Large Deviations in Quantum Information Theory

R. Ahlswede and V. Blinovsky

Abstract

We obtain asymptotic estimates on the probabilities of events of special types which are usefull in quantum information theory, especially in the theory of identification for noisy channels.

1 Introduction

Let T be the topological Hausdorff space with Borel σ–algebra $\mathcal{B}$. Also let $(P_n)_{n=1}^{\infty}$ be a sequence of distributions on T. We say that the large deviations principle (LDP) for $(P_n)_{n=1}^{\infty}$ is valid if there exists a functional $I : T \to \mathbb{R},\ I \not\equiv 0, \infty$ such that for arbitrary $B \in \mathcal{B}$ the following relations are valid

$$- \inf_{\xi \in B^o} I(\xi) \leq \liminf_{n \to \infty} \frac{\ln P_n(B)}{n} \leq \limsup_{n \to \infty} \frac{\ln P_n(B)}{n} \leq - \inf_{\xi \in \bar{B}} I(\xi), \tag{1}$$

where $\bar{B}$ (B^o) is the closure (open kernel) of the set B. In the case where T is a metric space we say that the local LDP for $(P_n)_{n=1}^{\infty}$ is valid if there exists a functional $I : T \to \mathbb{R},\ I \not\equiv 0, \infty$ such that for arbitrary $z \in T$

$$\lim_{\delta \to 0} \lim_{n \to \infty} \frac{\ln P_n(B_{z,\delta})}{n} = -I(z),$$

where $B_{z,\delta} = \{y \in T : d(y, z) \leq \delta\}$ is the ball of radius δ centered in z.

Let H_d be the space of Hermitian $d \times d$ matrices and let P_d be a distribution on it. Space H_d is naturally isomorphic to the linear space $\mathbb{R}^{d^2}$ and the dimension of it equals d^2. The large deviations theory in $\mathbb{R}^{d^2}$ is properly developed. We recall it. Let $t \in R^{d^2}$ and let

$$\Lambda(t) = \mathbb{E}e^{(t,z)} < \infty. \tag{2}$$

Then the value

$$\Lambda^*(\xi) = \sup_{t \in R^{d^2}} ((t, \xi) - \ln \Lambda(t))$$

is the rate function for the sequence of probabilities $(P_n)_{n=1}^{\infty}$, where

$$P_n(B) = \left(\frac{1}{n} \sum_{i=1}^{n} Z_i \in B \right)$$

$Z_i \in H_d, B \in \mathcal{B}$. If we consider in addition the exponential tightness of the sequence of probabilities $(P_n)_{n=1}^{\infty}$ then the LDP for $(P_n)_{n=1}^{\infty}$ is still valid (with rate function Λ^*), when $d = \infty$.

Actually one of the interesting problems for Quantum Information Theory is the estimation of the probability of the events (see [2])

$$B(C) = \{Z \not\leq C\}, \tag{3}$$

$$B'(C) = \{Z \not< C\}, \tag{4}$$

where $Z = \frac{1}{n} Z^n = \frac{1}{n} \sum_{i=1}^{n} Z_i$ and $(Z_i)_{n=1}^{\infty}$ is the sequence of i.i.d. random Hermitian matrices of finite dimension $d \times d$ and C is some Hermitian matrix. Note that $B_n(C)$ is open and when $d < \infty$ then $B'_n(C)$ is closed. Then we also consider the case when Z_i is random Hilbert-Schmidt operator on an infinite dimensional separable Hilbert space $\mathcal{H}$.

Let's consider at first that the dimension d of the ground space $\mathcal{H}$ is finite. Note that in this case sequence $(P_n)_{n=1}^{\infty}$ is exponentially compact.[1] It is easy to see that the sets $B_n(C), B'_n(C)$ are Borel sets.

Next we prove that if

$$EZ_i \notin B'(C), \tag{5}$$

then

$$-\inf_{x \in \mathcal{C}^d} \lim_{\delta \to 0} \Lambda'(C - I\delta, x) \leq \liminf_{n \to \infty} \frac{\ln P_n(B(C))}{n} \leq \limsup_{n \to \infty} \frac{\ln P_n(B(C))}{n} \leq -\inf_{x \in \mathcal{C}^d} \Lambda'(C, x), \tag{6}$$

and

$$\lim_{n \to \infty} \frac{\ln P_n(B'(C))}{n} = -\inf_{x \in C^d} \Lambda'(C, x), \tag{7}$$

where

$$\Lambda'(C, x) = \sup_{t \in R} (t(x, Cx) - \ln \bar{\Lambda}(t, x)), \quad \bar{\Lambda}(x) = E e^{t(x, Z_i x)}$$

and I is the unit matrix. It is enough to optimize the expressions in the last relations only over unit vectors $x \in C^d$.

Note that if

$$EZ_i \in B(C),$$

then

$$\lim_{n \to \infty} \frac{\ln P_n(B(C))}{n} = \lim_{n \to \infty} \frac{\ln P_n(B'(C))}{n} = 0.$$

We will produce the proof of (6), (7) in such a way that it will be valid for $d = \infty$ after minor changes. First of all we prove that the limit in the LHS of (6) exists. Note that the

[1] This means that for every nonnegative real number δ there exists a compact set K_δ with $P_n(K_\delta) < e^{-\delta n}$ for all large n.

relations

$$Z^n = \sum_{i=1}^{n} Z_i \nleq Cn,$$
$$Z^n = \sum_{i=1}^{n} Z_i \nless Cn$$

mean that for some $x \in R^d$

$$(x, Z^n x) > (x, Cx)n$$

or

$$(x, Z^n x) \geq (x, Cx)n$$

corresspondingly.

The closure of the set $S = \{Z \nless nC\}$ is contained in the set

$$\bigcup_{x \in R^d} \{(x, Zx) > n((x, Cx) - \epsilon_x)\}, t > 0,$$

where ϵ_x are choosen in such a way that

$$(x, EZ_i x) < (x, Cx) - \epsilon_x.$$

Since the sequence $(P_n)_{n=1}^{\infty}$ is exponentially compact, for every $L < \infty$ one can choose the compact set $K \subset R^{d^2}$ such that for large enough n,

$$\frac{\ln(P_n(R^{d^2} \setminus K_L))}{n} n < -L. \tag{8}$$

Then the set $S' = \bar{S} \bigcap K_L$ is also compact and one can choose finitely many $x_1, x_2, \ldots, x_m \in R^d$ such that

$$S' \subset \bigcup_{j=1}^{m} \{(x_j, Zx_j) > (x_j, Cx_j) - \epsilon_{x_j}\}.$$

For every set $\{(x_j, Zx_j) > (x_j, Cx_j)\}$ the LDP for $(P_n)_{n=1}^{\infty}$ is valid with rate function $\Lambda'(C, x_j)$. Indeed in this case we deal with the one dimensional random variable (x_j, Zx_j) and by the previous condition $Ee^{((x_j, Z_i, x_j)t)} < \infty$. Therefore

$$\begin{aligned}
-\Lambda'(C - I\delta, x_j) &\leq \liminf_{n\to\infty} \frac{\ln P_n\left(t(x_j, Z^n x_j) > nt(x_j, Cx_j)\right)}{n} \leq \\
&\leq \limsup_{n\to\infty} \frac{\ln P_n\left(t(x_j, Z^n x_j) > nt(x_j, Cx_j)\right)}{n} \leq \\
&\leq \limsup_{n\to\infty} \frac{\ln P_n\left(t(x_j, Z^n x_j) > nt((x_j, Cx_j) - \epsilon_{x_j})\right)}{n} \leq \\
tag9 \\
&\leq -\Lambda'(C - I\epsilon, x_j)
\end{aligned}$$

and

$$
\begin{aligned}
-\Lambda'(C,x_j) &\le \liminf_{n\to\infty} \frac{\ln P_n\left(t(x_j, Z^n x_j) \ge nt(x_j, Cx_j)\right)}{n} \le \\
&\le \limsup_{n\to\infty} \frac{\ln P_n\left(t(x_j, Z^n x_j) \ge nt(x_j, Cx_j)\right)}{n} \le \\
&\le \limsup_{n\to\infty} \frac{\ln P_n\left(t(x_j, Z^n x_j) \ge nt((x_j, Cx_j) - \epsilon_{x_j})\right)}{n} \le \\
&\le -\Lambda'(C - I\epsilon_{x_j}, x_j).
\end{aligned} \tag{10}
$$

The RHS equalities in the relations (9) (10) are a consequence of the Chebyshev inequality. The LHS equalities in the relations (9) (10) are a consequence Cramer's theorem for semi-infinite intervals (see [1]: for an arbitrary $a \in R$ and a sequence of i.i.d. random variables $(\chi_i)_{i=1}^{\infty}$

$$
\lim_{n\to\infty} \frac{\ln P\left(\frac{1}{n}\sum_{i=1}^{n}\chi_i \in [a,\infty)\right)}{n} = -\inf_{x\ge a}\Gamma(x),
$$

where for every i

$$
\Gamma(x) = \sup_{t\in R}(xt - \ln Ee^{t\chi_i}).
$$

Now we choose

$$
L > \inf_{x\in\mathcal{C}^d} \Lambda'(C,x) \tag{11}
$$

and then we have

$$
\begin{aligned}
&-\inf_{x\in\mathcal{C}^d}\lim_{\delta\to 0}\Lambda'(C - I\delta, x) \le \liminf_{n\to\infty}\frac{\ln P_n(B(C))}{n} \le \limsup_{n\to\infty}\frac{\ln P_n(B(C))}{n} \le \\
&\le \limsup_{n\to\infty}\frac{\ln\left(P_n(B(C)\bigcap K_L) + P_n(R^d\setminus K_L)\right)}{n} \le && (12)\\
&\le -\inf_{x_j}(\Lambda'(C - I\epsilon_{x_j}), x_j)) + o(1), && (13)
\end{aligned}
$$

$$
\begin{aligned}
&-\inf_{x\in\mathcal{C}}\Lambda'(C,x) \le \lim_{n\to\infty}\frac{\ln P_n(B'(C))}{n} \le \\
&\le \limsup_{n\to\infty}\frac{\ln\left(P_n(B'(C)\bigcap K_L) + P_n(R^d\setminus K_L)\right)}{n} \le && (14)\\
&\le -\inf_{x_j}(\Lambda'(C - I\epsilon_{x_j}, x_j)) + o(1), && (15)
\end{aligned}
$$

where $o(t)$ is the rest term which follows from the relation (11) and the fact that we take $\inf_{x_j}$. Because we can change the range over which the infimum at the RHS of the last chain of relations is taken to $x \in R^d$, we obtain that

$$
\lim_{n\to\infty}\frac{\ln P_n(B'(C))}{n} = -\inf_{x\in\mathcal{C}^d}\Lambda'(C,x),
$$

$$
\inf_{x\in\mathcal{C}^d}\lim_{\delta\to 0}\Lambda'(C - I\delta, x) \le \liminf_{n\to\infty}\frac{\ln P_n(B(C))}{n} \le \limsup_{n\to\infty}\frac{\ln P_n(B(C))}{n} \le -\inf_{x_j}(\Lambda'(C - I\epsilon_{x_j}, x_j)).
$$

Since for every x_j, $(x_j, EZ_i x_j) < (x_j, Cx_j) - \epsilon_{x_j}$, the function $\Lambda'(C - I\epsilon_{x_j}, x_j)$ monotonically increase as ϵ_{x_j} decreases we can omit ϵ in the RHS of the relations (13), (15). Now we consider the case when $d = \infty$ and we consider the matrices, which corresspond to self-adjoint Hilbert-Schmidt operators on Hilbert space $\mathcal{H}$. We suppose that EZ_i is also H-S matrix. In this case all previous considerations are still valid under the assumption that $(P_n)_{n=1}^{\infty}$ is exponentially tight. Hence we should find out in which cases this assumption is valid. Recall that for the self-adjoint Hilbert-Schmidt matrix $Z = (z_{ij})$ the following relations are valid

$$||Z|| \leq ||Z||_2 \leq ||Z||_1,$$

where $||\cdot||$ is uniform norm ,

$$||Z||_2^2 \triangleq tr(Z^2) = \sum_{i,j} |a_{ij}|^2$$

is the square of the Hilbert-Schmidt norm and

$$||Z||_1 = tr|Z|$$

is the trace-norm (if it exists). Hence to every Hilbert-Schmidt (H-S) matrix corressponds in the natural order the sequence of reals which is ℓ_2-sequence and vice versa. If one considers the set of all ℓ_2-sequences $\mathcal{A} = \{a\}$ such that

$$|a_j| \leq b_j$$

for some given ℓ_2- sequence b of positive reals, then the set $\mathcal{A}$ is compact (in ℓ_2- norm). Hence the exponential tightness of the sequence $(P_n)_{n=1}^{\infty}$ is a consequence of the following condition. Let $b = \{b_1, b_2, \ldots, b_n, \ldots, \}$ be some ℓ_2- sequence of positive reals and let $Z = (z_i)$ be the R^{∞}- representation of self-adjoint H-S matrix Z. If for some natural n the sum

$$\pi_n = \sum_{j=1}^{\infty} \left(e^{-n\Omega_j^+(b_j)} + e^{-n\Omega_j^-(b_j)} \right), \tag{16}$$

where

$$\Omega_i^{\pm}(b) = \sup_{t \in R}((t, Ez_j \pm b_j) - \ln Ee^{(t,z_j)}), \tag{17}$$

converges, then the sequence $(P_n)_{n=1}^{\infty}$ is exponentially compact. Note that the RHS of inequality (16) is nothing else but the additive upper bound for the probability that $|Ez_i - z_i| > b_i$ for some i. It is easy to see that if (17) is valid, then for any given $L > 0$ one can choose i such that

$$\sum_{j>i} \left(e^{-\Omega_j^+(b_j)n} + e^{-\Omega_j^-(b_j)n} \right) < \frac{1}{i+1} e^{-Ln}.$$

Then we use the relation, which follows from the finiteness of $\Lambda(t)$,

$$\lim_{|b| \to \infty} \frac{\Omega_m^{\pm}(b)}{|b|} = \infty$$

to choose the values $\bar{b}_1, \bar{b}_2, \ldots, \bar{b}_i > 0$ such that

$$e^{-\Omega_m^{\pm}(\bar{b}_m)n} < \frac{1}{i+1} e^{-Ln}, m = 1, 2, \ldots, i.$$

Then we have

$$\pi_n < e^{-Ln} \tag{18}$$

and $\bar{b}_1, \bar{b}_2, \ldots, \bar{b}_i, b_{i+1}, \ldots$ is ℓ_2-sequence. From (18) it follows that P_n is exponentially compact.

Note that even if one considers the diagonal matrices $\mathcal{A}$ and distribution P on them (all elements on the diagonals of matrices from $\mathcal{A}$ are ℓ_2-sequences) and we have no information about the joint distributions of different diagonal elements the only possible upper bound for the probability $P(|Z^n - EZ^n| \not\leq Cn)$, where C is some self-adjoint bounded operator, is the additive bound (16), which can be obtained by using the Chebyshev inequality for the estimation of deviations $P(|z_j - Ez_j| > c_j n)$.

The last thing we should like to note is a convenient upper asymptotic bound on the probability of the event $Z^n \not\leq Cn$, which follows from our considerations. As in estimations (13) we use the Chebyshev estimation (x is unit vector)

$$\begin{aligned} P_n(B_n(C)) &\leq e^{-n\left(\inf_{x\in R^d}\left(t(x,Cx)-\ln Ee^{t(x,Z_j,x)}+\epsilon\right)\right)} = \\ &= e^{\epsilon n}\sup_{x\in R^d}\left(Ee^{t(x,Z_j,x)}e^{-t(x,Cx)}\right)^n \leq \\ &\leq e^{\epsilon n}\sup_{x\in R^d}\left((x, Ee^{t(Z_j-C)}, x)\right)^n \leq e^{\epsilon n}||Ee^{tZ_j-tC}||^n. \end{aligned}$$

Here the second inequality follows from the inequality

$$e^{(x,Zx)} \leq (x, e^Z x),$$

which in turn is a consequence of the convexity of e^y. Since $\epsilon > 0$ is arbitrary we have

$$\limsup_{n\to\infty} \frac{\ln P_n(B(C))}{n} \leq \inf_{t\geq 0} \ln ||E^{tZ_j-tC}||.$$

This formula under the assumption of exponential compactness is also valid for $d = \infty$.

Note also that the unitary transformation $U = \{u_{p,q}\}$ preserve the compactness and if for some H-S matrices $A = \{a_{i,j}\}, B = \{b_{i,j}\}$, $|a_{i,j}| \leq |b_{i,j}|$, then for the matrices $A' = UAU^* = \{a'_{i,j}\}$ and $B' = UBU^* = \{b'_{i,j}\}$ we have

$$|a'_{i,j}| = |\sum_{p,q} u_{i,p}a_{p,q}u^*_{q,j}| \leq \sum_{p,q} |u_{i,p}||a_{p,q}||u^*_{q,j}| \leq \sum_{p,q} |u_{i,p}||b_{p,q}||u^*_{q,j}|$$

and

$$\sum_{i,j} |a'_{i,j}|^2 \leq \sum_{p,q} |b_{p,q}|^2 < \infty.$$

Hence, we can view $\{|a_{i,j}|\}$ as a sequence (in any order). It has ℓ_2-majorant $\{|b_{i,j}|\}$ in some basis the conjugate sequence $a'_{i,j}$ has the property that $|a'_{i,j}|$ has the majorant $|UBU^*_{ij}|$. Hence one could check the validness of the relations (16), (17) in some most convenient basis for the corresponding estimations.

At the end we should like to state an open problem. Let $\{Z_j\}_{n=1}^{\infty}$ be a sequence of i.i.d. Hermitian $d \times d$ matrices. Then the following chain of relations is valid

$$Tr\left(E\left(e^{\sum_{i=1}^{n} Z_i}\right)\right) = E\left(Tr\left(e^{\sum_{i=1}^{n} Z_i}\right)\right) \leq$$
$$\leq E\left(Tr\left(e^{\sum_{i=1}^{m} Z_i}\right) Tr\left(e^{\sum_{i=m+1}^{n} Z_i}\right)\right)\right) = E\left(Tr\left(e^{\sum_{i=1}^{m} Z_i}\right)\right) E\left(Tr\left(e^{\sum_{i=m+1}^{n} Z_i}\right)\right) =$$
$$= Tr\left(E\left(e^{\sum_{i=1}^{m} Z_i}\right)\right) Tr\left(E\left(e^{\sum_{i=m+1}^{n} Z_i}\right)\right).$$

Here the first inequality is a consequence of the Golden - Thompson inequality (see [4], [3])

$$Tr\left(e^{A+B}\right) \leq Tr\left(e^A e^B\right).$$

Hence the sequence $(a_n)_{n=1}^{\infty}$ with

$$a_n = \ln Tr\left(E\left(e^{\sum_{i=1}^{n} Z_i}\right)\right)$$

is subadditive

$$a_n \leq a_m + a_{n-m}$$

and there exists the limit

$$a \triangleq \lim_{n\to\infty} \frac{a_n}{n}.$$

The problem now is to find an explicit expression for this limit in terms of a 'single letter characterization' i.e. in terms of the marginal distribution of Z_j.

If Z_j are commutative, then

$$\ln\left|\left|\left(Ee^{Z_i}\right)^n\right|\right| \leq a_n \leq \ln\left(d||E^{Z_i}||^n\right). \tag{19}$$

Because $E = Ee^{Z_i}$ is A self-adjoint operator

$$||E||^2 = ||E^2||$$

and hence from (19) and the convergence of $(a_n/n)_{n=1}^{\infty}$ it follows that

$$\lim_{n\to\infty} \frac{a_n}{n} = \lim_{k\to\infty} \frac{a_{2^k}}{2^k} = \ln ||Ee^{Z_i}||.$$

Also in any case

$$a_n \leq \ln d||Ee^{Z_i}||^n$$

and hence

$$\lim_{n\to\infty} \frac{a_n}{n} \leq \ln ||Ee^{Z_i}||.$$

From the inequalities

$$||Ee^{\sum_{i=1}^{n} Z_i}|| \leq TrEe^{\sum_{i=1}^{n} Z_i} \leq d||Ee^{\sum_{i=1}^{n} Z_i}||$$

it follows that

$$\lim_{n\to\infty} \frac{\ln ||Ee^{\sum_{i=1}^n Z_i}||}{n} = a.$$

References

[1] J.D. Deuschel and D.W. Strook, Large Deviations, Academic press, Boston, 1989.

[2] R. Ahlswede and A. Winter, Strong converse for identification via quantum channels, IEEE Trans. on Inf. Theory, Vol. 48, No. 3, 569–579, 2002.

[3] M. Reed and B. Simon, Functional Analysis, Academic Press. Vol. 1,2, New York and London, 1972.

[4] C.J. Thompson, Inequality with applications in statistical mechanics, J. Math. Phys. Vol. 6, No. 11, 1812-1823, 1965.