frac{jT}{RVQ} (RA^2 + B^2Q). \quad (4.10)$$

Finally, the CC optimal formula (4.5) has no solution if $B^2j/(RV) < c$ in the $A > 0$ case and if $j(RA^2 + B^2Q)/(RVQ) < c$ in the $A < 0$ case. (Here we assume, as above, that $\epsilon(0; T) = \exp(-cT)$.) When ϵ is chosen sufficiently small (i.e. c is sufficiently large),

the feasibility domain in equation (4.5) is not empty and one distinguishes two regimes depending on how K_ε , defined by

$$K_\varepsilon = \frac{1}{1-\kappa} \left(\frac{A}{B} + \sqrt{\kappa \left(\frac{A^2}{B^2} + (1-\kappa) \frac{Q}{R} \right)} \right), \quad \kappa = \frac{cVR}{B^2j}, \quad (4.11)$$

compares with K_0 , which is the bare LQ optimal value corresponding to K_θ from equation (4.7) evaluated at $\theta = 0$. One derives

$$K_{CC} = \max(K_\varepsilon, K_0), \quad (4.12)$$

where of the two regimes, one is achieved within the interior of the optimization domain (the tail constraint is not restrictive) while the other one corresponds to the tail imposed by the boundary of the domain. It is worth noting that (4.11) is valid for both signs of A .

5. The example of a string

In this section we discuss an explicitly solvable example of an infinite system where the set of zeros (of the determinant in the condition of equation (3.10)) forms a quasi-continuous spectrum. Consider a string, defined as an overdamped system of multiple beads on a line connected to each other by elastic springs of strength D , stretched by the linear force of the strength A and subject to Langevin driving:

$$\partial_t x_j = Ax_j + D(x_{j+1} + x_{j-1} - 2x_j) + Bu_j + \xi_j, \quad (5.1)$$

$$J = \frac{1}{2} \int_0^T dt \sum_j (Qx_j^2 + Ru_j^2), \quad (5.2)$$

where $j = 1, \dots, N$, x_j marks the position of the j th bead on the string, and the zero-mean white Gaussian noise is distributed as in equation (1.2) with $V_{ij} = V\delta_{ij}$. u_j in equation (5.1) stands for the control. We are looking for a time-independent linear in x control, assuming that the control acts uniformly on all beads on the string, i.e. $u_j = -Kx_j$. Let us also assume that the string is periodic with the period N . Then, the solution of equation (5.1) allows expansion in a series over spatial harmonics:

$$x_j = \sum_{q=1}^N \exp(iq(j/N)) x_q, \quad (5.3)$$

with the wavevector, q , from the interval $-\pi < q < \pi$, and resulting in the following separated equations for the individual harmonics:

$$\partial_t x_q = Ax_q - 2D(1 - \cos q)x_q - BKx_q + \xi_q. \quad (5.4)$$

Repeating the steps leading to (3.8) one arrives at

$$\begin{aligned} \log \mathcal{Z} = & \frac{T}{2} \sum_q (BK - A + 2D(1 - \cos q) \\ & - \sqrt{(BK - A + 2D(1 - \cos q))^2 + V(Q + RK^2)\theta}). \end{aligned} \quad (5.5)$$

We choose to analyze only the most interesting regime, $D \gg BK - A$, where a nontrivial collective behavior emerges. Then, in the long wavelength, $1 - \cos q \rightarrow q^2/2$, and

Tail-constraining stochastic linear–quadratic control: a large deviation and statistical physics approach

continuous, $\sum_q \rightarrow (N/2\pi) \int dq$, limits, one derives

$$\log \mathcal{Z} = -\frac{TN}{3\pi} \frac{(BK - A)^{3/2}}{D^{1/2}} (1 + s)^{1/4} i \times \left((1 + \sqrt{1 + s}) \mathcal{K} \left(\frac{1}{2} - \frac{1}{2\sqrt{1 + s}} \right) - 2\mathcal{E} \left(\frac{1}{2} - \frac{1}{2\sqrt{1 + s}} \right) \right), \quad (5.6)$$

$$s = \frac{V(Q + RK^2)\theta}{(BK - A)^2}, \quad (5.7)$$

where one utilizes the standard \mathcal{K}, \mathcal{E} notation for the elliptic functions.

The expression on the rhs of equation (5.6) shows a singularity at $s = -1$, coinciding with the singularity (in the complex θ plane) observed in the scalar case at θ_* . Substituting equation (5.6) into (3.2) and evaluating the integral over θ in the saddle-point approximation, one arrives at

$$\mathcal{S}_*(j) \approx \frac{\sqrt{2} N (BK - A)^{3/2}}{3\pi D^{1/2}} + \theta_* j. \quad (5.8)$$

Juxtaposing the string expression equation (5.8) to the scalar one equation (3.9), one notes different behaviors with respect to $BK - A$. Optimizing equation (5.8) over K at a given large value $J = jT$, one obtains the same tail expression, the second formula in (4.9)—however, with another optimal control,

$$K_{\text{str}}^{5/2} = \frac{2^{3/2} \pi A D^{1/2}}{RVTNB^{1/2}} J, \quad (5.9)$$

replacing the first formula in equation (4.9). Note that in the string case the optimal K scales as $J^{2/5}$, which should be contrasted with the $J^{1/2}$ scaling in the scalar case from equation (4.9).

6. Conclusions and the path forward

This paper contributes to a subject in control theory—designing control scheme with some guarantees not only on the average of the cost functions but also on fluctuations, and specifically extreme fluctuations related to the tail of the cost function PDF. We consider a linear stochastic system, of first order in time derivatives, of the Langevin type, subject to the minimization of a quadratic cost function and also with (chance) constraints imposed on the tail of the cost function PDF. For the stationary regime of large time, for when the control is sufficient to make the system stable, we reduce a stochastic dynamic problem of the ‘field theory’ type to static optimization analysis with objectives and constraints stated in a matrix form. This type of reduction is unusual for a system lacking a fine-tuned fluctuation-dissipation relation between the relaxational and stochastic terms. On the other hand, the progress made is linked to the linearity of the underlying stochastic systems which allowed us, as in some problems relating to passive scalar turbulence [39]–[41] and driven linear elastic systems [42, 38], to formally express the solution for the system trajectory as an explicit function of the noise realization. Besides that, the main technical ingredients which allowed us to derive the results consisted in making a plausible assumption about the structure of the control (linear in the state variable and frozen in time), and then performing asymptotic evaluations of

the cost function statistics conditioned on the value of the cost matrix. Techniques of path integration, spectral analysis and large deviation estimations were used. We tested results on the simple scalar case and illustrated the utility of the method on an example system of high dimension (a 1D chain of particles connected in a string).

We plan to continue exploring the interface between control theory and statistical physics, addressing the following challenges.

- The computational feasibility of the main formulas of the paper, stating RS, TO and CC controls in equations (4.1), (4.3) and (4.5) as static optimization problems, needs to be analyzed for large systems and networks. After all, the main efforts in applied control theory go into designing efficient algorithms for discovering optimal, or close to optimal, control, and we do plan to contribute on this important task. Therefore, further analysis is required to answer an important practical question: that of whether the static formulations of the newly introduced TO-QG and CC-QG controls allow computationally favorable exact or approximate expressions in terms of convex optimizations.
- We also plan to study weakly non-linear stochastic systems through a singular perturbation stochastic diagrammatic technique of the Martin–Siggia–Rose type [43]. Besides this, some of the methods that we used in the paper, especially those related to large deviation analysis, are not restricted to linear systems. Our preliminary tests show that the effects of the non-linearity on the PDF tail are seriously enhanced in comparison with how the same non-linearity influences the average case control.
- It will be interesting to study TO and CC versions of the path integral non-linear control problems discussed in [13]–[16]. These problems, in their standard ‘min-cost’ formulations, allow reduction (under some relations like those of the fluctuation-dissipation theorem among the form of control, the covariance matrix of the noise and the cost function) from the generally non-linear Hamilton–Jacobi–Bellman equations for the optimal cost function to a linear equation of a Schrödinger type.
- The effects of partial observability and noise in the observations can be easily incorporated in both the TO and CC schemes discussed in the paper. In fact this type of generalization is standard and widespread in control theory, where for example the LQG (linear–quadratic–Gaussian) control generalizes the LQ control.
- In terms of relevance to an application, this work was motivated by recent interest in and discussion related to developing new optimization and control paradigms for power networks, so-called smart grids. In this application, strong fluctuations associated with loads and renewable generation, electromechanical control of generation, and the desire to make energy production cheaper while also (and most importantly) maintaining probabilistic security limitations of the chance-constrained type all make the theoretical model discussed in this paper an ideal framework for consideration. In particular, we plan to extend the approaches of [44, 45] and modify and apply the theory developed in this paper to design a multi-objective chance-constrained optimum power flow including better control of the generation, loads and storage resources in power grids.
- We also anticipate that some of the models and results discussed in this paper will be of interest in relation to problems in statistical microfluidics and biofluidics, focusing on adjusting characteristics of individual molecules (polymers, membranes, etc) and also

aimed at modifying properties of the medium (non-Newtonian flows) macroscopically. The time-independent and linear nature of the control schemes discussed in this paper make them especially attractive for these applications. Natural constraints, e.g. associated with the force field (optical or mechanical) as well as with some other physical limitations, could be incorporated into control as single-objective or multi-objective cost functions.

Acknowledgments

We are grateful to D Bienstock, L Gurvits, H J Kappen, K Turitsyn and participants of the ‘Optimization and Control Theory for Smart Grids’ project at LANL for motivating discussions and remarks. The research at LANL was carried out under the auspices of the National Nuclear Security Administration of the US Department of Energy at Los Alamos National Laboratory under Contract No. DE C52-06NA25396.

References

- [1] Weiner N, 1949 *Extrapolation, Interpolation, and Smoothing of Stationary time Series with Engineering Applications* (New York: Wiley)
- [2] Kalman R E, *Contributions to the theory of optimal control*, 1960 *Bol. Soc. Mat. Mex.* **5** 102
- [3] Kalman R E, *A new approach to linear filtering and prediction problems*, 1960 *Trans. ASME J. Basic Eng.* **82** 35
- [4] Stratonovich R L, *Conditional Markov process theory*, 1960 *Theory Prob. Appl. (USSR)* **5** 156
- [5] Kalman R E and Bucy R S, *New results in lineal filtering and prediction theory*, 1961 *Trans. ASME J. Basic Eng. D* **83** 95
- [6] Kalman R E, *New methods of Wiener filtering theory*, 1963 *Proc. 1st Symp., Engineering Applications of Random Function Theory and Probability* ed J L Bogdanoff and F Korin (New York: Wiley) pp 270–385
- [7] Kailath T, *A view of three decades of linear filtering theory*, 1974 *IEEE Trans. Inform. Theory* **20** 146
- [8] Einstein A, *Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt*, 1905 *Ann. Phys.* **322** 132
- [9] von Smoluchowski M, *Zur kinetischen theorie der Brownschen molekularebewegung und der suspensionen*, 1906 *Ann. Phys.* **326** 756
- [10] Ulam S, *Marian Smoluchowski and the theory of probabilities in physics*, 1957 *Am. J. Phys.* **25** 475
- [11] van Kampen N G, 1992 *Stochastic Processes in Physics and Chemistry* (Amsterdam: Elsevier)
- [12] Gardiner C W, 2004 *Handbook of Stochastic Methods: for Physics, Chemistry and the Natural Sciences* (Berlin: Springer)
- [13] Kappen H J, *Path integrals and symmetry breaking for optimal control theory*, 2005 *J. Stat. Mech.* **P11011**
- [14] Kappen H J, *Optimal control theory and the linear bellman equation*, 2011 *Inference Learn. Dyn. Models* 363
- [15] van den Broek J L, Wiegerinck W A J J and Kappen H J, *Risk sensitive path integral control*, 2010 *Uncert. Artif. Intell.* **26** 1
- [16] Dvijotham K and Todorov E, *A unifying framework for linearly solvable control*, 2011 *Proc. 27th Conf. Annual Conf. on Uncertainty in Artificial Intelligence (UAI-11)* (Corvallis: AUAI Press) pp 179–86
- [17] Jarzynski C, *Equalities and inequalities: irreversibility and the second law of thermodynamics at the nanoscale*, 2011 *Annu. Rev. Condens. Matter Phys.* **2** 329
- [18] Bustamante C, Liphardt J and Ritort F, *The nonequilibrium thermodynamics of small systems*, 2005 *Phys. Today* **58** 43
- [19] McAdams H and Arkin A, *Stochastic mechanisms in gene expression*, 1997 *Proc. Natl Acad. Sci.* **94** 814
- [20] Schmiedl T and Seifert U, *Stochastic thermodynamics of chemical reaction networks*, 2007 *J. Chem. Phys.* **126**
- [21] Chernyak V, Chertkov M, Goldberg D A and Turitsyn K, *Non-equilibrium statistical physics of currents in queuing networks*, 2010 *J. Stat. Phys.* **140** 819
- [22] Chernyak V Y, Chertkov M and Sinityn N A, *The geometric universality of currents*, 2011 *J. Stat. Mech.* **P09006**

- [23] Altarelli F, Braunstein A, Ramezanpour A and Zecchina R, *Stochastic optimization by message passing*, 2011 *J. Stat. Mech.* **P11009**
- [24] Willems J and Mitter S, *Controllability, observability, pole allocation, and state reconstruction*, 1971 *IEEE Trans. Automat. Control* **16** 582
- [25] Athans M, *The role and use of the stochastic linear–quadratic–Gaussian problem in control system design*, 1971 *IEEE Trans. Automat. Control* **16** 529
- [26] Jacobson D, *Optimal stochastic linear systems with exponential performance criteria and their relation to deterministic differential games*, 1973 *IEEE Trans. Automat. Control* **18** 124
- [27] Whittle P, *Risk-sensitive linear/quadratic/Gaussian control*, 1981 *Adv. Appl. Probab.* **13** 764
- [28] Whittle P and Kuhn J, *A Hamiltonian formulation of risk-sensitive linear/quadratic/Gaussian control*, 1986 *Int. J. Control* **43** 1
- [29] Glover K and Doyle J C, *State-space formulae for all stabilizing controllers that satisfy an H -infinity norm bound and relations to risk sensitivity*, 1988 *Syst. Control Lett.* **11** 167
- [30] Petersen I R, James M R and Dupuis P, *Minimax optimal control of stochastic uncertain systems with relative entropy constraints*, 2000 *IEEE Trans. Automat. Control* **45** 398
- [31] Charnes A, Cooper W W and Symonds G H, *Cost horizons and certainty equivalents: an approach to stochastic programming of heating oil*, 1958 *Manage. Sci.* **4** 235
- [32] Miller L B and Wagner H M, *Chance-constrained programming with joint constraints*, 1965 *Oper. Res.* **13** 930
- [33] Ben-Tal A, El Ghaoui L and Nemirovski A, 2009 *Robust Optimization* (Princeton, NJ: Princeton University Press)
- [34] Franklin G F, Powell J D and Emami-Naeini A, 2002 *Feedback Control of Dynamic Systems* 4th edn (Upper Saddle River, NJ: Prentice-Hall)
- [35] Zhou K, Doyle J C and Glover K, 1996 *Robust Control and Optimal Control* (Englewood Cliffs, NJ: Prentice-Hall)
- [36] Hespanha J P, 2009 *Linear Systems Theory* (Princeton, NJ: Princeton University Press)
- [37] Zhou K, 1998 *Essentials of Robust Control* (Englewood Cliffs, NJ: Prentice-Hall)
- [38] Turitsyn K, Chertkov M, Chernyak V Y and Puliafito A, *Statistics of entropy production in linearized stochastic systems*, 2007 *Phys. Rev. Lett.* **98** 180603
- [39] Chertkov M, Falkovich G, Kolokolov I and Lebedev V, *Statistics of a passive scalar advected by a large-scale two-dimensional velocity field: analytic solution*, 1995 *Phys. Rev. E* **51** 5609
- [40] Chertkov M, Falkovich G and Kolokolov I, *Intermittent dissipation of a passive scalar in turbulence*, 1998 *Phys. Rev. Lett.* **80** 2121
- [41] Falkovich G, Gawędzki K and Vergassola M, *Particles and fields in fluid turbulence*, 2001 *Rev. Mod. Phys.* **73** 913
- [42] Kwon C, Ao P and Thouless D J, *Structure of stochastic dynamics near fixed points*, 2005 *Proc. Natl Acad. Sci. USA* **102** 13029
- [43] Martin P C, Siggia E D and Rose H A, *Statistical dynamics of classical systems*, 1973 *Phys. Rev. A* **8** 423
- [44] Dvijotham K, Backhaus S and Chertkov M, *Operations-based planning for placement and sizing of energy storage in a grid with a high penetration of renewables*, 2011 arXiv:1107.1382
- [45] Harnett S, Bienstock D and Chertkov M, *Chance-constrained optimal power flow: risk-aware network control under uncertainty*, 2012 LANL preprint