# Energy Saving Approximations For Random Processes

M. Lifshits

August, 22 2016

(VI International Conference Modern Problems in Theoretical and Applied Probability)

This is a joint work with I. Ibragimov from PDMI, E. Setterqvist from Linköping university, Sweden, and Z. Kabluchko from Münster university, Germany.

# Example: running after a Brownian dog

How to keep the Brownian dog on a leash in the energy saving mode?

Let the dog walk in $\mathbb{R}$ according to a Brownian motion $W(t)$.

You must follow it by moving with a finite speed and always stay not more than 1 away from the dog.

If $x(t)$ is your trajectory, then the goal is to follow the dog, i.e. keep $|x(t) - W(t)| \leq 1$ and expend minimal kinetic energy per unit of time

$$\frac{1}{T} \int_0^T x'(t)^2 \, dt$$

in a long run, $T \to \infty$.

# Diffusion strategy for the pursuit

Let $X(t) := x(t) - W(t)$ be the signed distance to the dog.
A reasonable strategy is to determine the speed $x'(t)$ as a function of $X(t)$ by accelerating when $X(t)$ approaches the boundary $\pm 1$. So let $x'(t) := b(X(t))$
Then $X$ becomes a stationary diffusion satisfying

$$dX = b(X)dt - dW.$$

One-dimensional diffusions are well understood. The density of the invariant measure is

$$p(x) = C\, e^{B(x)}, \qquad \text{where } B(x) := 2\int^x b(y)dy.$$

By ergodic theorem, in the stationary regime

$$\tfrac{1}{T}\int_0^T x'(t)^2\, dt \to \frac{1}{4}\int_{-1}^{1} b(x)^2\, p(x)\, dx = \frac{1}{4}\int_{-1}^{1} \tfrac{p'(x)^2}{p(x)^2}\, p(x)\, dx := \frac{1}{4}\, I(p).$$

We have to minimize Fisher information I(p) !

# Solution: optimal strategy

Minimizing Fisher information on the interval is a classical problem arising in Statistics, Data Analysis, etc (Zipkin, Huber, Levit, Shevlyakov, etc).
By simple variational calculus we obtain the optimal density

$$p(x) = \cos^2(\pi x/2), \qquad x \in [-1, 1],$$

and the optimal speed strategy

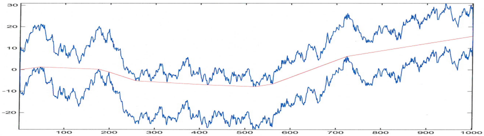
$$b(x) = -\pi \tan(\pi x/2)$$

exploding at the boundary.
This leads to the asymptotic minimal reduced energy

$$\frac{1}{T} \int_0^T x'(t)^2 \, dt \to \frac{1}{4} \, I(p) = \frac{\pi^2}{4}.$$

# Non-adaptive setting: taut string

$$\begin{cases} \int_0^T f'(t)^2 dt \searrow \min & \text{or(!!)} \int_0^T \varphi(f'(t))dt \searrow \min \\ f(0) = w(0), \quad f(T) = w(T), & \\ w(t) - r \le f(t) \le w(t) + r, \qquad 0 \le t \le T. & \end{cases}$$



# Formal setting

We consider uniform norms

$$||h||_T := \sup_{0 \le t \le T} |h(t)|, \qquad h \in \mathbb{C}[0,T],$$

and Sobolev-type norms (average kinetic energy)

$$|h|_T^2 := \int_0^T h'(t)^2 dt, \qquad h \in AC[0,T].$$

Let $W$ be a Brownian motion. We are mostly interested in its approximation characteristics

$$I_W(T,r) := \inf\{|h|_T; h \in AC[0,T], ||h-W||_T \le r, h(0)=0\}.$$

# Main results for non-adaptive approximation

**Theorem**

*There exists $\mathcal{C} \in (0,\infty)$ such that for any $q > 0$ if $\frac{r}{\sqrt{T}} \to 0$, then*

$$\frac{r}{T^{1/2}}\, I_W(T,r) \xrightarrow{L_q} \mathcal{C}$$

We may complete the mean convergence with a.s. convergence to $\mathcal{C}$.

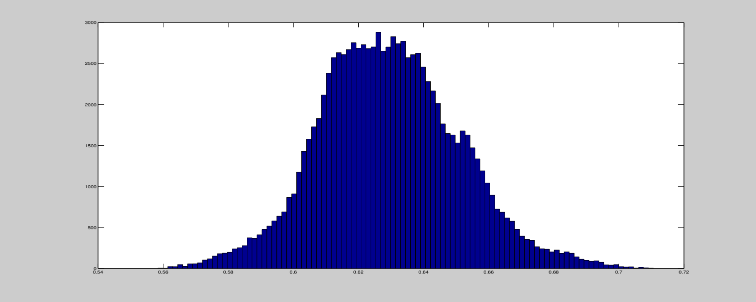
**Theorem**

*For any fixed $r > 0$, when $T \to \infty$, we have*

$$\frac{r}{T^{1/2}}\, I_W(T,r) \xrightarrow{a.s.} \mathcal{C}.$$

Main proof ideas: Gaussian concentration and subadditivity in time.

# Empirical modelling of $\mathcal{C}$




$$\mathcal{C} \approx 0.63$$

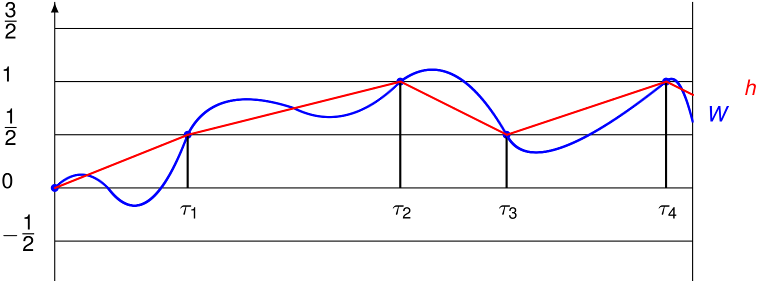
Comparing to the optimal pursuit,

$$0.63 \approx \mathcal{C} \le \frac{\pi}{2} \approx 1.51.$$

This is a price to pay for not knowing the future.
Theoretical lower and upper bounds for $\mathcal{C}$ are also available.

# Upper bound: free-knot approximation

Let $\tau_{n+1} := \inf\left\{t \geq \tau_n \middle| |W(t) - W(\tau_n)| \geq \frac{1}{2}\right\}$ Let $h(t)$ interpolate between the points $(\tau_n, W(\tau_n))$.




Then $\forall t$ we have $|h(t) - W(t)| \leq 1$ and

$$\int_{\tau_n}^{\tau_{n+1}} h'(t)^2 dt = \frac{(h(\tau_{n+1}) - h(\tau_n))^2}{\tau_{n+1} - \tau_n} = \frac{1}{4(\tau_{n+1} - \tau_n)}$$

are i.i.d. random variables.

# Free-knot approximation - numbers

On the long interval $[0, T]$ we have approximately $\frac{T}{\mathbb{E}\tau_1}$ cycles, and the average energy of $h$ on a cycle is $\mathbb{E}\frac{1}{4\tau_1}$. By the Law of Large Numbers,

$$\mathcal{C}^2 \le \lim_{T\to\infty} \frac{|h|_T^2}{T} = \frac{\mathbb{E}(\frac{1}{\tau_1})}{4\mathbb{E}\tau_1}.$$

We are able to calculate both expectations. First, by Wald identity,

$$\mathbb{E}\tau_1 = \mathbb{E}W(\tau_1)^2 = 1/4.$$

Second, it is easy to see that $\frac{1}{\tau_1}$ is equidistributed with $4\sup_{0\le t\le 1}|W(t)|^2$. It remains to evaluate $\mathbb{E}\sup_{0\le t\le 1}|W(t)|^2$. For exponential moment $\theta$ independent of $W$ we have

$$\mathbb{E}\sup_{0\le t\le 1}|W(t)|^2 = \mathbb{E}\sup_{0\le t\le \theta}|W(t)|^2 = \int_0^\infty \frac{x\,dx}{cosh(x)} \approx 1.832.$$

Thus $\mathcal{C} \le 2\sqrt{1.832} \approx 2.7$.

# An extended setting: "Pursuit under Potential"

Consider a fixed time horizon $[0, T]$, introduce a penalty function (potential) $Q(\cdot)$. Problem: find a pursuit process $X(\cdot)$ such that

$$\mathbb{E} \int_0^T \left[ X'(t)^2 + Q(X(t) - W(t)) \right] dt \searrow \min$$

among all adapted absolutely continuous random functions $X$.
We also consider an infinite horizon problem stated as

$$\lim_{T\to\infty} T^{-1}\, \mathbb{E} \int_0^T \left[ X'(t)^2 + Q(X(t) - W(t)) \right] dt \searrow \min$$

By appropriate interpretation of $Q$ this setting formally includes the Brownian dog problem, whenever

$$Q(y) := \begin{cases} 0, & |y| \le 1, \\ +\infty, & |y| > 1. \end{cases}$$

# A strategy of optimal pursuit

Strategy: $\quad X'(t) := b(X - W, T - t)$.
At every moment we determine the pursuit speed as a prescribed function of two arguments: the current distance from the target $W$ and the remaining time $T - t$. We show that this kind of strategy is the best among all adapted strategies on every finite interval of time provided that the drift function $b(\cdot, \cdot)$ is chosen properly.
Consider the expected penalty function achievable on the time interval of length $t$ when starting at the point $X(0) = y$,

$$\begin{aligned} F(y,t) &:= \mathbb{E} \int_0^t \Big[ X'(s)^2 + Q(X(s) - W(s)) \Big] \, ds \\ &= \mathbb{E} \int_0^t \Big[ b(Y(s), t-s)^2 + Q(Y(s)) \Big] \, ds. \end{aligned}$$

A version of Feynman-Kac formula leads to an equation quite close to Burgers equation. Therefore, Hopf–Cole transform
$F(y,t) := -2 \ln V(y,t)$ leads to some form of heat equation.

# Heat equation and survival probability

For the heat equation, we can find a good probabilistic solution, see Borodin and Salminen's "Handbook of Brownian motion". We find there

$$V(y,t) = \mathbb{E}\exp\left\{-\frac{1}{2}\int_0^t Q(W_y(s))\,ds\right\},$$

where $W_y$ stands for a Brownian motion starting at a point $y$. This is the survival probability under killing rate $Q$ ! For the Brownian dog problem we just have $V(y,t) = \mathbb{P}\left(|W_y(s)| \le 1, 0 \le s \le t\right)$ which, for large $t$, is nothing but small ball probability. Once the optimal energy $F(y,t)$ is found, we may found the optimal speed strategy $b(y,t)$. Looking at the final result, we discover that the distortion $Y = X - W$ of the optimal pursuit coincides with the Brownian motion conditioned to survive under the killing rate $Q$ ! For the Brownian dog problem, we get the Brownian motion conditioned to stay in the strip $[-1,1]$. For quadratic potential $Q(y) = y^2$ we get $b(y,t) = -\tanh(t)\,y \sim -y$ (for large $t$) which corresponds to the Ornstein – Uhlenbeck process.

# Infinite intervals

We search an adapted and absolutely continuous pursuit $X$ minimizing asymptotic energy per unit of time

$$\lim_{T\to\infty} T^{-1}\, \mathbb{E} \int_0^T \Big[ X'(t)^2 + Q(X(t)-W(t)) \Big]\, dt.$$

Again, a natural candidate for being an optimal pursuit is a process $X$ satisfying $X'(t) := b(X-W)$, where now the speed depends only on the distortion. This strategy is optimal provided that the drift function $b(\cdot)$ is chosen properly.
Optimization arguments and the variable change $b = V'/V$ lead to the eigenvalue problem for 1-dimensional Shrödinger equation

$$V''(y) - Q(y)\, V(y) = -\lambda\, V(y).$$

We conclude that the minimal asymptotic energy in the stationary regime is equal to the minimal eigenvalue of the respective Shrödinger equation, while the optimal speed function $b(y)$ is equal to the log-derivative of the corresponding eigenfunction.

# Generalization

Brownian motion $\nearrow$ general process with stationary increments or a stationary process.

Kinetic energy $\nearrow$ general form of energy.

General potential $Q$ $\searrow$ quadratic potential $Q(y) = \alpha y^2$.
This makes possible to consider the $L_2$ (or wide sense) setting.

# Problem setting

Let $(B(t))_{t\in\Theta}$ with $\Theta = \mathbb{Z}$ or $\Theta = \mathbb{R}$ be a wide sense stationary process with discrete or continuous time.

The classical linear prediction problem consists of finding an element in $\overline{\mathrm{span}}\{B(s), s \le t\}$ providing the best possible mean square approximation to the variable $B(\tau)$ with $\tau > t$.

We investigate this and some other similar problems where, in addition to prediction quality, optimization takes into account other features of the objects we search for.

One of the most motivating examples of this kind is an approximation of $B$ by a stationary differentiable process $X$ taking into account the kinetic energy that $X$ spends in its approximation efforts. The goals of the approximation quality and energy saving may be naturally combined with averaging in time by minimization of the functional

$$\lim_{N\to\infty} \frac{1}{N} \int_0^N \left[ |X(t) - B(t)|^2 + \alpha^2 |X'(t)|^2 \right] dt,$$

where $\alpha > 0$ is a balancing parameter.

# Problem setting: continued

If, additionally, the process $X(t) - B(t)$ and the derivative $X'(t)$ are stationary processes in the strict sense, in many situations ergodic theorem applies and the limit

$$\lim_{N\to\infty} \frac{1}{N} \int_0^N \left[ |X(t) - B(t)|^2 + \alpha^2 |X'(t)|^2 \right] dt,$$

is equal to $\mathbb{E}|X(0) - B(0)|^2 + \alpha^2 \mathbb{E}|X'(0)|^2$. Setting aside ergodicity issues, we may solve the problem

$$\mathbb{E}|X(0) - B(0)|^2 + \alpha^2 \mathbb{E}|X'(0)|^2 \to \min .$$

This problem makes sense either in a linear non-adaptive setting, i.e. with $X(0) \in \overline{\mathrm{span}}\{B(s), s \in \mathbb{R}\}$, or in a linear adaptive setting by requiring additionally $X(0) \in \overline{\mathrm{span}}\{B(s), s \leq 0\}$.

# Generalized energy

More generally, consider $H := \overline{\mathrm{span}}\{B(t), t \in \Theta\}$ as a Hilbert space equipped with the scalar product $(\xi, \eta) = \mathbb{E}(\xi\overline{\eta})$. For $T \subset \Theta$ let $H(T) := \overline{\mathrm{span}}\{B(t), t \in T\}$.

Furthermore, let $L$ be a linear operator with values in $H$ and defined on a linear subspace $\mathcal{D}(L) \subset H$. Consider the extremal problem

$$\mathbb{E}|Y - B(0)|^2 + \mathbb{E}|L(Y)|^2 \to \min,$$

where the minimum is taken over all $Y \in H(T) \bigcap \mathcal{D}(L)$. The first term in the sum describes approximation, prediction, or interpolation quality while the second term stands for additional properties of the object we are searching for, e.g. for the smoothness of the approximating process.

This is the most general form of the problem we are interested in. It includes (with $L = 0$) the classical prediction and interpolation problems.

# Operators $L$: spectral representation

Recall the spectral representation

$$B(t) = \int e^{itu} \mathcal{W}(du)$$

where $\mathcal{W}$ is an orthogonal random measure with $\mathbb{E}|\mathcal{W}(A)|^2 = \mu(A)$, $\mu$ being the spectral measure of $B$.
The operators $L$ we handle are those of the form

$$L\left(\int \phi(u)\mathcal{W}(du)\right) := \int \ell(u)\phi(u)\mathcal{W}(du).$$

For example, in continuous time case differentiation (kinetic energy operator) corresponds to $\ell(u) = iu$.
Similarly, in discrete time case difference operator corresponds to $\ell(u) = e^{iu} - 1$.

# Optimal non-adaptive approximation

For non-adaptive approximation the unique solution of the problem

$$\mathbb{E}|Y - B(0)|^2 + \mathbb{E}|L(Y)|^2 \to \min,$$

exists and is given by

$$Y = \int \frac{1}{1 + |\ell(u)|^2} \, \mathcal{W}(du)$$

and the corresponding minimum is equal to

$$\int \frac{|\ell(u)|^2}{1 + |\ell(u)|^2} \, \mu(du).$$

The same result holds for the processes with stationary increments.

# Optimal non-adaptive approximation: kinetic energy

In continuous time case the problem

$$\mathbb{E}|X(0) - B(0)|^2 + \alpha^2 \mathbb{E}|X'|^2 \to \min$$

is solved by the double-sided exponential moving average

$$X(t) = \frac{1}{2\alpha} \int_{\mathbb{R}} \exp\{-|\tau|/\alpha\} B(t+\tau) d\tau.$$

This is indeed non-adaptive! Interestingly, the form of the solution does not depend on the spectral measure of $B$.

Analogously, in the discrete time case the problem

$$\mathbb{E}|X(0) - B(0)|^2 + \alpha^2 \mathbb{E}|X(1) - X(0)|^2 \to \min$$

is solved by the double-sided series

$$X(t) = \frac{1}{\sqrt{1+4\alpha^2}} \left( B(t) + \sum_{k=1}^{\infty} \beta^{-k} \left( B(t+k) + B(t-k) \right) \right)$$

with $\beta = \frac{2\alpha^2+1+\sqrt{1+4\alpha^2}}{2\alpha^2}$ (the golden section while $\alpha = 1$).

# Kolmogorov-type criterion of error-free prediction

Let $(B(t))_{t\in\mathbb{Z}}$, be a wide sense centered stationary sequence and $\mu$ is its spectral measure. Let us represent $\mu$ as the sum $\mu = \mu_a + \mu_s$ of its absolutely continuous and singular components. We denote by $f$ the density of $\mu_a$ with respect to the Lebesgue measure and $H_t = \overline{\mathrm{span}}\{X(s), s \le t\}$ while $H = \overline{\mathrm{span}}\{X(s), s \in \mathbb{Z}\}$. Let

$$\sigma^2(t) := \inf_{Y\in H_t} \left\{ \mathbb{E}|Y - B(0)|^2 + \mathbb{E}|LY|^2 \right\}, \qquad t \in \mathbb{Z},$$

be the corresponding prediction errors and $\sigma^2(\infty)$ be the similar quantity with $H_t$ replaced by $H$. Then $\sigma^2(t) \searrow \sigma^2(\infty)$ as $t \nearrow +\infty$.
For the classical prediction problem, i.e. for $L = 0$, by Kolmogorov's singularity criterion we have $\sigma^2(t) = \sigma^2(\infty) = 0$ for all $t \in \mathbb{Z}$ iff $\int_{-\pi}^{\pi} |\ln f(u)| du = \infty$.
For $L \neq 0$, we have $\sigma^2(t) \ge \sigma^2(\infty) > 0$ and we state the problem as follows: when $\sigma^2(t) = \sigma^2(\infty)$ holds? When approximation based on the knowledge of the process up to time $t$ works as well as the one based on the knowledge of the whole process?

# Criterion of the singular prediction

### Theorem

*Let $B$ be a discrete time, wide sense stationary process. Let $L$ be a linear operator corresponding to a frequency characteristic $\ell(\cdot)$. Then for $t \in \mathbb{Z}$ the equality $\sigma^2(t) = \sigma^2(\infty)$ holds iff either $\int_{-\pi}^{\pi} |\ln f(u)| du = \infty$ holds, or $\int_{-\pi}^{\pi} |\ln f(u)| du < \infty$ holds, $t \geq 0$ and $\frac{1}{1+|\ell(u)|^2}$ is a trigonometric polynomial of degree not exceeding $t$, i.e.*

$$(1 + |\ell(u)|^2)^{-1} = \sum_{|j| \leq t} b_j \, e^{iju}$$

*Lebesgue-a.e. with some coefficients $b_j \in \mathcal{C}$.*

For continuous time processes the results are similar, based on Krein's singularity criterion $\int_{-\infty}^{\infty} \frac{|\ln f(u)|}{1+u^2} du = \infty$.

# Interpolation

Consider the simplest case of interpolation problem in discrete time. Let $(B(t))_{t\in\mathbb{Z}}$, be a wide sense stationary sequence having spectral density $f$ and let $L$ be a linear operator with frequency characteristic $\ell(\cdot)$. Consider the extremal problem

$$\mathbb{E}|Y - B(0)|^2 + \mathbb{E}|LY|^2 \to \min, \qquad Y \in H_1^\circ,$$

where $H_1^\circ = \overline{\mathrm{span}}\{B(s), |s| \geq 1\}$. Let

$$\sigma_{\mathrm{int}}^2 = \inf_{Y\in H_1^\circ} \left(\mathbb{E}|Y - B(0)|^2 + \mathbb{E}|LY|^2\right)$$

denote the interpolation error.

The classical case of this problem, i.e. $L = 0$, was considered by A.N. Kolmogorov. He proved that precise extrapolation with $\sigma_{\mathrm{int}}^2 = 0$ is possible iff $\int_{-\pi}^{\pi} \frac{du}{f(u)} = \infty$. If the integral is convergent, then $\sigma_{\mathrm{int}}^2 = 4\pi^2 \left(\int_{-\pi}^{\pi} \frac{du}{f(u)}\right)^{-1}$.

We extend this result to the case of general $L$ as follows.

# Interpolation: extended setting

## Theorem

*If* $\int_{-\pi}^{\pi} \frac{du}{f(u)} = \infty$ *holds, then*

$$\sigma_{\rm int}^2 = \int_{-\pi}^{\pi} \frac{|\ell(u)|^2}{1+|\ell(u)|^2}\, f(u)\, du.$$

*Otherwise,*

$$\sigma_{\rm int}^2 = \int_{-\pi}^{\pi} \frac{|\ell(u)|^2 f(u) du}{1+|\ell(u)|^2} + \left( \int_{-\pi}^{\pi} \frac{du}{1+|\ell(u)|^2} \right)^2 \left( \int_{-\pi}^{\pi} \frac{du}{f(u)(1+|\ell(u)|^2)} \right)^{-1}.$$

# References


- Ibragimov, I.A., Kabluchko, Z., Lifshits M.A. Some extensions of linear approximation and prediction problems for stationary processes. In preparation.
- Kabluchko, Z., Lifshits M.A. Least energy approximation for processes with stationary increments. Preprint `arxiv 1506.08369`. To appear in J. Theor. Probab.
- Lifshits, M., Setterqvist, E. Energy of taut string accompanying Wiener process, Stoch. Proc. Appl., 125, 401–427 (2015).