

Συγκέντρωση του μέτρου στον διακριτό κύβο με μεθόδους Στοχαστικής Ανάλυσης

Διπλωματική Εργασία

Ιάσων Ασκούνης

Τμήμα Μαθηματικών

Πανεπιστήμιο Αθηνών

Αθήνα – 2024

Περιεχόμενα

1	Πρόλογος	1
1.1	Boolean συναρτήσεις	1
1.2	Ισοπεριμετρικές ανισότητες στον διακριτό κύβο	2
2	Εισαγωγή	7
2.1	Συναρτήσεις στον διακριτό κύβο και η ανισότητα Poincaré	7
2.2	Υπερσυσταλτότητα: η ανισότητα της Bonami	13
2.3	Η ανισότητα Kahn-Kalai-Linial	19
2.4	Η εικασία του Talagrand	24
2.5	Ευστάθεια της ανισότητας του Talagrand	28
2.6	Αρμονική επέκταση Boolean συναρτήσεων	31
3	Εργαλεία από τον Στοχαστικό Λογισμό	37
3.1	Δεσμευμένη μέση τιμή και martingales	37
3.2	Στοχαστικές διαδικασίες και τετραγωνική κύμανση	40
3.3	Ανελίξεις στον συνεχή κύβο και σκιαγράφηση της μεθόδου	42
4	Το κύριο εργαλείο: Μια στοχαστική διαδικασία με άλματα	49
4.1	Κατασκευή και βασικές ιδιότητες	49
4.2	Η διαδικασία της επιρροής	53
4.3	Απόδειξη της βελτιωμένης εικασίας Talagrand	56
4.3.1	Περίπτωση 1n-Μικρά άλματα	58
4.3.2	Περίπτωση 2n-Μεγάλα άλματα	65
5	Η ανισότητα επιρροών του Talagrand και η ευστάθειά της	69
5.1	Η ανισότητα επιρροών του Talagrand	69
5.2	Απόδειξη του Θεωρήματος 2.4.5	75
5.3	Απόδειξη του Θεωρήματος 2.5.1	75
5.4	Ευσαιθησία θορύβου και επιρροές	79

6	Η μέθοδος των τυχαίων περιορισμών	85
6.1	Εισαγωγή	85
6.2	Τυχαίοι περιορισμοί και ισοπεριμετρικές ανισότητες στον διακριτό κύβο	88
6.3	Επεκτάσεις και ισχυρότερες εκδοχές	93
7	Ακριβείς ισοπεριμετρικές ανισότητες στον διακριτό κύβο	97
7.1	Εισαγωγή	97
7.1.1	Μονόπλευρο σύνορο	97
7.1.2	Διαμερίσεις του κύβου	100
7.1.3	Αμφίπλευρο σύνορο	101
7.1.4	Οι ισοπεριμετρικές ανισότητες του Talagrand για τη διακριτή κλίση	102
7.1.5	Hamming μπάλες	105
7.2	Αποδείξεις των θεωρημάτων	106
7.2.1	Απόδειξη του Θεωρήματος 7.1.1: λογαριθμική συνάρτηση	108
7.2.2	Απόδειξη του Θεωρήματος 7.1.3: τετραγωνικό πολυώνυμο	113
7.2.3	Απόδειξη του Θεωρήματος 7.1.4: κυβικό πολυώνυμο	115
7.2.4	Απόδειξη των Θεωρημάτων 7.1.11 και 7.1.13	118
8	Summary	121
8.1	Boolean functions	121
8.2	Isoperimetric inequalities on the discrete cube	122
	Βιβλιογραφία	127

1.1 Boolean συναρτήσεις

Θεωρούμε τον διακριτό κύβο $\{-1, 1\}^n$ εφοδιασμένο με το ομοιόμορφο μέτρο πιθανότητας

$$\left(\frac{1}{2}\delta_1 + \frac{1}{2}\delta_{-1}\right)^{\otimes n}.$$

Boolean συνάρτηση είναι κάθε συνάρτηση $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$, δηλαδή κάθε συνάρτηση που απεικονίζει διανύσματα n ψηφίων σε ένα ψηφίο. Οι Boolean συναρτήσεις θα είναι το κύριο μαθηματικό αντικείμενο που θα μελετήσουμε σε αυτή τη διπλωματική εργασία.

Παρόλη την απλότητα του ορισμού τους, οι Boolean συναρτήσεις μπορούν να αναπαριστούν πολλά πράγματα απο τελείως διαφορετικούς κλάδους και δεν προκαλεί έκπληξη το γεγονός ότι εμφανίζονται παντού. Παραθέτουμε μερικά παραδείγματα που δείχνουν κάποια από τα πεδία εφαρμογής τους:

Κοινωνική επιλογή: Έστω ότι έχουμε n -το πλήθος πολίτες της Αγγλίας και το ψηφίο $x_i \in \{-1, 1\}$ αντιπροσωπεύει την άποψη του i -οστού πολίτη για κάποιο ζήτημα πολιτικής φύσεως, για παράδειγμα, αν συμφωνεί ή διαφωνεί με την έξοδο της Αγγλίας απο την Ευρωπαϊκή Ένωση (Brexit). Το διάνυσμα $x = (x_1, \dots, x_n) \in \{-1, 1\}^n$ αντιπροσωπεύει την άποψη των n αυτών πολιτών, και κάθε Boolean συνάρτηση f αντιπροσωπεύει κάποιον τρόπο με τον οποίο λαμβάνεται απόφαση για το πώς πρέπει να δράσει η πολιτεία βασιζόμενη στις απόψεις αυτών των n πολιτών. Προφανώς υπάρχουν πάρα πολλοί τρόποι για να παρθούν αποφάσεις, κάτι που φαίνεται από την ποικιλία των διαφορετικών εκλογικών νόμων στις διάφορες χώρες του κόσμου. Ο πιο φυσιολογικός ίσως τρόπος είναι μέσω της *συνάρτησης πλειοψηφίας*, η οποία δίνει αποτέλεσμα 1 αν $\sum_{i=1}^n x_i > 0$ και αποτέλεσμα -1 αν $\sum_{i=1}^n x_i < 0$. Με άλλα λόγια, μέσω αυτής της συνάρτησης, η απόφαση παίρνεται σε συμφωνία με το τι επιθυμεί η πλειοψηφία των n πολιτών. Ένας άλλος τρόπος, λιγότερο δημοκρατικός, είναι μέσω της *δικτατορικής συνάρτησης* η οποία λαμβάνει υπόψιν της την άποψη μόνο ενός εκ των πολιτών, έστω του k -οστού πολίτη, και δίνει την τιμή $f(x) = x_k$ που επέλεξε αυτός.

Αλγόριθμοι και Θεωρία Γραφημάτων: Μεγάλο μέρος της επιστήμης της Πληροφορικής και των Αλγορίθμων ασχολείται με συγκεκριμένα προβλήματα στα οποία δίνεται κάποια διακριτή δομή και το ζητούμενο είναι να αποφανθούμε αν αυτή η δομή έχει κάποια συγκεκριμένη ιδιότητα. Για παράδειγμα, για δοθέν γράφημα με N το πλήθος κορυφές, πώς μπορούμε να αποφασίσουμε αν περιέχει Χαμιλτονιανό μονοπάτι; Αν είναι δένδρο; Αν μπορούμε να το χρωματίσουμε με 4 χρώματα; Αν είναι ισομορφικό με κάποιο άλλο γράφημα; Αν είναι επίπεδο; και πολλά άλλα. Ένα γράφημα με N το πλήθος κορυφές έχει το πολύ $n = \binom{N}{2}$ το πλήθος ακμές και περιγράφεται πλήρως αν γνωρίζουμε ποιες από αυτές τις ακμές εμφανίζονται στο γράφημα και ποιες όχι. Επομένως, μπορούμε να ταυτίσουμε το γράφημα με ένα διάνυσμα $x \in \{-1, 1\}^n$, όπου κάθε ψηφίο x_i αντιστοιχεί σε κάποια πιθανή ακμή: έχουμε $x_i = 1$ αν η ακμή είναι στο γράφημα και $x_i = -1$ αν δεν είναι. Έτσι, για κάθε «ιδιότητα γραφήματος» που μπορούμε να σκεφτούμε προκύπτει μια Boolean συνάρτηση.

Συνδυαστική: Θεωρούμε το σύνολο $S = [n] := \{1, 2, \dots, n\}$ το οποίο αποτελείται από του πρώτους n -το πλήθος φυσικούς αριθμούς. Τότε, μπορούμε να θέσουμε πλήθος ερωτημάτων για τον τρόπο με τον οποίο αλληλεπιδρούν τα διάφορα υποσύνολά του. Για παράδειγμα, ποια είναι η μεγαλύτερη οικογένεια υποσυνόλων του S με την ιδιότητα ότι κανένα από αυτά τα υποσύνολα δεν περιέχει κάποιο άλλο; Μπορούμε να σκεφτόμαστε κάθε οικογένεια $\mathcal{F}$ υποσυνόλων του S σαν μια Boolean συνάρτηση, με τον εξής τρόπο. Κάθε $x \in \{-1, 1\}^n$ αναπαριστά ένα υποσύνολο του S , όπου $x_i = 1$ αν ο i ανήκει στο υποσύνολο και $x_i = -1$ διαφορετικά. Με τη σειρά της, κάθε οικογένεια $\mathcal{F}$ υποσυνόλων του S ορίζει μια Boolean συνάρτηση $f_{\mathcal{F}}$: για κάθε $x \in \{-1, 1\}^n$ έχουμε ότι $f_{\mathcal{F}}(x) = 1$ αν το υποσύνολο που αναπαριστά το x ανήκει στην $\mathcal{F}$ και $f_{\mathcal{F}}(x) = -1$ αλλιώς.

1.2 Ισοπεριμετρικές ανισότητες στον διακριτό κύβο

Σε αυτή την εργασία δεν ενδιαφερόμαστε τόσο για κάποιες συγκεκριμένες Boolean συναρτήσεις και τον υπολογισμό τους, αλλά για τις γενικές ιδιότητες αυτής της κλάσης συναρτήσεων και τον τρόπο με τον οποίο αυτές οι ιδιότητες συνδέονται μεταξύ τους.

Θα μελετήσουμε ισοπεριμετρικές ανισότητες στον διακριτό κύβο. Κλασικά αποτελέσματα του είδους οφείλονται στον Margulis [42], τον Talagrand [48, 51, 49] και τους Kahn, Kalai και Linial [34]. Όλα αυτές οι ανισότητες αφορούν διάφορες έννοιες συνόρων και τις σχέσεις μεταξύ τους για Boolean συναρτήσεις.

Για μια συνάρτηση $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$ και ένα σημείο $x \in \{-1, 1\}^n$, ορίζουμε τη συνάρτηση ευαισθησίας της f στο x , και τη συμβολίζουμε με $h_f(x)$, να είναι το πλήθος των συντεταγμένων $i \in [n]$ για τις οποίες $f(\sigma_i(x)) \neq f(x)$, όπου $\sigma_i(x)$ είναι η κορυφή του κύβου που διαφέρει από το x μόνο στην i -οστή συντεταγμένη. Το σύνολο κορυφών V_f μιας συνάρτησης f ορίζεται να είναι το σύνολο όλων των x για τα οποία $h_f(x) > 0$, και το σύνολο ακμών E_f της f ορίζεται να είναι το σύνολο των ακμών $(x, \sigma_i(x))$ του διακριτού κύβου στα άκρα των οποίων η f παίρνει διαφορετικές τιμές. Το πιο βασικό ισοπεριμετρικό αποτέλεσμα για Boolean συναρτήσεις, η

ανισότητα Poincaré, ισχυρίζεται ότι

$$(1.2.1) \quad \frac{|E_f|}{2^n} \geq \text{Var}(f),$$

δηλαδή ότι το σύνολο ακμών μιας συνάρτησης f δεν μπορεί να είναι πολύ μικρό. Ο Margulis [42] βελτίωσε την ανισότητα Poincaré, δείχνοντας ότι για κάθε συνάρτηση f ισχύει ότι

$$(1.2.2) \quad \frac{|V_f|}{2^n} \frac{|E_f|}{2^n} \geq \Omega(\text{Var}(f)^2).$$

Ο Talagrand θεώρησε μια διαφορετική έννοια συνόρου για δοθείσα συνάρτηση f , την ποσότητα $\mathbb{E}[\sqrt{h_f}]$, την οποία θα ονομάζουμε σύνολο Talagrand της f . Απέδειξε διάφορα αποτελέσματα γι' αυτήν την ποσότητα [48, 51], το πιο βασικό από τα οποία ισχυρίζεται ότι

$$(1.2.3) \quad \mathbb{E}[\sqrt{h_f}] \geq \Omega(\text{Var}(f)).$$

Αυτή η ανισότητα ισχυροποιεί το θεώρημα του Margulis, όπως φαίνεται με απλή εφαρμογή της ανισότητας Cauchy-Schwarz. Στη συνέχεια, ο Talagrand ισχυροποίησε αυτό το αποτέλεσμα για συναρτήσεις με μικρή διασπορά, αποδεικνύοντας το ακόλουθο θεώρημα.

Θεώρημα 1.2.1 (Talagrand [48]). *Υπάρχει απόλυτη σταθερά $c > 0$ τέτοια ώστε, για κάθε συνάρτηση $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$ ισχύει ότι*

$$(1.2.4) \quad \mathbb{E}[\sqrt{h_f}] \geq c \text{Var}(f) \sqrt{\log\left(\frac{1}{\text{Var}(f)}\right)}.$$

Παρόλο που αυτή η ανισότητα είναι ακριβής έως απόλυτες σταθερές (κάτι που μπορούμε να δούμε θεωρώντας υπο-κύβους), αφήνει κάποια εκκρεμότητα καθώς είναι μη συγκρίσιμη με ένα άλλο σημαντικό ισοπεριμετρικό αποτέλεσμα, το θεώρημα (KKL) των Kahn, Kalai και Linial [34].

Το θεώρημα (KKL) αφορά μια άλλη ισοπεριμετρική ποσότητα που σχετίζεται με δοθείσα συνάρτηση f , τη λεγόμενη επιρροή. Η επιρροή μιας συντεταγμένης $i \in [n]$, την οποία συμβολίζουμε με $I_i(f)$, είναι το πλήθος των ακμών στο E_f που βρίσκονται στην κατεύθυνση i , διαιρεμένο με 2^n . Δηλαδή,

$$(1.2.5) \quad I_i(f) = \mathbb{P}[f(x) \neq f(\sigma_i(x))].$$

Το θεώρημα (KKL) [34] ισχυρίζεται ότι κάθε συνάρτηση πρέπει να έχει τουλάχιστον μία μεταβλητή που έχει μη αμελητέα επιρροή. Συγκεκριμένα, οι Kahn, Kalai και Linial απέδειξαν ότι

$$(1.2.6) \quad \max_i I_i(f) \geq \Omega\left(\frac{\log n}{n} \text{Var}(f)\right).$$

Καθώς τα δύο ισοπεριμετρικά θεωρήματα, αυτό του Talagrand και αυτό των Kahn, Ka-

laï και Linial, μοιάζουν πολύ διαφορετικά (ομοίως και οι τεχνικές που υπεισέρχονται στις αποδείξεις τους), ο Talagrand ρώτησε αν υπάρχει κάποια μοναδική ισοπεριμετρική ανισότητα η οποία να είναι αρκετά ισχυρή ώστε να συνεπάγεται και τα δύο αυτά αποτελέσματα ταυτόχρονα. Στο [51], ο Talagrand έκανε μια εικασία για τη μορφή μιας τέτοιας ανισότητας, και σημείωσε κάποια πρόοδο προς την απόδειξή της. Συγκεκριμένα, απέδειξε ότι υπάρχει $\alpha \in (0, 1/2]$ ώστε για κάθε συνάρτηση $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$ να ισχύει ότι

$$(1.2.7) \quad \mathbb{E} \left[\sqrt{h_f} \right] \geq c \text{Var}(f) \left(\log \left(\frac{1}{\text{Var}(f)} \right) \right)^{1/2-\alpha} \left(\log \left(1 + 1 / \left(\sum_{i=1}^n I_i(f)^2 \right) \right) \right)^\alpha,$$

για κάποια απόλυτη σταθερά $c > 0$. Αυτή η ανισότητα συνεπάγεται ένα ασθενέστερο αποτέλεσμα τύπου (KKL): Υπάρχει $\beta > 0$ τέτοιος ώστε αν η f έχει μέση τιμή 0 να ισχύει ότι

$$(1.2.8) \quad \max_i I_i(f) \geq \Omega((\log n)^\beta / n).$$

Όμως, από την (1.2.7) δεν έπεται το πλήρες θεώρημα (KKL). Ο Talagrand έκανε την εικασία ότι η (1.2.7) ισχύει και για $\alpha = 1/2$, μια πρόταση που είναι αρκετά ισχυρή για να δώσει ως άμεσο πόρισμα το θεώρημα (KKL).

Οι Eldan και Gross [19] απέδειξαν ότι η εικασία του Talagrand ισχύει, δηλαδή μπορούμε να πάρουμε $\alpha = 1/2$ στην (1.2.7).

Θεώρημα 1.2.2. Υπάρχει απόλυτη σταθερά $c > 0$ τέτοια ώστε, για κάθε συνάρτηση $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$ να ισχύει ότι

$$(1.2.9) \quad \mathbb{E} \left[\sqrt{h_f} \right] \geq c \text{Var}(f) \sqrt{\log \left(1 + 1 / \left(\sum_{i=1}^n I_i(f)^2 \right) \right)}.$$

Παρουσιάζουμε την απόδειξη του Θεωρήματος 1.2.2 στα Κεφάλαια 3, 4 και 5. Η μέθοδος που εισάγεται από τους Eldan και Gross βασίζεται στον στοχαστικό λογισμό και αποφεύγει τη χρήση της υπερσυσταλτότητας, η οποία έπαιξε ουσιαστικό ρόλο σε όλα τα προηγούμενα αποτελέσματα. Μια συνοπτική περιγραφή της μεθόδου δίνεται στο Κεφάλαιο 3.

Παρουσιάζουμε επίσης άλλα αποτελέσματα από το [19], τα οποία αποδεικνύονται με την ίδια τεχνική και εξασφαλίζουν ότι οι προσεγγιστικά-ακραίες συναρτήσεις για διάφορες κλασικές ανισότητες που αφορούν τις επιρροές Boolean συναρτήσεων πρέπει να έχουν μεγάλο σύνολο κορυφών. Για παράδειγμα, δίνουμε κάτω φράγμα για το μέγεθος του συνόλου κορυφών μιας συνάρτησης που κάνει την

$$\text{Var}(f) \leq C \sum_{i=1}^n \frac{I_i(f)}{1 + \log(1/I_i(f))}$$

σχεδόν ισότητα. Πόρισμα αυτού του αποτελέσματος είναι το γεγονός ότι αν για κάποιο σύνολο η ανισότητα Kahn-Kalai-Linial ισχύει περίπου ως ισότητα τότε μεγάλο ποσοστό του συνόλου

περιέχεται στο εσωτερικό σύνολο κορυφών του. Περιγράφουμε επίσης μια βελτιωμένη εκδοχή μιας ανισότητας των Keller και Kindler που συνδέει ποσοτικά τις επιρροές με την ευστάθεια θορύβου.

Στο Κεφάλαιο 6 παρουσιάζουμε τη μεταγενέστερη εργασία [21] των Eldan, Kindler, Lifshitz και Minzer, στην οποία μεταξύ άλλων δίνεται μια εναλλακτική απόδειξη του Θεωρήματος 1.2.2. Για κάθε $f : \{-1, 1\}^n \rightarrow \mathbb{R}$ και κάθε $x = (x_1, \dots, x_n) \in \{-1, 1\}^n$ θέτουμε

$$D_j f(x) = \frac{f(x_1, \dots, x_j, \dots, x_n) - f(x_1, \dots, -x_j, \dots, x_n)}{2}.$$

Το βασικό τεχνικό αποτέλεσμα είναι το εξής θεώρημα, το οποίο αποδεικνύεται με τη «μέθοδο των τυχαίων περιορισμών».

Θεώρημα 1.2.3. *Για κάθε $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$ και κάθε $p \in [1, 2]$ ισχύει ότι*

$$(1.2.10) \quad \|\nabla f\|_p \gtrsim \sup_{d \geq 0} \left(\sum_{|S| \geq d} |\widehat{f}(S)|^2 \right)^{1/p} \sqrt{d},$$

όπου $\widehat{f}(S)$, $S \subseteq [n]$ είναι οι συντελεστές Walsh της f .

Επιλέγοντας $d \approx \log\left(\frac{e}{\text{Var}(f)}\right)$, από την ανισότητα (1.2.10) παίρνουμε την ακριβή ισοπεριμετρική ανισότητα του Talagrand

$$(1.2.11) \quad \|\nabla f\|_p \gtrsim (\text{Var}(f))^{1/p} \sqrt{\log\left(\frac{e}{\text{Var}(f)}\right)}$$

για κάθε $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$, και επιλέγοντας $d \approx \log\left(\frac{e}{W(f)}\right)$, όπου

$$W(f) = \sum_{j=1}^n \left(\mathbb{E} [|D_j f|] \right)^2,$$

από την ανισότητα (1.2.10) παίρνουμε την εικασία του Talagrand

$$(1.2.12) \quad \|\nabla f\|_p \gtrsim (\text{Var}(f))^{1/p} \sqrt{\log\left(\frac{e}{W(f)}\right)}$$

για κάθε $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$, που είναι ισοδύναμη με το Θεώρημα 1.2.2.

Στο Κεφάλαιο 7 παρουσιάζουμε την πρόσφατη εργασία [4] των Beltran, Ivanisvili και Madrid στην οποία αποδεικνύονται ακριβείς ισοπεριμετρικές ανισότητες στον διακριτό κύβο. Ένα παράδειγμα είναι η ανισότητα

$$\mathbb{E} \left[h_A^{\log_2(3/2)} \right] \geq \mu_n(A)^* (\log_2(1/\mu_n(A)^*))^{\log_2(3/2)}$$

για κάθε $A \subseteq \{-1, 1\}^n$, όπου $\mu_n(A)^* = \min\{\mu_n(A), 1 - \mu_n(A)\}$ και η συνάρτηση h_A έχει φορέα το A και σε κάθε x αντιστοιχίζει το πλήθος των γειτονικών κορυφών του x που ανήκουν στο

συμπλήρωμα του A . Δίνονται επίσης κάτω φράγματα για την $\mathbb{E}[h_A^\beta]$ συναρτήσει του $\mu_n(A)$ για κάθε $\beta \in [1/2, 1]$, τα οποία βελτιώνουν παλιότερα γνωστά αποτελέσματα και σε κάποιες περιπτώσεις είναι βέλτιστα. Ένα παράδειγμα είναι η ανισότητα $\mathbb{E}[h_A^{0.53}] \geq 2\mu_n(A)(1 - \mu_n(A))$ που ισχύει για κάθε σύνολο A με $\mu_n(A) \geq 1/2$, η οποία ισχυροποιεί μια ανισότητα των Kahn και Park.

Στην ίδια εργασία δίνεται επίσης νέα απόδειξη της (1.2.10), η οποία σε σύγκριση με τη μέθοδο των τυχαίων περιορισμών του [21] έχει το πλεονέκτημα ότι επεκτείνει τα αποτελέσματα σε συναρτήσεις με τιμές σε τυχόντα χώρο Banach που έχει πεπερασμένο συν-τύπο. Έστω $\Delta = \sum_{j=1}^n D_j$, και έστω $e^{-t\Delta}$ η ημιομάδα θερμότητας.

Θεώρημα 1.2.4. Έστω X χώρος Banach με σταθερά συν-τύπου $C_q(X) < \infty$ για κάποιον $q \geq 2$. Για κάθε $p \in [1, 2]$ και για κάθε $f : \{-1, 1\}^n \rightarrow X$ έχουμε ότι

$$(1.2.13) \quad \|f - e^{-t\Delta} f\|_p \lesssim C_q(X) q^{3/2} (1 - e^{-2t})^{\frac{1}{q}} \|Df\|_p, \quad t \geq 0$$

όπου $\|Df\|_p^p = \mathbb{E} \left[\left\| \sum_{1 \leq j \leq n} x'_j D_j f(x) \right\|^p \right]$ και x' είναι ένα ανεξάρτητο αντίγραφο του x .

Αν η f είναι Boolean συνάρτηση, τότε το Θεώρημα 1.2.4 μαζί με την απλή εκτίμηση

$$\|f - e^{-\Delta/d} f\|_p \gtrsim \left(\sum_{|S| \geq d} \widehat{f}(S)^2 \right)^{1/p}, \quad p \in [1, 2]$$

συνεπάγεται την (1.2.10), απ' όπου έπεται το Θεώρημα 1.2.2.

ΚΕΦΑΛΑΙΟ 2

Εισαγωγή

2.1 Συναρτήσεις στον διακριτό κύβο και η ανισότητα Poincaré

Χρησιμοποιούμε τον όρο Boolean συνάρτηση για κάθε συνάρτηση $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$ με πεδίο ορισμού τον διακριτό κύβο $\{-1, 1\}^n$ που παίρνει τις τιμές ± 1 . Για κάθε Boolean συνάρτηση ορίζονται οι «επιρροές» της που κατά κάποιον τρόπο δίνουν ένα «μέτρο σημαντικότητας» κάθε συντεταγμένης x_i για τον υπολογισμό της τιμής $f(x)$.

Ορισμός 2.1.1 (επιρροές μιας Boolean συνάρτησης). Έστω $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$ μια Boolean συνάρτηση. Ορίζουμε την *i-οστή επιρροή* της f , την οποία συμβολίζουμε με $I_i(f)$, ως εξής:

$$I_i(f) = \mu_n(\{x \in \{-1, 1\}^n : f(x) \neq f(\sigma_i(x))\}),$$

όπου με μ_n συμβολίζουμε το ομοιόμορφο μέτρο πιθανότητας στον διακριτό κύβο το οποίο δίνει μάζα $\frac{1}{2^n}$ σε κάθε $x \in \{-1, 1\}^n$, και με $\sigma_i(x)$ συμβολίζουμε το διάνυσμα που προκύπτει από το x αν αλλάξουμε το πρόσημο της *i*-οστής συντεταγμένης. Με άλλα λόγια, η *i*-οστή επιρροή της f δείχνει κατά πόσον η αλλαγή της *i*-οστής συντεταγμένης του $x \in \{-1, 1\}^n$ επηρεάζει την τιμή της $f(x)$.

Μία από τις κύριες ποσότητες οι οποίες θα εμφανιστούν στις διάφορες ανισότητες που θα μελετήσουμε είναι η ολική επιρροή μιας Boolean συνάρτησης.

Ορισμός 2.1.2 (ολική επιρροή μιας Boolean συνάρτησης). Έστω $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$ μια Boolean συνάρτηση. Ορίζουμε ως *ολική επιρροή* της f , και την συμβολίζουμε με $I(f)$, την ποσότητα

$$I(f) := \sum_{i=1}^n I_i(f).$$

Παράδειγμα 2.1.3. Αν θεωρήσουμε τη συνάρτηση $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$ που λαμβάνει υπόψην της μόνο την απόφαση του *j*-παίκτη, δηλαδή τη συνάρτηση

$$f(x) = x_j, \quad x \in \{-1, 1\}^n$$

όπου $j \in [n] := \{1, \dots, n\}$, τότε παρατηρούμε ότι ισχύουν τα εξής:

- Η επιρροή του j -παίκτη καθορίζει πλήρως την τιμή της f , δηλαδή

$$I_j(f) = 1,$$

κάτι που συμβαδίζει με τον διαισθητικό ορισμό που δώσαμε για τις επιρροές της f .

- Η επιρροή κάθε άλλου παίκτη $i \neq j$ είναι μηδενική αφού η f καθορίζεται πλήρως μόνο μέσω της απόφασης του j -παίκτη, δηλαδή

$$I_i(f) = 0.$$

Μια άλλη βασική έννοια που ορίζεται για κάθε Boolean συνάρτηση είναι η μέση τιμή της, η οποία σε αυτό το πλαίσιο συγκρίνει το πλήθος των φορών που εμφανίζονται οι αριθμοί 1 και -1 ως τιμές της f . Γενικότερα, δίνουμε τον ακόλουθο ορισμό.

Ορισμός 2.1.4 (μέση τιμή). Έστω $f : \{-1, 1\}^n \rightarrow \mathbb{R}$. Ορίζουμε τη «μέση τιμή» της f , την οποία συμβολίζουμε με $\mathbb{E}[f]$, ως εξής:

$$\mathbb{E}[f] = \int_{\{-1, 1\}^n} f(x) d\mu_n(x) = \frac{1}{2^n} \sum_{x \in \{-1, 1\}^n} f(x).$$

Παρατήρηση 2.1.5. Στην περίπτωση που η f είναι Boolean συνάρτηση, δηλαδή παίρνει μόνο τις τιμές ± 1 , είναι φανερό ότι

$$-1 \leq \mathbb{E}[f] \leq 1.$$

Δηλαδή, η μέση τιμή της f είναι ένας αριθμός του διαστήματος $[-1, 1]$ που εκφράζει διαισθητικά τα παρακάτω:

- Αν $\mathbb{E}[f] > 0$, τότε το πλήθος των εμφανίσεων του 1 ως τιμής της f είναι μεγαλύτερο από το πλήθος των εμφανίσεων του -1 ως τιμής της f .
- Αν $\mathbb{E}[f] < 0$, τότε το πλήθος των εμφανίσεων του -1 ως τιμής της f είναι μεγαλύτερο από το πλήθος των εμφανίσεων του 1 ως τιμής της f .
- Αν $\mathbb{E}[f] = 0$, τότε το πλήθος των εμφανίσεων της τιμής -1 είναι ίσο με αυτό των εμφανίσεων της τιμής 1 και τότε η f καλείται «ισορροπημένη».

Είναι εύκολο να αποδειχθεί ότι ο χώρος $\mathcal{H}_n$ των συναρτήσεων $f : \{-1, 1\}^n \rightarrow \mathbb{R}$ είναι γραμμικός χώρος με τις κατά σημείο πράξεις, τον οποίο μπορούμε να εφοδιάσουμε με μια συνάρτηση εσωτερικού γινομένου:

Ορισμός 2.1.6. Στον χώρο H_n ορίζουμε το εσωτερικό γινόμενο $\langle \cdot, \cdot \rangle$, όπου

$$\langle f, g \rangle := \mathbb{E}[fg] = \frac{1}{2^n} \sum_{x \in \{-1, 1\}^n} f(x)g(x)$$

για κάθε $f, g : \{-1, 1\}^n \rightarrow \mathbb{R}$. Αποδεικνύεται εύκολα ότι αυτή η συνάρτηση ικανοποιεί τις ιδιότητες του εσωτερικού γινομένου.

Το παραπάνω εσωτερικό γινόμενο επάγει μια νόρμα στον $\mathcal{H}_n$, την οποία συμβολίζουμε με $\|\cdot\|_2$: για κάθε $f : \{-1, 1\}^n \rightarrow \mathbb{R}$ ορίζουμε

$$\|f\|_2 := \sqrt{\langle f, f \rangle} = \sqrt{\mathbb{E}[f^2]}.$$

Ορισμός 2.1.7 (συναρτήσεις Walsh). Για κάθε $S \subseteq [n]$ ορίζουμε την Boolean συνάρτηση $w_S : \{-1, 1\}^n \rightarrow \{-1, 1\}$ ως εξής:

$$(2.1.1) \quad w_S(x) := \begin{cases} \prod_{i \in S} x_i, & \text{αν } S \neq \emptyset \\ 1, & \text{αν } S = \emptyset \end{cases}$$

για κάθε $x \in \{-1, 1\}^n$. Οι συναρτήσεις w_S καλούνται συναρτήσεις Walsh και, όπως θα δούμε παρακάτω, αποτελούν ορθοκανονική βάση του $\mathcal{H}_n$.

Παρατηρήσεις 2.1.8. Για κάθε $i \in [n]$ θέτουμε

$$w_i = w_{\{i\}}.$$

Παρατηρούμε ότι για κάθε $x \in \{-1, 1\}^n$ ισχύει ότι

$$w_i(x) = x_i = \begin{cases} 1, & \text{με πιθανότητα } 1/2 \\ -1, & \text{με πιθανότητα } 1/2 \end{cases}.$$

Ελέγχουμε εύκολα ότι οι w_i είναι ανεξάρτητες συμμετρικές τυχαίες μεταβλητές Bernoulli(1/2), δηλαδή η ακολουθία των $\{w_1, w_2, \dots, w_n\}$ είναι μια πεπερασμένη ακολουθία ανεξάρτητων τυχαίων μεταβλητών Rademacher.

Επίσης, από τον ορισμό βλέπουμε ότι

$$\mathbb{E}[w_i] = 0$$

και ότι για κάθε $\emptyset \neq S \subseteq [n]$ ισχύει ότι

$$w_S = \prod_{i \in S} w_i.$$

Από την ανεξαρτησία των w_i έπεται ότι

$$\mathbb{E}[w_S] = \prod_{i \in S} \mathbb{E}[w_i] = 0.$$

Για $S = \emptyset$ έχουμε $w_\emptyset \equiv 1$, άρα

$$\mathbb{E}[w_\emptyset] = 1.$$

Προχωρώντας λίγο περισσότερο, παρατηρούμε ότι για κάθε $S, T \subseteq [n]$ ισχύει ότι

$$w_S w_T = w_{S \Delta T}$$

όπου $S \Delta T = (S \setminus T) \cup (T \setminus S)$ είναι η συμμετρική διαφορά των S και T , και άρα

$$\langle w_S, w_T \rangle = \mathbb{E}[w_S w_T] = \mathbb{E}[w_{S \Delta T}] = \begin{cases} 1, & \text{αν } S = T \\ 0, & \text{αν } S \neq T \end{cases}.$$

Αυτό δείχνει ότι το σύνολο $\{w_S : S \subseteq [n]\}$ αποτελεί ένα ορθοκανονικό σύνολο του χώρου $\mathcal{H}_n$ με 2^n στοιχεία.

Παρατηρούμε επίσης ότι, για κάθε $x, y \in \{-1, 1\}^n$,

$$\mathbb{1}_{\{y\}}(x) = \prod_{i=1}^n \frac{1 + x_i y_i}{2} = \frac{1}{2^n} \prod_{i=1}^n (1 + x_i y_i) = \frac{1}{2^n} \sum_{S \subseteq [n]} w_S(x) w_S(y).$$

Συνεπώς, αν θεωρήσουμε τυχούσα συνάρτηση $f : \{-1, 1\}^n \rightarrow \mathbb{R}$, μπορούμε να γράψουμε

$$\begin{aligned} f(x) &= \sum_{y \in \{-1, 1\}^n} f(y) \mathbb{1}_{\{y\}}(x) = \sum_{S \subseteq [n]} \left(\frac{1}{2^n} \sum_{y \in \{-1, 1\}^n} f(y) w_S(y) \right) w_S(x) \\ &= \sum_{S \subseteq [n]} \langle f, w_S \rangle w_S(x) = \sum_{S \subseteq [n]} \widehat{f}(S) w_S(x) \end{aligned}$$

όπου

$$\widehat{f}(S) := \langle f, w_S \rangle$$

είναι οι συντελεστές Fourier-Walsh της f .

Αποδείξαμε επομένως ότι κάθε $f \in \mathcal{H}_n$ γράφεται ως

$$f = \sum_{S \subseteq [n]} \widehat{f}(S) w_S$$

και άρα το σύνολο $\{w_S : S \subseteq [n]\}$ αποτελεί ορθοκανονική βάση του χώρου $\mathcal{H}_n$. Συνοψίζουμε στο επόμενο θεώρημα.

Θεώρημα 2.1.9. *Ο χώρος $\mathcal{H}_n$ έχει ορθοκανονική βάση το σύνολο των συναρτήσεων Walsh $\{w_S : S \subseteq [n]\}$ και μάλιστα για κάθε $f \in \mathcal{H}_n$ ισχύει ότι*

$$f = \sum_{S \subseteq [n]} \widehat{f}(S) w_S$$

όπου

$$\widehat{f}(S) = \langle f, w_S \rangle = \mathbb{E}[f w_S]$$

για κάθε $S \subseteq [n]$ είναι οι συντελεστές Fourier της f . Ειδικότερα,

$$\widehat{f}(\emptyset) = \langle f, w_\emptyset \rangle = \mathbb{E}[f].$$

Πόρισμα 2.1.10 (ταυτότητα Parseval). Για κάθε $f \in \mathcal{H}_n$ ισχύει η ισότητα

$$\|f\|_2^2 = \mathbb{E}[f^2] = \sum_{S \subseteq [n]} (\widehat{f}(S))^2.$$

Μία ακόμα χρήσιμη ποσότητα που θα παίξει μάλιστα καθοριστικό ρόλο παρακάτω, αφού θα μελετήσουμε ανισότητες που τη συνδέουν με τις επιρροές μιας Boolean συνάρτησης, είναι η διασπορά της f :

Ορισμός 2.1.11 (διασπορά Boolean συνάρτησης). Έστω $f : \{-1, 1\}^n \rightarrow \mathbb{R}$. Ορίζουμε την «διασπορά» της f , την οποία θα συμβολίζουμε με $\text{Var}[f]$, ως την ποσότητα:

$$\text{Var}(f) = \mathbb{E}(f - \mathbb{E}[f])^2.$$

Χρησιμοποιώντας την ταυτότητα Parseval παίρνουμε την απλοποιημένη μορφή

$$\text{Var}(f) = \mathbb{E}[f^2] - (\mathbb{E}[f])^2 = \sum_{S \subseteq [n]} \widehat{f}(S)^2 - \widehat{f}(\emptyset)^2 = \sum_{\emptyset \neq S \subseteq [n]} \widehat{f}(S)^2.$$

Για κάθε $f : \{-1, 1\}^n \rightarrow \mathbb{R}$ ορίζουμε επίσης την έννοια της διακριτής μερικής παραγώγου.

Ορισμός 2.1.12 (διακριτή μερική παράγωγος). Έστω $f : \{-1, 1\}^n \rightarrow \mathbb{R}$. Ορίζουμε την i -οστή διακριτή μερική παράγωγο της f , την οποία συμβολίζουμε με $\partial_i f$, ως εξής:

$$\partial_i f(x) := \frac{f(x^{i \rightarrow 1}) - f(x^{i \rightarrow -1})}{2}$$

για κάθε $x \in \{-1, 1\}^n$, όπου το $x^{i \rightarrow a}$ έχει τον a ως i -οστή συντεταγμένη και συμπίπτει με το x σε όλες τις άλλες συντεταγμένες.

Παρατηρήσεις 2.1.13. Έστω $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$ μια Boolean συνάρτηση.

(α) Παρατηρούμε ότι η μερική παράγωγος $\partial_i f(x)$ έχει τρεις το πλήθος δυνατές τιμές: 1 ή -1 αν η αλλαγή της i -οστής συντεταγμένης επηρεάζει την τιμή της $f(x)$ και 0 αν η αλλαγή της i -οστής συντεταγμένης του x δεν επηρεάζει την τιμή της $f(x)$. Με άλλα λόγια, για κάθε $x \in \{-1, 1\}^n$ ισχύει ότι

$$\partial_i f(x) \in \{-1, 0, 1\}.$$

Άρα, για κάθε $x \in \{-1, 1\}^n$ το τετράγωνο $(\partial_i f(x))^2$ της $\partial_i f(x)$ μπορεί να πάρει δύο τιμές, 1 και 0, δηλαδή

$$\partial_i f(x)^2 \in \{0, 1\}$$

και μάλιστα ισχύει ότι

$$(2.1.2) \quad \|\partial_i f\|_2^2 = \mathbb{E} [\partial_i f(x)^2] = I_i(f)$$

για κάθε $i \in [n]$. Αυτό ισχύει διότι $(\partial_i f(x))^2 = 1$ αν η αλλαγή της i -οστής συντεταγμένης επηρεάζει την τιμή της $f(x)$ και 0 αλλιώς.

(β) Για κάθε $f, g : \{-1, 1\}^n \rightarrow \mathbb{R}$ και για κάθε $a, b \in \mathbb{R}$ ισχύει ότι

$$\partial_i(af + bg) = a\partial_i f + b\partial_i g, \quad i \in [n].$$

(γ) Για κάθε $S \subseteq [n]$, θεωρώντας την συνάρτηση Walsh w_S , έχουμε ότι

$$\partial_i w_S = \begin{cases} w_{S \setminus \{i\}}, & \text{αν } i \in S \\ 0, & \text{αν } i \notin S \end{cases}.$$

(δ) Χρησιμοποιώντας τις δύο παραπάνω παρατηρήσεις έχουμε ότι αν η $f : \{-1, 1\}^n \rightarrow \mathbb{R}$ έχει ανάλυση Fourier

$$f = \sum_{S \subseteq [n]} \widehat{f}(S) w_S$$

τότε

$$\partial_i f = \sum_{S \subseteq [n]} \widehat{f}(S) \partial_i w_S = \sum_{\{S \subseteq [n] : i \in S\}} \widehat{f}(S) w_{S \setminus \{i\}}$$

είναι η ανάλυση Fourier της $\partial_i f$.

Ορισμός 2.1.14 (μονότονη συνάρτηση). Μια Boolean συνάρτηση $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$ καλείται μονότονη αν $f(x) \leq f(y)$ για κάθε $x, y \in \{-1, 1\}^n$ με $x_i \leq y_i$ για κάθε $i \in [n]$.

Παρατήρηση 2.1.15. Αν $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$ είναι μια Boolean συνάρτηση που είναι μονότονη τότε, από τον ορισμό της μερικής παραγώγου $\partial_i f$ έχουμε ότι για κάθε $x \in \{-1, 1\}^n$ η διαφορά

$$f(x^{i \rightarrow 1}) - f(x^{i \rightarrow -1}) \in \{0, 2\}$$

άρα $\partial_i f(x) \in \{0, 1\}$, και έπεται ότι

$$(2.1.3) \quad I_i(f) = \mathbb{E} [\partial_i f(x)] = \widehat{f}(i)$$

για κάθε $i \in [n]$, αφού $\partial_i f(x) = 0$ για εκείνα ακριβώς τα $x \in \{-1, 1\}^n$ για τα οποία δεν η τιμή της $f(x)$ δεν επηρεάζεται με την αλλαγή της i -οστής συντεταγμένης.

Ορισμός 2.1.16 (κλίση). Έστω $f : \{-1, 1\}^n \rightarrow \mathbb{R}$ και $\partial_i f$ οι μερικές παράγωγοι της f . Ορίζουμε την κλίση της f , την οποία θα συμβολίζουμε με ∇f , ως εξής:

$$\nabla f = (\partial_1 f, \partial_2 f, \dots, \partial_n f) : \{-1, 1\}^n \rightarrow \mathbb{R}^n.$$

Θεώρημα 2.1.17 (ανισότητα Poincaré). Έστω $f : \{-1, 1\}^n \rightarrow \mathbb{R}$. Τότε, ισχύει ότι

$$\text{Var}(f) \leq \int_{E_2^n} \|\nabla f\|_2^2 d\mu_n = \sum_{i=1}^n I_i(f).$$

Απόδειξη. Παρατηρούμε αρχικά ότι για την ανάλυση Fourier της $\partial_i f$ έχουμε ότι

$$\partial_i f(x) = \sum_{\{S \subseteq [n] : i \in S\}} \widehat{f}(S) w_{S \setminus \{i\}}(x)$$

και άρα, λόγω της ταυτότητας Parseval,

$$\|\partial_i f\|_2^2 = \mathbb{E}[(\partial_i f(x))^2] = \sum_{\{S \subseteq [n] : i \in S\}} \widehat{f}(S)^2.$$

Λόγω της προηγούμενης παρατήρησης, έχουμε ότι

$$\sum_{i=1}^n I_i(f) = \sum_{i=1}^n \mathbb{E}[(\partial_i f(x))^2] = \sum_{i=1}^n \sum_{\{S \subseteq [n] : i \in S\}} \widehat{f}(S)^2 = \sum_{S \neq \emptyset} |S| \widehat{f}(S)^2$$

και, αφού $|S| \geq 1$ για κάθε $S \neq \emptyset$, έπεται ότι

$$\sum_{i=1}^n I_i(f) = \sum_{S \neq \emptyset} |S| \widehat{f}(S)^2 \geq \sum_{S \neq \emptyset} \widehat{f}(S)^2 = \text{Var}(f)$$

δηλαδή το ζητούμενο. □

2.2 Υπερσυσταλτότητα: η ανισότητα της Bonami

Σε αυτήν την ενότητα αποδεικνύουμε την ανισότητα υπερσυσταλτότητας για συναρτήσεις ορισμένες στον διακριτό κύβο. Για κάθε $f : \{-1, 1\}^n \rightarrow \mathbb{R}$ θεωρούμε τη διακριτή Λαπλασιανή $L(f)$ της f , η οποία ορίζεται από την

$$L(f)(x) = \frac{1}{2} \sum_{i=1}^n (f(\sigma_i(x)) - f(x)).$$

Παρατηρούμε ότι $w_S(\sigma_i(x)) = w_S(x)$ αν $i \notin S$ και $w_S(y) = -w_S(x)$ αν $i \in S$. Έπεται ότι

$$L(w_S) = -|S|w_S,$$

δηλαδή οι συναρτήσεις Walsh είναι ιδιοσυναρτήσεις της διακριτής Λαπλασιανής. Για κάθε $t \geq 0$ θεωρούμε τον τελεστή $P_t : C(\{-1, 1\}^n) \rightarrow C(\{-1, 1\}^n)$ με

$$P_t(f) = (e^{tL})(f) = \sum_{k=0}^{\infty} \frac{t^k L^k(f)}{k!},$$

όπου $L^k = L \circ \dots \circ L$ (k φορές). Παρατηρούμε ότι, για κάθε $S \subseteq \{1, \dots, n\}$,

$$L^k(w_S) = (-1)^k |S|^k w_S$$

(αυτό προκύπτει επαγωγικά από την $L(w_S) = -|S|w_S$). Συνεπώς,

$$P_t(w_S) = \sum_{k=0}^{\infty} \frac{(-1)^k t^k |S|^k w_S}{k!} = (e^{-t|S|})(w_S).$$

Έπεται ότι

$$P_t f = \sum_{S \subseteq [n]} \widehat{f}(S) e^{-t|S|} w_S.$$

Θεώρημα 2.2.1 (Bonami). Έστω $1 < p < \infty$ και $t \geq 0$. Θέτουμε $q(t) = 1 + (p-1)e^{2t}$. Τότε, για κάθε $f : \{-1, 1\}^n \rightarrow \mathbb{R}$ ισχύει ότι

$$\|P_t(f)\|_{q(t)} \leq \|f\|_p.$$

Για κάθε $1 < p < \infty$ ορίζουμε $r_p = \frac{1}{\sqrt{p-1}}$. Η απόδειξη του Θεωρήματος 2.2.1 βασίζεται στο ακόλουθο θεώρημα.

Θεώρημα 2.2.2. Έστω $1 < p < q < \infty$ και έστω $\{x_S : S \subseteq [n]\}$ μια οικογένεια διανυσμάτων σε έναν χώρο με νόρμα $(X, \|\cdot\|)$. Τότε,

$$\left\| \sum_S r_q^{|S|} w_S x_S \right\|_{L^q(X)} \leq \left\| \sum_S r_p^{|S|} w_S x_S \right\|_{L^p(X)},$$

όπου w_S είναι οι συναρτήσεις Walsh.

Έχοντας το Θεώρημα 2.2.2 στη διάθεσή μας αποδεικνύουμε το Θεώρημα 2.2.1.

Απόδειξη του Θεωρήματος 2.2.1. Παρατηρούμε ότι αν $f = \sum_S \widehat{f}(S) w_S$, τότε

$$P_t f = \sum_S \widehat{f}(S) e^{-t|S|} w_S = \sum_S r_p^{|S|} e^{-t|S|} w_S \frac{\widehat{f}(S)}{r_p^{|S|}} = \sum_S r_{q(t)}^{|S|} e^{-t|S|} w_S \frac{\widehat{f}(S)}{r_p^{|S|}},$$

αφού $r_{q(t)} = e^{-t|S|} r_p^{|S|}$. Από το Θεώρημα 2.2.2 έχουμε ότι

$$\|P_t f\|_{q(t)} = \left\| \sum_S r_{q(t)}^{|S|} w_S \frac{\widehat{f}(S)}{r_p^{|S|}} \right\|_{q(t)} \leq \left\| \sum_S r_p^{|S|} w_S \frac{\widehat{f}(S)}{r_p^{|S|}} \right\|_p = \|f\|_p$$

όπως θέλαμε. □

Περνάμε τώρα στην απόδειξη του Θεωρήματος 2.2.2. Το αρχικό βήμα είναι μια «ανισότητα δύο σημείων».

Θεώρημα 2.2.3. Για κάθε $x, y \in \mathbb{R}$ και $1 < p < \infty$ ορίζουμε

$$F_p(x, y) = \left(\frac{1}{2}|x + r_p y|^p + \frac{1}{2}|x - r_p y|^p \right)^{1/p},$$

όπου

$$r_p = \frac{1}{\sqrt{p-1}}.$$

Τότε, η $F_p(x, y)$ είναι φθίνουσα συνάρτηση του p στο $(1, \infty)$.

Απόδειξη. Έστω $x, y \in \mathbb{R}$ και $1 < p < q < \infty$. Η ανισότητα $F_q(x, y) \leq F_p(x, y)$ είναι ομογενής, οπότε μπορούμε να υποθέσουμε ότι $x = 1$. Διακρίνουμε τρεις περιπτώσεις:

(α) Υποθέτουμε πρώτα ότι $1 < p < q \leq 2$ και ότι $0 \leq |r_p y| \leq 1$. Χρησιμοποιώντας το διωνυμικό θεώρημα βλέπουμε ότι

$$\frac{1}{2}|1 + r_q y|^q + \frac{1}{2}|1 - r_q y|^q = 1 + \sum_{k=1}^{\infty} \binom{q}{2k} \left(\frac{y^2}{q-1} \right)^k.$$

Χρησιμοποιώντας την ανισότητα $(1+x)^\alpha \leq 1 + \alpha x$ (η οποία ισχύει για $0 < \alpha \leq 1$ και $x \geq 0$) με $\alpha = p/q$, παίρνουμε

$$\left(1 + \sum_{k=1}^{\infty} \binom{q}{2k} \left(\frac{y^2}{q-1} \right)^k \right)^{p/q} \leq 1 + \frac{p}{q} \sum_{k=1}^{\infty} \binom{q}{2k} \left(\frac{y^2}{q-1} \right)^k.$$

Έπεται ότι

$$F_q(1, y) = \left(1 + \sum_{k=1}^{\infty} \binom{q}{2k} \left(\frac{y^2}{q-1} \right)^k \right)^{1/q} \leq \left(1 + \frac{p}{q} \sum_{k=1}^{\infty} \binom{q}{2k} \left(\frac{y^2}{q-1} \right)^k \right)^{1/p}.$$

Παρατηρούμε ότι

$$\begin{aligned} \frac{p}{q} \binom{q}{2k} \left(\frac{1}{q-1} \right)^k &= \frac{p}{q} \frac{q(q-1) \cdots (q-2k+1)}{(2k)! (q-1)^k} = \frac{p(q-2) \cdots (q-2k+1)}{(2k)! (q-1)^{k-1}} \\ &= \frac{p(2-q) \cdots (2k-1-q)}{(2k)! (q-1)^{k-1}} \leq \frac{p(2-p) \cdots (2k-1-p)}{(2k)! (p-1)^{k-1}} = \binom{p}{2k} \left(\frac{1}{p-1} \right)^k. \end{aligned}$$

[Παρατηρήστε ότι $(q-2) \cdots (q-2k+1) = (2-q) \cdots (2k-1-q)$ γιατί το πλήθος των όρων στο γινόμενο είναι άρτιο, και ότι $\frac{(2-q) \cdots (2k-1-q)}{(q-1)^{k-1}} \leq \frac{(2-p) \cdots (2k-1-p)}{(p-1)^{k-1}}$ γιατί $p < q$.] Επιστρέφοντας στην εκτίμηση για την $F_q(x, y)$ παίρνουμε

$$F_q(1, y) \leq \left(1 + \frac{p}{q} \sum_{k=1}^{\infty} \binom{q}{2k} \left(\frac{y^2}{q-1} \right)^k \right)^{1/p} \leq \left(1 + \sum_{k=1}^{\infty} \binom{p}{2k} \left(\frac{y^2}{p-1} \right)^k \right)^{1/p} = F_p(1, y).$$

(β) Υποθέτουμε τώρα ότι $1 < p < q \leq 2$ και ότι $|r_p y| \geq 1$. Θέτουμε $\lambda = r_q/r_p$ και $\mu = 1/|r_p y|$.

Γράφουμε

$$F_q(1, y) = \left(\frac{1}{2} |1 + \lambda r_p y|^q + \frac{1}{2} |1 - \lambda r_p y|^q \right)^{1/q} = \frac{1}{\mu} \left(\frac{1}{2} |\lambda + \mu|^q + \frac{1}{2} |\lambda - \mu|^q \right)^{1/q}.$$

Παρατηρήστε ότι $0 < \lambda, \mu \leq 1$. Άρα, $|\lambda - \mu| \leq 1 - \lambda\mu$ και $\lambda + \mu \leq 1 + \lambda\mu$. Μπορούμε λοιπόν να γράψουμε

$$F_q(1, y) \leq \frac{1}{\mu} \left(\frac{1}{2} |1 + \lambda\mu|^q + \frac{1}{2} |1 - \lambda\mu|^q \right)^{1/q}.$$

Έχουμε $\lambda\mu = r_q z$, όπου $z = \frac{1}{r_p |r_p y|}$ και $|r_p z| = \frac{1}{|r_p y|} \leq 1$. Από το (α) συμπεραίνουμε ότι

$$\begin{aligned} \frac{1}{\mu} \left(\frac{1}{2} |1 + \lambda\mu|^q + \frac{1}{2} |1 - \lambda\mu|^q \right)^{1/q} &= \frac{1}{\mu} \left(\frac{1}{2} |1 + r_q z|^q + \frac{1}{2} |1 - r_q z|^q \right)^{1/q} \leq \frac{1}{\mu} \left(\frac{1}{2} |1 + r_p z|^p + \frac{1}{2} |1 - r_p z|^p \right)^{1/p} \\ &= \frac{1}{\mu} \left(\frac{1}{2} |1 + \mu|^p + \frac{1}{2} |1 - \mu|^p \right)^{1/p} = \left(\frac{1}{2} \left| 1 + \frac{1}{\mu} \right|^p + \frac{1}{2} \left| 1 - \frac{1}{\mu} \right|^p \right)^{1/p} \\ &= \left(\frac{1}{2} |1 + r_p y|^p + \frac{1}{2} |1 - r_p y|^p \right)^{1/p} = F_p(1, y). \end{aligned}$$

Δηλαδή, $F_q(1, y) \leq F_p(1, y)$.

(γ) Τέλος, υποθέτουμε ότι $2 \leq p < q < \infty$. Θα χρησιμοποιήσουμε δυϊσμό. Θεωρούμε τους συζυγείς εκθέτες p' και q' των p και q . Παρατηρήστε ότι $1 < q' < p' \leq 2$. Αν $\lambda = r_{p'}/r_{q'} = \frac{\sqrt{q'-1}}{\sqrt{p'-1}}$ και αν $\kappa(1, 1) = \kappa(-1, -1) = 1 + \lambda$ και $\kappa(1, -1) = \kappa(-1, 1) = 1 - \lambda$, τότε τα (α) και (β) δείχνουν ότι για τον τελεστή $T : L^{q'}([-1, 1]^1) \rightarrow L^{p'}([-1, 1]^1)$, $1 < p' < q' \leq 2$, που ορίζεται μέσω της

$$T(f)(x) = \int_{[-1, 1]^1} \kappa(x, y) f(y) dy$$

ισχύει η ανισότητα

$$\|T(f)\|_{p'} \leq \|f\|_{q'}.$$

Η κ είναι συμμετρική στον $\{-1, 1\}^1 \times \{-1, 1\}^1$, άρα $T^* = T$, όπου $T^* : L^p([-1, 1]^1) \rightarrow L^q([-1, 1]^1)$ είναι ο συζυγής τελεστής του T . Από την $\|T^*\| = \|T\|$ έπεται ότι

$$\|T(f)\|_q \leq \|f\|_p$$

για κάθε $f : [-1, 1]^1 \rightarrow \mathbb{R}$. Παρατηρώντας ότι

$$\frac{q' - 1}{p' - 1} = \frac{p - 1}{q - 1}$$

έχουμε το ζητούμενο. □

Η ανισότητα του Θεωρήματος 2.2.3 επεκτείνεται εύκολα στην περίπτωση που τα x, y είναι διανύσματα σε έναν χώρο με νόρμα.

Θεώρημα 2.2.4. Έστω $(X, \|\cdot\|)$ ένας χώρος με νόρμα. Για κάθε $x, y \in X$ και $1 < p < \infty$ ορίζουμε

$$F_p(x, y) = \left(\frac{1}{2} \|x + r_p y\|^p + \frac{1}{2} \|x - r_p y\|^p \right)^{1/p},$$

όπου

$$r_p = \frac{1}{\sqrt{p-1}}.$$

Τότε, η $F_p(x, y)$ είναι φθίνουσα συνάρτηση του p στο $(1, \infty)$.

Θα χρειαστούμε το εξής λήμμα.

Λήμμα 2.2.5. Έστω $(X, \|\cdot\|)$ ένας χώρος με νόρμα. Αν $x, z \in X$ και $-1 \leq \lambda < 1$, τότε

$$\|x + \lambda z\| \leq \frac{1}{2} (\|x + z\| + \|x - z\|) + \frac{\lambda}{2} (\|x + z\| - \|x - z\|).$$

Απόδειξη. Γράφουμε

$$x + \lambda z = \left(\frac{1+\lambda}{2} \right) (x + z) + \left(\frac{1-\lambda}{2} \right) (x - z),$$

οπότε

$$\begin{aligned} \|x + \lambda z\| &\leq \left(\frac{1+\lambda}{2} \right) \|x + z\| + \left(\frac{1-\lambda}{2} \right) \|x - z\| \\ &= \frac{1}{2} (\|x + z\| + \|x - z\|) + \frac{\lambda}{2} (\|x + z\| - \|x - z\|). \end{aligned}$$

□

Απόδειξη του Θεωρήματος 2.2.4. Έστω $1 < p < q < \infty$. Θέτουμε $u = x + r_p y$, $v = x - r_p y$ και $\lambda = r_q / r_p$. Παρατηρήστε ότι $0 < \lambda < 1$, οπότε το Λήμμα 2.2.5 μας δίνει

$$\begin{aligned} \|x + \lambda r_p y\| &\leq \frac{1}{2} (\|x + r_p y\| + \|x - r_p y\|) + \frac{\lambda}{2} (\|x + r_p y\| - \|x - r_p y\|) \\ &= \frac{1}{2} (\|u\| + \|v\|) + \frac{\lambda}{2} (\|u\| - \|v\|) \end{aligned}$$

και, εντελώς ανάλογα,

$$\|x - \lambda r_p y\| \leq \frac{1}{2} (\|u\| + \|v\|) - \frac{\lambda}{2} (\|u\| - \|v\|).$$

Μπορούμε λοιπόν να γράψουμε

$$\begin{aligned}
\left(\frac{1}{2}\|x + r_q y\|^q + \frac{1}{2}\|x - r_q y\|^q\right)^{1/q} &= \left(\frac{1}{2}\|x + \lambda r_p y\|^q + \frac{1}{2}\|x - \lambda r_p y\|^q\right)^{1/q} \\
&\leq \left(\frac{1}{2}\left(\frac{1}{2}(\|u\| + \|v\|) + \frac{\lambda}{2}(\|u\| - \|v\|)\right)^q + \frac{1}{2}\left(\frac{1}{2}(\|u\| + \|v\|) - \frac{\lambda}{2}(\|u\| - \|v\|)\right)^q\right)^{1/q} \\
&\leq \left(\frac{1}{2}\left(\frac{1}{2}(\|u\| + \|v\|) + \frac{1}{2}(\|u\| - \|v\|)\right)^p + \frac{1}{2}\left(\frac{1}{2}(\|u\| + \|v\|) - \frac{1}{2}(\|u\| - \|v\|)\right)^p\right)^{1/p} \\
&= \left(\frac{1}{2}\|u\|^p + \frac{1}{2}\|v\|^p\right)^{1/p} = \left(\frac{1}{2}\|x + r_p y\|^p + \frac{1}{2}\|x - r_p y\|^p\right)^{1/p},
\end{aligned}$$

όπου, για την δεύτερη ανισότητα, χρησιμοποιήσαμε το Θεώρημα 2.2.3 με $x = \frac{\|u\| + \|v\|}{2}$ και $y = \frac{1}{2r_p}(\|u\| - \|v\|)$. $\square$

Απόδειξη του Θεωρήματος 2.2.2. Με επαγωγή ως προς n . Για $n = 1$ το ζητούμενο είναι ακριβώς η ανισότητα που αποδείξαμε στο Θεώρημα 2.2.3 (θεωρούμε $x_\emptyset = x$, $x_{\{1\}} = y$ και $w_{\{1\}}(-1) = 1$, $w_{\{1\}}(1) = 1$.)

Υποθέτουμε ότι το αποτέλεσμα ισχύει για $k = n-1$ και γράφουμε $\{-1, 1\}^n = \{-1, 1\}^{n-1} \times \{-1, 1\}^1$, $\mu_n = \mu_{n-1} \times \mu_{\{n\}}$. Γράφουμε $\mathcal{P}(n-1)$ για το σύνολο των υποσυνόλων του $\{1, \dots, n-1\}$ και $\mathcal{P}(n)$ για το σύνολο των υποσυνόλων του $\{1, \dots, n\}$. Για κάθε $T \in \mathcal{P}(n-1)$ ορίζουμε $T^+ = T \cup \{n\}$. Με αυτόν τον συμβολισμό, $\mathcal{P}(n) = \mathcal{P}(n-1) \cup \{T^+ : T \in \mathcal{P}(n-1)\}$. Γράφουμε

$$\sum_S r_p^{|S|} w_S(\omega) x_S = u_p(x) + x_n(t) r_p v_p(x),$$

όπου $x \in \{-1, 1\}^{n-1}$, $t \in \{-1, 1\}^{\{n\}} = \{-1, +1\}$ και $\omega = (x, t) \in \{-1, 1\}^n$. Ορίζουμε

$$u_p = \sum_{T \in \mathcal{P}(n-1)} r_p^{|T|} w_T x_T \quad \text{και} \quad v_p = \sum_{T \in \mathcal{P}(n-1)} r_p^{|T|} w_T x_{T^+}.$$

Εντελώς ανάλογα,

$$\sum_S r_q^{|S|} w_S x_S = u_q + x_n r_q v_q,$$

όπου

$$u_q = \sum_{T \in \mathcal{P}(n-1)} r_q^{|T|} w_T x_T \quad \text{και} \quad v_q = \sum_{T \in \mathcal{P}(n-1)} r_q^{|T|} w_T x_{T^+}.$$

Επομένως αρκεί να δείξουμε ότι

$$\|u_q + r_q x_n v_q\|_{L^q} \leq \|u_p + r_p x_n v_p\|_{L^p}.$$

Από την επαγωγική μας υπόθεση ισχύει

$$\begin{aligned}
\|u_q + r_q x_n v_q\|_{L^q} &= (\mathbb{E}_{n-1}(\|u_q + r_q x_n v_q\|^q))^{1/q} \\
&= \left(\mathbb{E}_{n-1} \left\| \sum_{T \in \mathcal{P}(n-1)} r_q^{|T|} w_T(x_T + x_n r_q x_{T^+}) \right\|^q \right)^{1/q} \\
&\leq \left(\mathbb{E}_{n-1} \left\| \sum_{T \in \mathcal{P}(n-1)} r_p^{|T|} w_T(x_T + x_n r_q x_{T^+}) \right\|^p \right)^{1/p} \\
&= (\mathbb{E}_{n-1} \|u_q + r_q x_n v_p\|^p)^{1/p}.
\end{aligned}$$

Σε αυτό το σημείο θα χρησιμοποιήσουμε την παρακάτω ανισότητα: Αν $(\Omega_1, \mathcal{S}_1, \mu_1)$, $(\Omega_2, \mathcal{S}_2, \mu_2)$ είναι χώροι μέτρου, $f : \Omega_1 \times \Omega_2 \rightarrow [0, +\infty]$, και $0 < p \leq q < +\infty$, τότε

$$\left(\int_{\Omega_1} \left(\int_{\Omega_2} f(x, y)^p d\mu(y) \right)^{q/p} d\mu_1(x) \right)^{1/q} \leq \left(\int_{\Omega_2} \left(\int_{\Omega_1} f(x, y)^q d\mu_1(x) \right)^{p/q} d\mu_2(y) \right)^{1/p}.$$

Άρα, αν $x \in \{-1, 1\}^{n-1}$, $t \in \{-1, +1\}$, τότε

$$\begin{aligned}
\|u_q(x) + x_n(t) r_q v_q(x)\|_{L^q} &= (\mathbb{E}_{\{n\}}(\mathbb{E}_{n-1} \|u_q(x) + x_n(t) r_q v_q(x)\|^q))^{1/q} \\
&\leq (\mathbb{E}_{\{\times\}}(\mathbb{E}_{n-1} \|u_p(x) + x_n(t) r_q v_p(x)\|^{q/p}))^{1/q} \\
&\leq (\mathbb{E}_{n-1}(\mathbb{E}_{\{n\}} \|u_p(x) + x_n(t) r_q v_p(x)\|^q))^{1/p} \\
&\leq (\mathbb{E}_{n-1} \mathbb{E}_{\{n\}} \|u_p(x) + x_n(t) r_p v_p(x)\|^p)^{1/p} \\
&= (\mathbb{E}_n \|u_p(x) + x_n(t) r_p v_p(x)\|^p)^{1/p} \\
&= \left\| \sum_{S \in \mathcal{P}(n)} r_p^{|S|} w_S x_S \right\|_{L^p}
\end{aligned}$$

όπου η τελευταία ανισότητα προέκυψε από την περίπτωση $k = 1$. □

2.3 Η ανισότητα Kahn-Kalai-Linial

Ένας από τους πιο σημαντικούς τρόπους μελέτης των επιρροών μιας Boolean συνάρτησης, είναι να δούμε πως αντιδρούν κάτω από τους «θορύβους». Ξεκινάμε αυτήν την ενότητα δίνοντας κάποιους βασικούς ορισμούς που θα μας οδηγήσουν στην έννοια του «θορύβου».

Ορισμός 2.3.1 (συσχετισμένα σημεία του διακριτού κύβου). Έστω $x \in \{-1, 1\}^n$ σημείο του διακριτού κύβου και $p \in [-1, 1]$. Ένα τυχαίο διάνυσμα y με τιμές στον διακριτό κύβο $\{-1, 1\}^n$

λέγεται p -συσχετισμένο με το x εάν για κάθε $i \in [n]$

$$y_i = \begin{cases} x_i, & \text{με πιθανότητα } \frac{1+p}{2} \\ -x_i, & \text{με πιθανότητα } \frac{1-p}{2}. \end{cases}$$

Με άλλα λόγια, το y είναι p -συσχετισμένο με το x αν κάθε συντεταγμένη του y προκύπτει αν αλλάξουμε πρόσημο στην αντίστοιχη συντεταγμένη του x με πιθανότητα $1-p$.

Για παράδειγμα, αν $p = 1$ τότε $y = x$, αφού η πιθανότητα να μεταβληθεί κάποια συντεταγμένη του x είναι μηδενική. Επίσης, για $p = 0$ έχουμε ότι το y είναι ομοιόμορφα τυχαίο, αφού για κάθε $1 \leq i \leq n$ η πιθανότητα να αλλάξει πρόσημο η συντεταγμένη x_i του x είναι ίση με $\frac{1}{2}$.

Χρησιμοποιώντας την έννοια του τελεστή θορύβου, μπορούμε ουσιαστικά αντί να κοιτάζουμε την τιμή $f(x)$ να κοιτάζουμε την τιμή $f(y)$ όπου y είναι μια θορυβημένη εκδοχή του x .

Ορισμός 2.3.2 (τελεστής θορύβου). Ο τελεστής θορύβου T_p είναι ένας τελεστής πάνω στον χώρο $\mathcal{H}_n$ των συναρτήσεων $f : \{-1, 1\}^n \rightarrow \mathbb{R}$, που ορίζεται ως εξής:

$$T_p f(x) = \mathbb{E}[f(y)]$$

όπου y είναι ένα τυχαίο διάνυσμα το οποίο είναι p -συσχετισμένο με το x .

Παρατήρηση 2.3.3. Έστω $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$ μια Boolean συνάρτηση με αναπαράσταση Fourier

$$f = \sum_{S \subseteq [n]} \widehat{f}(S) w_S.$$

Λόγω γραμμικότητας της μέσης τιμής έχουμε ότι

$$T_p f(x) = \sum_{S \subseteq [n]} \widehat{f}(S) T_p w_S(x)$$

όπου, για κάθε $S \subseteq [n]$,

$$\begin{aligned} T_p w_S(x) &= \prod_{i \in S} \mathbb{E}[y_i] = \prod_{i \in S} \left(\frac{1+p}{2} x_i - \frac{1-p}{2} x_i \right) \\ &= \prod_{i \in S} p x_i = p^{|S|} \prod_{i \in S} x_i = p^{|S|} w_S(x) \end{aligned}$$

και άρα

$$T_p f(x) = \sum_{S \subseteq [n]} \widehat{f}(S) p^{|S|} w_S(x).$$

Παραπάνω χρησιμοποιήθηκε το γεγονός ότι αν X, Y είναι ανεξάρτητες τυχαίες μεταβλητές τότε

$$\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$$

και άρα, αφού οι y_i είναι ανεξάρτητες, έπεται ότι

$$\mathbb{E} \left[\prod_{i \in S} y_i \right] = \prod_{i \in S} \mathbb{E}[y_i].$$

Από τα παραπάνω έχουμε άμεσα τους συντελεστές Fourier του τελεστή θορύβου T_p :

$$\widehat{T_p f}(S) = \widehat{f}(S) p^{|S|}.$$

Ορισμός 2.3.4 (p -νόρμα). Έστω $f : \{-1, 1\}^n \rightarrow \mathbb{R}$ και $p \geq 1$. Η p -νόρμα της f , την οποία θα συμβολίζουμε με $\|f\|_p$, ορίζεται ως η ποσότητα

$$\|f\|_p := (\mathbb{E}[|f|^p])^{\frac{1}{p}}$$

και ικανοποιεί της ιδιότητες της νόρμας στον χώρο $\mathcal{H}_n$.

Διασθητικά, κάθε $\|f\|_p$ «μετράει» την απόσταση της f από τη μηδενική συνάρτηση και ως εκ τούτου η απόσταση μεταξύ δύο συναρτήσεων $f, g \in \mathcal{H}_n$ μπορεί να εκφραστεί μέσω της ποσότητας $\|f - g\|_p$. Φυσικά, η φιλικότερη προς εμάς p -νόρμα είναι η $\|\cdot\|_2$, την οποία ήδη από την αρχή καταφέραμε να συνδέσουμε με τους συντελεστές Fourier της f . Πολύ συχνά όμως, απαιτείται και η μελέτη της p -νόρμας μιας συνάρτησης για τιμές $p \neq 2$. Από την ανισότητα Hölder προκύπτει άμεσα η μονοτονία της $\|f\|_p$ ως προς p : Αν $1 \leq p \leq q$ τότε

$$\|f\|_p \leq \|f\|_q.$$

Θα χρησιμοποιήσουμε μια συνέπεια της ανισότητας υπερσυσταλτότητας.

Θεώρημα 2.3.5 (ανισότητα υπερσυσταλτότητας). Για κάθε $p \in [0, 1]$ και $f : \{-1, 1\}^n \rightarrow \mathbb{R}$ ισχύει ότι

$$\|T_p f\|_2 \leq \|f\|_{1+p^2}.$$

Απόδειξη. Παρατηρούμε ότι αν $2 = 1 + ((1 + p^2) - 1)e^{2t} = 1 + p^2 e^{2t}$ τότε $p e^t = 1$, συνεπώς

$$P_t f = \sum_{S \subseteq [n]} \widehat{f}(S) e^{-t|S|} w_S = \sum_{S \subseteq [n]} \widehat{f}(S) p^{|S|} w_S = T_p f$$

για κάθε $f : \{-1, 1\}^n \rightarrow \mathbb{R}$. Από το Θεώρημα 2.2.1 συμπεραίνουμε ότι

$$\|T_p f\|_2 = \|P_t(f)\|_2 = \|P_t f\|_{1+((1+p^2)-1)e^{2t}} \leq \|f\|_{1+p^2}$$

όπως θέλαμε. □

Η ανισότητα υπερσυσταλτότητας παίζει βασικό ρόλο στην απόδειξη της ανισότητας Kahn-Kalai-Linial που συνδέει τη διασπορά μιας Boolean συνάρτησης με την συνολική επιρροή της.

Θεώρημα 2.3.6 (ανισότητα Kahn-Kalai-Linial). Έστω $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$ μια Boolean συνάρτηση και έστω δ η μέγιστη επιρροή της f , δηλαδή $\delta = \max_i I_i(f)$. Τότε,

$$\text{Var}(f) \leq \frac{C}{\log(1/\delta)} \sum_{i=1}^n I_i(f).$$

όπου $C > 0$ απόλυτη σταθερά.

Απόδειξη. Ξεκινάμε από τη σχέση

$$\text{Var}(f) = \sum_{1 \leq |S| \leq n} \widehat{f}(S)^2$$

και σταθεροποιούμε $K > 0$ που θα επιλεγεί κατάλληλα στη συνέχεια. Σπάζοντας το άθροισμα του δεξιού μέλους σε δύο επιμέρους αθροίσματα, έχουμε ότι

$$\sum_{1 \leq |S| \leq n} \widehat{f}(S)^2 = \sum_{1 \leq |S| \leq K} \widehat{f}(S)^2 + \sum_{K < |S| \leq n} \widehat{f}(S)^2$$

και τώρα για τα S με $1 \leq |S| \leq K$, πολλαπλασιάζοντας κάθε όρο του πρώτου αθροίσματος με $|S|$, έχουμε ότι μεγαλώνει το άθροισμα και άρα

$$\begin{aligned} \sum_{1 \leq |S| \leq n} \widehat{f}(S)^2 &= \sum_{1 \leq |S| \leq K} \widehat{f}(S)^2 + \sum_{K < |S| \leq n} \widehat{f}(S)^2 \\ &\leq \sum_{1 \leq |S| \leq K} |S| \widehat{f}(S)^2 + \sum_{K < |S| \leq n} \widehat{f}(S)^2. \end{aligned}$$

Επίσης, για το δεύτερο άθροισμα έχουμε ότι, αφού για τα S με $K < |S| \leq n$ ισχύει ότι $\frac{|S|}{K} > 1$, πολλαπλασιάζοντας με $\frac{|S|}{K}$ κάθε όρο του αθροίσματος, το άθροισμα μεγαλώνει και άρα

$$\begin{aligned} \sum_{1 \leq |S| \leq K} |S| \widehat{f}(S)^2 + \sum_{K < |S| \leq n} \widehat{f}(S)^2 &\leq \sum_{1 \leq |S| \leq K} |S| \widehat{f}(S)^2 + \frac{1}{K} \sum_{K < |S| \leq n} |S| \widehat{f}(S)^2 \\ &\leq \sum_{1 \leq |S| \leq K} |S| \widehat{f}(S)^2 + \frac{1}{K} \sum_{i=1}^n I_i(f). \end{aligned}$$

Επιστρέφοντας ξανά στο πρώτο άθροισμα της παραπάνω ανισότητας έχουμε ότι, αφού για τα S με $1 \leq |S| \leq K$ ισχύει η ανισότητα

$$2^{2(K-|S|)} \geq 1,$$

πολλαπλασιάζοντας κάθε όρο του πρώτου αθροίσματος με $2^{2(K-|S|)}$ παίρνουμε

$$\text{Var}(f) \leq 2^{2K} \sum_{1 \leq |S| \leq K} \frac{1}{2^{2|S|}} |S| \widehat{f}(S)^2 + \frac{1}{K} \sum_{i=1}^n I_i(f).$$

Τώρα λόγω της αναπαράστασης Fourier του $T_p \partial_i f$ έχουμε ότι

$$T_p \partial_i f = \sum_{\{S \subseteq [n]: i \in S\}} \widehat{f}(S) p^{|S|-1} w_{S \setminus \{i\}}.$$

Από την ταυτότητα Parseval έπεται ότι

$$\|T_p \partial_i f\|_2^2 = \sum_{\{S \subseteq [n]: i \in S\}} \widehat{f}(S)^2 p^{2|S|-2}$$

και αθροίζοντας πάνω από όλα τα $i = 1, \dots, n$ συμπεραίνουμε ότι

$$\sum_{i=1}^n \|T_p \partial_i f\|_2^2 = \sum_{1 \leq |S| \leq n} \widehat{f}(S)^2 |S| p^{2|S|-2}.$$

Επιλέγουμε $p = \frac{1}{2} > 0$. Από την ανισότητα

$$\sum_{1 \leq |S| \leq K} |S| \widehat{f}(S)^2 \leq 2^{2K} \sum_{1 \leq |S| \leq K} \frac{1}{2^{2|S|}} |S| \widehat{f}(S)^2 = 2^{2K-2} \sum_{i=1}^n \|T_{1/2} \partial_i f\|_2^2$$

έχουμε τελικά ότι

$$\text{Var}(f) \leq 2^{2K-2} \sum_{i=1}^n \|T_{1/2} \partial_i f\|_2^2 + \frac{1}{K} \sum_{i=1}^n I_i(f).$$

Χρησιμοποιώντας το Θεώρημα 2.3.5 για $p = \frac{1}{2}$ έχουμε ότι

$$\begin{aligned} \text{Var}(f) &\leq 2^{2K-2} \sum_{i=1}^n \|\partial_i f\|_{5/4}^2 + \frac{1}{K} \sum_{i=1}^n I_i(f) \\ &\leq 2^{2K-2} \sum_{i=1}^n I_i(f)^{8/5} + \frac{1}{K} \sum_{i=1}^n I_i(f) \\ &\leq 2^{2K-2} \delta^{3/5} \sum_{i=1}^n I_i(f) + \frac{1}{K} \sum_{i=1}^n I_i(f) \\ &= \left(\frac{1}{K} + 2^{2K-2} \delta^{3/5} \right) \sum_{i=1}^n I_i(f). \end{aligned}$$

Τώρα, επιλέγοντας το K ώστε να ελαχιστοποιηθεί ο συντελεστής

$$\frac{1}{K} + 2^{2K-2} \delta^{3/5},$$

δηλαδή ουσιαστικά επιλέγοντας

$$K = \frac{3}{10} \log_2(1/\delta)$$

καταλήγουμε στη ζητούμενη ανισότητα. □

2.4 Η εικασία του Talagrand

Στις προηγούμενες δύο ενότητες είδαμε την ανισότητα Poincaré και την ανισότητα Kahn-Kalai-Linial. Η ανισότητα Poincaré μπορεί να διατυπωθεί στη μορφή

$$(2.4.1) \quad \text{Var}(f) \leq \sum_{i=1}^n I_i(f).$$

Η συνολική επιρροή $\sum_i I_i(f)$ στο δεξιό μέλος ισούται με το πλήθος των ακμών του διακριτού κύβου που διαχωρίζουν τα σύνολα όπου $f(x) = 1$ και $f(x) = -1$. Μπορούμε λοιπόν να την σκεφτόμαστε ως κάποιου είδους εμβαδόν επιφάνειας της f .

Η ανισότητα Kahn, Kalai και Linial δείχνει ότι η ανισότητα Poincaré επιδέχεται «λογαριθμική» βελτίωση: Υπάρχει απόλυτη σταθερά $C > 0$ τέτοια ώστε

$$(2.4.2) \quad \text{Var}(f) \leq C \frac{\sum_i I_i(f)}{\log(1/\max_i(I_i(f)))}.$$

Η ανισότητα KKL είναι ακριβής για τη συνάρτηση Tribes, χάνει όμως έναν όρο της τάξης του $\sqrt{n}/\log n$ για τη συνάρτηση πλειοψηφίας, που έχει όλες τις επιρροές της τάξης $1/\sqrt{n}$, και αυτό υποδεικνύει ότι ίσως η συνολική επιρροή $\sum I_i(f)$ δεν είναι η σωστή έννοια εμβαδού επιφάνειας για όλες τις Boolean συναρτήσεις.

Ορισμός 2.4.1 (συνάρτηση ευαισθησίας). Για κάθε $f : \{-1, 1\}^n \rightarrow \mathbb{R}$ θεωρούμε τη συνάρτηση ευαισθησίας

$$h_f(y) = |\{i \in [n] \mid f(y) \neq f(\sigma_i(y))\}|$$

της f . Μπορούμε να σκεφτόμαστε την ποσότητα $\mathbb{E}[\sqrt{h_f}]$ ως ένα άλλο είδος εμβαδού επιφάνειας για την f .

Στο [48, Θεώρημα 1.1], ο Talagrand απέδειξε την εξής ανισότητα.

Θεώρημα 2.4.2. Για κάθε Boolean συνάρτηση $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$ ισχύει ότι

$$(2.4.3) \quad \text{Var}(f) \leq \frac{1}{\sqrt{2}} \mathbb{E}[\sqrt{h_f}].$$

Μια ακριβής, ισχυρότερη μορφή της ανισότητας (2.4.3) αποδείχθηκε από τον Bobkov [7]. Η ανισότητα που απέδειξε ο Bobkov οδηγεί σε μια στοιχειώδη απόδειξη της ισοπεριμετρικής ανισότητας στον χώρο του Gauss. Η ανισότητα (2.4.3) είναι ακριβής για γραμμικές συναρτήσεις κατωφλίου όπως η συνάρτηση πλειοψηφίας, όχι όμως για την Tribes. Έτσι, καμία από τις δύο ανισότητες (2.4.2) και (2.4.3) δεν συνεπάγεται την άλλη. Οδηγούμαστε λοιπόν στο εξής ερώτημα:

Ερώτημα 2.4.3. Ποια είναι η σωστή έννοια εμβαδού επιφάνειας για Boolean συναρτήσεις; Υπάρχει κάποια ανισότητα από την οποία να έπονται τόσο η (2.4.2) όσο και η (2.4.3);

Σαν ένα βήμα προς αυτή την κατεύθυνση, ο Talagrand διατύπωσε στο [51] την εικασία ότι ισχύει μια ανισότητα ισχυρότερη από την (2.4.3). Συγκεκριμένα, ότι υπάρχει σταθερά $\beta > 0$ τέτοια ώστε, για κάθε Boolean συνάρτηση $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$,

$$(2.4.4) \quad \mathbb{E} \left[\sqrt{h_f} \right] \geq \beta \cdot \text{Var}(f) \cdot \left(\log \left(\frac{e}{\sum I_i(f)^2} \right) \right)^{1/2}.$$

Ο Talagrand απέδειξε ότι υπάρχουν $\alpha \leq 1/2$ και σταθερά $\beta > 0$ τέτοια ώστε

$$\int_{\{-1,1\}^n} \sqrt{h_f(x)} d\mu_n(x) \geq \beta \cdot \text{Var}(f) \left(\log \frac{e}{\text{Var}(f)} \right)^{1/2-\alpha} \cdot \left(\log \left(\frac{e}{\sum I_i(f)^2} \right) \right)^\alpha,$$

όμως η απόδειξή του δεν μπορούσε να φτάσει μέχρι την τιμή $\alpha = 1/2$, και το αποτέλεσμα αυτό δεν ήταν (παραλίγο) αρκετό ώστε να δώσει τη λογαριθμική βελτίωση στην ανισότητα KKL.

Μια άλλη έννοια εμβαδού επιφάνειας είναι το *σύνολο κορυφών* ∂f της f .

Ορισμός 2.4.4 (σύνολο κορυφών). Έστω $f : \{-1, 1\}^n \rightarrow \mathbb{R}$. Το σύνολο κορυφών της f είναι το σύνολο

$$\partial f = \left\{ y \in \{-1, 1\}^n \mid \exists i \text{ τ.ω. } f(y) \neq f(\sigma_i(y)) \right\}.$$

Το ∂f είναι η ξένη ένωση του *εσωτερικού συνόρου κορυφών*

$$\partial^+ f = \left\{ y \in \{-1, 1\}^n \mid f(y) = 1 \text{ και υπάρχει } i \text{ τ.ω. } f(\sigma_i(y)) = -1 \right\}$$

και του *εξωτερικού συνόρου κορυφών*

$$\partial^- f = \left\{ y \in \{-1, 1\}^n \mid f(y) = -1 \text{ και υπάρχει } i \text{ τ.ω. } f(\sigma_i(y)) = 1 \right\}.$$

Από την ανισότητα Cauchy-Schwarz έπεται ότι

$$\mathbb{E} \left[\sqrt{h_f} \right] \leq \sqrt{\mathbb{E} [h_f] \cdot \mu_n(\partial f)} = \sqrt{\mu_n(\partial f) \sum_i I_i(f)},$$

άρα η παραπάνω εικασία ισχυροποιεί την ανισότητα KKL στο διάστημα $\text{Var}(f) = \Omega(1)$ (δείτε παρακάτω το Θεώρημα 2.4.7 για έναν υπολογισμό).

Η ανισότητα (2.4.2) γενικεύτηκε περαιτέρω σε μια άλλη κατεύθυνση από τον Talagrand [49], ο οποίος απέδειξε το εξής:

Θεώρημα 2.4.5. Υπάρχει απόλυτη σταθερά $C_T > 0$ τέτοια ώστε, για κάθε $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$,

$$(2.4.5) \quad \text{Var}(f) \leq C_T \sum_{i=1}^n \frac{I_i(f)}{1 + \log(1/I_i(f))}.$$

Είναι γνωστό ότι αυτή η ανισότητα είναι ακριβής με την έννοια ότι για κάθε n -άδα τιμών των επιρροών υπάρχουν παραδείγματα για τα οποία ισχύει περίπου ως ισότητα [39].

Ανισότητες όπως οι (2.4.2), (2.4.4) και (2.4.5), σε σύνδεση με την συγκέντρωση της επιρροής [23] και ακριβή φαινόμενα κατωφλίου [24], έχουν χρησιμοποιηθεί σε πολλές περιοχές των μαθηματικών και της επιστήμης υπολογιστών, όπως η θεωρία μάθησης [45], οι μετρικές εμφυτεύσεις [41], η θεωρία διήθησης [6], η κλασική και κβαντική πολυπλοκότητα επικοινωνίας [47], η δυσκολία της προσέγγισης [14], και επίσης η δυναμική κοινωνικών δικτύων [43] και η στατιστική φυσική [3]. Για μια γενική επισκόπηση δείτε το [35].

Η αρχική απόδειξη του Talagrand για το Θεώρημα 2.4.5, αλλά και άλλες μεταγενέστερες αποδείξεις (βλ. για παράδειγμα [13]), βασίζονται όλες στην αρχή της υπερσυσταλτότητας.

Οι Eldan και Gross ανέπτυξαν μια νέα προσέγγιση για την απόδειξη όλων αυτών των ανισοτήτων. Οι αποδείξεις τους βασίζονται στον στοχαστικό λογισμό (pathwise analysis), αποφεύγουν τη χρήση της υπερσυσταλτότητας, και μάλιστα χρησιμοποιούν κλασική Boolean ανάλυση Fourier με φειδώ. Με αυτές τις τεχνικές αποδεικνύουν ότι η εικασία του Talagrand ισχύει:

Θεώρημα 2.4.6. Υπάρχει απόλυτη σταθερά $C > 0$ τέτοια ώστε για κάθε $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$,

$$\mathbb{E} \left[\sqrt{h_f} \right] \geq C \cdot \text{Var}(f) \cdot \sqrt{\log \left(2 + \frac{e}{\sum_i I_i(f)^2} \right)}.$$

Απέδειξαν μάλιστα ένα ισχυρότερο θεώρημα, το οποίο έχει ως άμεσο πόρισμα το Θεώρημα 2.4.6.

Θεώρημα 2.4.7. Υπάρχει απόλυτη σταθερά $C > 0$ τέτοια ώστε να ισχύει το εξής. Για κάθε $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$, υπάρχει μια συνάρτηση $g_f : \{-1, 1\}^n \rightarrow [0, 1]$ με $\mathbb{E}[g^2] \leq 2\text{Var}(f)$ τέτοια ώστε για κάθε $1/2 \leq p < 1$,

$$(2.4.6) \quad \mathbb{E} \left[h_f^p \right] \geq C \text{Var}(f) \cdot \left(\log \left(2 + \frac{e}{\sum_i I_i(f)^2} \right) \right)^p.$$

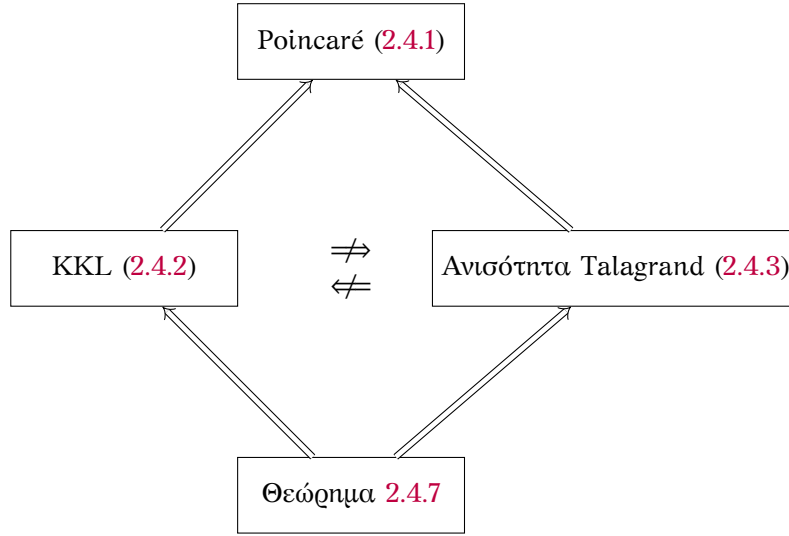
Επιλέγοντας $p = 1/2$ στο Θεώρημα 2.4.7 παίρνουμε τόσο την ανισότητα KKL σε πλήρη γενικότητα όσο και ένα νέο κάτω φράγμα για την συνολική επιρροή στο πνεύμα της ισοπεριμετρικής ανισότητας. Από την ανισότητα Cauchy-Schwarz, έχουμε ότι

$$\mathbb{E} \left[\sqrt{h_f g} \right] \leq \sqrt{\mathbb{E}[h_f]} \sqrt{\mathbb{E}[g^2]} \leq \sqrt{I(f)} \sqrt{2\text{Var}(f)}$$

και εισάγοντας αυτήν την ανισότητα στην (2.4.6) παίρνουμε

$$(2.4.7) \quad \text{Var}(f) \leq C \cdot \frac{\sum_i I_i(f)}{\log \left(\frac{e}{\sum_i I_i(f)^2} \right)}.$$

Από αυτήν την ανισότητα μπορούμε να δούμε τη σχέση ανάμεσα στην ανισότητα Poincaré, την ανισότητα KKL, το θεώρημα του Talagrand και το Θεώρημα 2.4.7, την οποία συνοψίζουμε στο Σχήμα 2.1.



Σχήμα 2.1: Σχέσεις των ανισοτήτων

Θεώρημα 2.4.7 $\Rightarrow$ **KKL**. Αν θέσουμε $\delta = \max_i I_i(f)$ τότε η (2.4.7) μας δίνει την

$$(2.4.8) \quad \text{Var}(f) \leq C \cdot \frac{\sum_i I_i(f)}{\log\left(\frac{1}{\delta \sum_i I_i(f)}\right)}.$$

Διακρίνουμε δύο περιπτώσεις:

- Αν $\delta \sum_i I_i(f) \leq \delta^{1/2}$ τότε από την (2.4.8) προκύπτει άμεσα ότι

$$\text{Var}(f) \leq 2C \frac{\sum_i I_i(f)}{\log(1/\delta)}.$$

- Αν $\delta \sum_i I_i(f) \geq \delta^{1/2}$ τότε έχουμε

$$\sum_i I_i(f) \geq \frac{1}{\delta^{1/2}} \geq \text{Var}(f) \frac{1}{\delta^{1/2}} \geq \frac{1}{2} \text{Var}(f) \log\left(\frac{1}{\delta}\right)$$

για $\delta < 1$ (αλλιώς δεν έχουμε τίποτα να δείξουμε), το οποίο μας δίνει πάλι

$$\text{Var}(f) \leq 2 \frac{\sum_i I_i(f)}{\log(1/\delta)}.$$

Θεώρημα 2.4.7 $\Rightarrow$ **Ευστάθεια της ισοπεριμετρικής ανισότητας**. Η ισοπεριμετρική ανισότητα ακμών [28, Ενότητα 3] αποδείχθηκε από τον Hart.

Θεώρημα 2.4.8. Έστω $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$ μια Boolean συνάρτηση και έστω

$$A = \{x \in \{-1, 1\}^n \mid f(x) = 1\}.$$

Τότε,

$$(2.4.9) \quad \sum_{i=1}^n I_i(f) \geq 2\mu_n(A) \log_2 \left(\frac{1}{\mu_n(A)} \right),$$

με ισότητα αν και μόνο αν το A είναι υποκύβος.

Υποθέτουμε ότι $\mathbb{E}[f] \leq 0$, οπότε $\mu_n(A) \leq 1/2$, και ότι η f είναι ακραία για την ισοπεριμετρική ανισότητα έως μία σταθερά, δηλαδή

$$\sum_{i=1}^n I_i(f) \leq C\mu_n(A) \log_2 \left(\frac{1}{\mu_n(A)} \right)$$

για κάποια σταθερά $C > 0$. Εφόσον $\text{Var}(f) = 4\mu_n(A)(1 - \mu_n(A)) \geq 2\mu_n(A)$, συμπεραίνουμε ότι

$$(2.4.10) \quad \sum_{i=1}^n I_i(f) \leq C\text{Var}(f) \log_2 \left(\frac{2}{\text{Var}(f)} \right).$$

Υποθέτουμε επίσης ότι η f είναι μονότονη. Τότε, $\text{Var}(f) \geq \sum_i I_i(f)^2$, και από τις (2.4.7) και (2.4.10) παίρνουμε ότι

$$\text{Var}(f) \leq C \frac{\sum_i I_i(f)}{\log \left(\frac{2}{\sum_i I_i(f)^2} \right)} \leq C \frac{\sum_i I_i(f)}{\log \left(\frac{2}{\text{Var}(f)} \right)} \leq C' \text{Var}(f).$$

Ειδικότερα, οι δύο παρονομαστές διαφέρουν κατά έναν σταθερό παράγοντα:

$$\log \left(\frac{2}{\sum_i I_i(f)^2} \right) = \Theta \left(\log \left(\frac{2}{\text{Var}(f)} \right) \right),$$

και αυτό συνεπάγεται ότι η μάζα Fourier του πρώτου επιπέδου είναι ανάλογη με μια δύναμη της διασποράς.

2.5 Ευστάθεια της ανισότητας του Talagrand

Στο Κεφάλαιο 5 δίνουμε μια νέα απόδειξη του Θεωρήματος 2.4.5 χρησιμοποιώντας τεχνικές στοχαστικού λογισμού, και αποδεικνύουμε μια ισχυρότερη μορφή που μπορούμε να την βλέπουμε ως μια εκδοχή ευστάθειας της (2.4.5) συναρτήσεως του συνόρου κορυφών της f : Αν έχουμε περίπου ισότητα στην (2.4.5), τότε τόσο το εσωτερικό όσο και το εξωτερικό σύνορο κορυφών της f είναι μεγάλα. Η διατύπωση του θεωρήματος είναι η εξής.

Θεώρημα 2.5.1. Υπάρχει απόλυτη σταθερά $C_B > 0$ τέτοια ώστε: για κάθε $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$, αν θέσουμε

$$(2.5.1) \quad T(f) = \sum_{i=1}^n \frac{I_i(f)}{1 + \log(1/I_i(f))} \quad \text{και} \quad r_{\text{Tal}} = \frac{\text{Var}(f)}{T(f)}$$

τότε

$$\mu_n(\partial^\pm f) \geq \frac{r_{\text{Tal}}}{C_B \log \frac{C_B}{r_{\text{Tal}}}} \text{Var}(f).$$

Μπορούμε να εφαρμόσουμε άμεσα το Θεώρημα 2.5.1 σε δύο από τις συναρτησιακές ανισότητες που συζητήσαμε στις προηγούμενες ενότητες – την ισοπεριμετρική ανισότητα και την ανισότητα KKL – δείχνοντας ότι αν οποιαδήποτε από αυτές τις ανισότητες είναι ακριβής έως σταθερά, τότε η συνάρτηση πρέπει να έχει μεγάλο σύνορο κορυφών.

Η ισοπεριμετρική ανισότητα και το σύνορο κορυφών. Είναι φυσιολογικό να ρωτήσουμε πόσο ισχυρή είναι η ισοπεριμετρική ανισότητα του Θεωρήματος 2.4.8. Είναι σωστό ότι αν έχουμε περίπου ισότητα στην (2.4.9) τότε το A είναι με κάποια έννοια κοντά σε υποκύβο; Μια απάντηση σε αυτό το ερώτημα δόθηκε στο [22] για σύνολα A τα οποία είναι $(1 + \varepsilon)$ -κοντά στο να ικανοποιούν την ανισότητα. Εικασίες σχετικά με τα σύνολα για τα οποία η ανισότητα είναι ακριβής έως έναν σταθερό πολλαπλασιαστικό παράγοντα έχουν διατυπωθεί στο [32]. Θα δούμε ότι αν η συνάρτηση είναι $O(1)$ -κοντά στο να ικανοποιεί την ανισότητα (2.4.9) τότε ένα σταθερό ποσοστό του συνόλου A περιέχεται στο εσωτερικό σύνορο κορυφών της (ενώ για τα ακραία σύνολα, το σύνορο κορυφών είναι ολόκληρο το σύνολο A).

Πόρισμα 2.5.2. Έστω $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$ μια Boolean συνάρτηση με $\mathbb{E}[f] \leq 0$ και $A = \{x \in \{-1, 1\}^n \mid f(x) = 1\}$. Θέτουμε

$$r_{\text{Iso}} = \frac{2\mu_n(A) \log_2 \frac{1}{\mu_n(A)}}{\sum_{i=1}^n I_i(f)}.$$

Τότε,

$$\mu_n(\partial A) \geq \frac{r_{\text{Iso}}}{2C_B \log \left(\frac{2C_B}{r_{\text{Iso}}} \right)} \mu_n(A)$$

όπου C_B είναι η απόλυτη σταθερά του Θεωρήματος 2.5.1.

Απόδειξη. Όπως στο Θεώρημα 2.5.1, θέτουμε $r_{\text{Tal}} = \frac{\text{Var}(f)}{T(f)}$. Παρατηρούμε ότι $I_i(f) \leq 2\mu_n(A)$ για κάθε $1 \leq i \leq n$. Εφόσον $\mu_n(A) \leq 1/2$, έχουμε

$$\text{Var}(f) = 4\mu_n(A)(1 - \mu_n(A)) \geq 2\mu_n(A).$$

Τότε, παίρνουμε το εξής κάτω φράγμα για την παράμετρο r_{Tal} :

$$\begin{aligned} r_{\text{Tal}} &= \frac{\text{Var}(f)}{\sum_i \frac{I_i(f)}{1 + \log(1/I_i(f))}} \geq \frac{\text{Var}(f)}{\sum_i \frac{I_i(f)}{1 + \log(1/2\mu_n(A))}} = \frac{r_{\text{Iso}} \left(1 + \log \left(\frac{1}{2\mu_n(A)} \right) \right)}{\frac{2}{\log 2} \mu_n(A) \log \frac{1}{\mu_n(A)}} \text{Var}(f) \\ &\geq \frac{r_{\text{Iso}} \log 2}{2\mu_n(A)} \text{Var}(f) \geq \frac{r_{\text{Iso}}}{2}. \end{aligned}$$

και από το Θεώρημα 2.5.1 συμπεραίνουμε ότι

$$\mu_n(\partial^\pm f) \geq \frac{r_{\text{Tal}}}{C_B \log \frac{C_B}{r_{\text{Tal}}}} \text{Var}(f) \geq \frac{r_{\text{Iso}}}{2C_B \log \left(\frac{2C_B}{r_{\text{Iso}}} \right)} \mu_n(A).$$

□

Ανισότητα KKL και σύνορο κορυφών. Στην αρχική του διατύπωση, το Θεώρημα KKL [34, Θεώρημα 3.1], το οποίο είναι άμεση συνέπεια της (2.4.2), ισχυρίζεται ότι κάθε Boolean συνάρτηση πρέπει να έχει μια μεταβλητή με σχετικά μεγάλη επιρροή: Υπάρχει μια απόλυτη σταθερά $C > 0$ τέτοια ώστε για κάθε $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$, υπάρχει ένας δείκτης $i \in [n]$ με

$$I_i(f) \geq C \cdot \text{Var}(f) \frac{\log n}{n}.$$

Το δεύτερο πόρισμά μας ισχυρίζεται ότι αν όλες οι επιρροές είναι της τάξης $\text{Var}(f) \frac{\log n}{n}$ τότε η συνάρτηση πρέπει να έχει μεγάλο (εσωτερικό και εξωτερικό) σύνορο κορυφών.

Πόρισμα 2.5.3. Υποθέτουμε ότι για κάποιον $C \leq \sqrt{n}$, έχουμε $I_i(f) \leq C \cdot \text{Var}(f) \frac{\log n}{n}$ για κάθε i . Τότε, υπάρχει μια σταθερά c_{KKL} που εξαρτάται μόνο από την τιμή του C , τέτοια ώστε

$$\mu_n(\partial^\pm f) \geq c_{\text{KKL}} \text{Var}(f).$$

Απόδειξη. Σε αυτή την περίπτωση, έχουμε

$$\begin{aligned} r_{\text{Tal}} &= \frac{\text{Var}(f)}{\sum_i^n \frac{I_i(f)}{1+\log(1/I_i(f))}} \geq \frac{\text{Var}(f)}{\sum_i^n \frac{I_i(f)}{1+\log(C \cdot \text{Var}(f) \frac{\log n}{n})}} \geq \frac{\text{Var}(f) \left(1 + \log \left(\frac{n}{C \text{Var}(f) \log n} \right)\right)}{C \cdot \text{Var}(f) \log n} \\ &\geq \frac{\log n - \log(C \text{Var}(f) \log n)}{C \log n} > \frac{1}{4C}. \end{aligned}$$

Έτσι, από το Θεώρημα 2.5.1, υπάρχει μια σταθερά c_{KKL} που εξαρτάται μόνο από το C τέτοια ώστε

$$\mu_n(\partial^\pm f) \geq c_{\text{KKL}} \text{Var}(f).$$

□

Τέλος, βελτιώνουμε μια ανισότητα των Keller και Kindler [36]. Έστω $S_\varepsilon(f)$ η ευστάθεια θορύβου της f , δηλαδή

$$(2.5.2) \quad S_\varepsilon(f) = \text{Cov}_{x \sim \mu_n, y \sim N_\varepsilon(x)} [f(x), f(y)],$$

όπου $N_\varepsilon(x)$ είναι ένα τυχαίο διάνυσμα που η i -οστή συντεταγμένη του είναι ίση με x_i με πιθανότητα $1 - \varepsilon$ και ίση με ένα ομοιόμορφα τυχαίο ψηφίο με πιθανότητα ε .

Θεώρημα 2.5.4. Υπάρχουν απόλυτες σταθερές $C, c > 0$ τέτοιες ώστε

$$(2.5.3) \quad S_\varepsilon(f) \leq C \cdot \text{Var}(f) \left(\sum_{i=1}^n I_i(f)^2 \right)^{c\varepsilon}.$$

Το φράγμα που αποδεικνύεται στο [36] είναι παρόμοιο αλλά στη θέση του όρου $\text{Var}(f)$ υπάρχει μια σταθερά. Έτσι, το Θεώρημα 2.5.4 είναι ισχυρότερο όταν $\text{Var}(f) = o(1)$. Το θεώρημα αυτό χρησιμοποιείται στην απόδειξη του Θεωρήματος 2.4.5. Η σχέση ανάμεσα στις επιρροές και την ευαισθησία θορύβου αποδείχθηκε για πρώτη φορά στο [5], όπου αποδεικνύεται ένα ποιοτικό φράγμα της ίδιας φύσης.

2.6 Αρμονική επέκταση Boolean συναρτήσεων

Σε αυτήν την ενότητα εισάγουμε την έννοια της αρμονικής επέκτασης μιας Boolean συνάρτησης. Γνωρίζουμε ότι κάθε Boolean συνάρτηση $f : \{-1, 1\}^n \rightarrow \mathbb{R}$ γράφεται μονοσήμαντα ως άθροισμα μονωνύμων

$$(2.6.1) \quad f(y) = \sum_{S \subseteq [n]} \widehat{f}(S) w_S = \sum_{S \subseteq [n]} \widehat{f}(S) \prod_{i \in S} y_i,$$

όπου οι συντελεστές Fourier $\widehat{f}(S)$ δίνονται από τις

$$(2.6.2) \quad \widehat{f}(S) = \mathbb{E}_y \left[f(y) \prod_{i \in S} y_i \right].$$

Χρησιμοποιώντας την (2.6.1) μπορούμε να επεκτείνουμε το πεδίο ορισμού μιας Boolean συνάρτησης από τον διακριτό κύβο $\{-1, 1\}^n$ σε ολόκληρον τον $\mathbb{R}^n$. Η συνάρτηση αυτή λέγεται αρμονική επέκταση της και συμβολίζεται πάλι με f . Εύκολα ελέγχουμε ότι

$$\frac{\partial^2 f}{\partial y_i^2} = 0$$

για κάθε $i = 1, \dots, n$, άρα η f είναι όντως αρμονική συνάρτηση. Επίσης, με αυτόν τον συμβολισμό έχουμε ότι $f(0) = \widehat{f}(\emptyset) = \mathbb{E}[f]$. Γενικά, αν $x \in [-1, 1]^n$ τότε η τιμή $f(x)$ της αρμονικής επέκτασης στο x είναι κυρτός συνδυασμός των τιμών της f στα σημεία $y \in \{-1, 1\}^n$. Πιο συγκεκριμένα,

$$(2.6.3) \quad f(x) = \sum_{y \in \{-1, 1\}^n} \kappa(x, y) f(y),$$

όπου $\kappa(x, y) = \prod_{i=1}^n (1 + x_i y_i) / 2$.

Όπως κάναμε και για την f , γράφουμε $\partial_i f$ για την αρμονική επέκταση της $\partial_i f$, και την θεωρούμε ως συνάρτηση ορισμένη στον κύβο $[-1, 1]^n$. Με απευθείας υπολογισμό ελέγχουμε ότι η αρμονική επέκταση της παραγώγου $\partial_i f$ ισούται με την μερική παράγωγο $\frac{\partial}{\partial x_i}$ της αρμονικής

επέκτασης της f .

Στον ορισμό του συντελεστή Fourier στην (2.6.2), η μέση τιμή παίρνεται ως προς το ομοιόμορφο μέτρο $\mu_n(\{y\}) = \frac{1}{2^n}$. Μπορούμε επίσης να αναλύσουμε μια συνάρτηση σε συντελεστές Fourier ως προς ένα μεροληπτικό μέτρο. Θα χρησιμοποιήσουμε ανάλυση αυτού του τύπου μόνο στην απόδειξη του Θεωρήματος 2.5.4.

Για κάθε $p = (p_1, \dots, p_n) \in [0, 1]^n$, συμβολίζουμε με $\mu_{n,p}$ το μέτρο

$$\mu_{n,p}(y) = \prod_{i=1}^n \frac{1 + y_i(2p_i - 1)}{2} = w_{(2p-1)}(y),$$

το οποίο δίνει στο i -οστό ψηφίο την τιμή 1 με πιθανότητα p_i . Ορίζουμε

$$(2.6.4) \quad \omega_i(y) = \frac{1}{2} \left(\frac{1 - 2p_i}{\sqrt{p_i(1-p_i)}} + y_i \frac{1}{\sqrt{p_i(1-p_i)}} \right),$$

και για κάθε $S \subseteq [n]$, ορίζουμε $\omega_S(y) = \prod_{i \in S} \omega_i(y)$. Τότε, κάθε συνάρτηση f γράφεται ως

$$(2.6.5) \quad \begin{aligned} f(y) &= \sum_{S \subseteq [n]} \widehat{f}_p(S) \omega_S(y) := \sum_{S \subseteq [n]} (\mathbb{E}_{\mu_{n,p}} [f \cdot \omega_S]) \omega_S(y) \\ &= \sum_{S \subseteq [n]} \left(\sum_{y \in \{-1,1\}^n} f(y) \omega_S(y) w_{2p-1}(y) \right) \omega_S(y) \end{aligned}$$

Οι συντελεστές $\widehat{f}_p := \mathbb{E}_{\mu_{n,p}} [f \cdot \omega_S]$ είναι οι « p -μεροληπτικοί» συντελεστές Fourier της f .

Η p -μεροληπτική επιρροή του i -οστού ψηφίου είναι η ποσότητα

$$I_i^p(f) = 4p_i(1-p_i) \mathbb{P}_{y \sim \mu_{n,p}} [f(x) \neq f(\sigma_i(x))].$$

Αν η f είναι μονότονη, τότε

$$(2.6.6) \quad I_i^p(f) = 2\sqrt{p_i} \sqrt{1-p_i} \widehat{f}_p(\{i\}).$$

Η επόμενη πρόταση συνδέει τους p -μεροληπτικούς συντελεστές Fourier με τις παραγώγους της f . Μια απόδειξη (με κάπως διαφορετικό συμβολισμό) δίνεται στο [44, Ενότητα 8].

Πρόταση 2.6.1. Έστω $S = \{i_1, \dots, i_k\} \subseteq [n]$ ένα σύνολο δεικτών, $x \in (-1, 1)^n$, και $p = \frac{1+x}{2}$. Τότε,

$$\partial_{i_1} \dots \partial_{i_k} f(x) = \left(\prod_{i \in S} \frac{4}{\sqrt{1-x_i^2}} \right) \widehat{f}_p(S).$$

Θα χρειαστούμε επίσης δύο λήμματα που συσχετίζουν τα βάρη των συντελεστών Fourier ανωτέρων επιπέδων με αυτά των κατώτερων επιπέδων. Έτσι προκύπτουν ανισότητες που συγκρίνουν την αρμονική επέκταση μιας συνάρτησης με αυτές των παραγώγων της. Το πρώτο λήμμα είναι άμεση εφαρμογή του γεγονότος ότι η $\kappa(x, \cdot)$ είναι υποκανονική. Ουσιαστικά,

φράσσει τα βάρη Fourier του πρώτου επιπέδου από μια συνάρτηση των βαρών του μηδενικού επιπέδου.

Λήμμα 2.6.2 (ανισότητα επιπέδου-1). *Υπάρχει μια σταθερά L τέτοια ώστε να ισχύει το εξής: Έστω $g : [-1, 1]^n \rightarrow [0, 1]$ η αρμονική επέκταση μιας Boolean συνάρτησης, και έστω $x \in (-1, 1)^n$ με $|x_i| = t$ για κάθε i . Τότε,*

$$(2.6.7) \quad \|\nabla g(x)\|_2^2 \leq \frac{L}{(1-t)^4} g(x)^2 \log\left(\frac{e}{g(x)}\right).$$

Απόδειξη. Το λήμμα είναι όμοιο με την [50, Πρόταση 2.2], αν την εφαρμόσουμε σε ένα μεροληπτικό μέτρο γινόμενο αντί για το ομοιόμορφο μέτρο πιθανότητας στο $\{-1, 1\}^n$. Για λόγους πληρότητας, παρουσιάζουμε εδώ μια γενική απόδειξη, η οποία δεν χρησιμοποιεί απευθείας την υπερσυσταλτικότητα.

Χρησιμοποιώντας την (2.6.3), γράφουμε την αρμονική επέκταση $g(x)$ ως

$$g(x) = \sum_y \kappa(x, y) g(y),$$

όπου $\kappa(x, y) = \prod_i (1 + x_i y_i) / 2$. Αφού η παραγωγήιση αντιμετατίθεται με την αρμονική επέκταση, έχουμε ότι

$$\begin{aligned} \partial_i g(x) &= \frac{\partial}{\partial x_i} \sum_y \kappa(x, y) g(y) = \sum_y \kappa(x, y) y_i \frac{g(y)}{1 + x_i y_i} \\ &= \frac{1}{1-t^2} \sum_y \kappa(x, y) y_i g(y) (1 - x_i y_i) = \frac{1}{1-t^2} \nu(A) \left(\frac{\int_A y_i d\nu}{\nu(A)} - x_i \right), \end{aligned}$$

όπου ν είναι το μέτρο με πυκνότητα $\kappa(x, \cdot)$, και $A = \text{supp}(g)$. Με αυτόν τον συμβολισμό,

$$g(x) = \nu(A) \quad \text{και} \quad \partial_i g(x) = \int_A \frac{y_i}{1 + x_i y_i} d\nu.$$

Έστω $\{\alpha_i\}_{i=1}^n$ βάρη τέτοια ώστε $\sum_{i=1}^n \alpha_i^2 = 1$, και έστω $h : \{-1, 1\}^n \rightarrow \mathbb{R}$ η συνάρτηση

$$h(y) = \sum_{i=1}^n \alpha_i \frac{y_i}{1 + x_i y_i}.$$

Έστω $Y \in \{-1, 1\}^n$ τυχαίο διάνυσμα με κατανομή το ν . Θυμηθείτε ότι η υποκανονική νόρμα μιας τυχαίας μεταβλητής $R \in \mathbb{R}$ δίνεται από την $\|R\|_{\psi_2} = \inf \left\{ s > 0 \mid \mathbb{E} \exp(R^2/s^2) \leq 2 \right\}$, ενώ η υποκανονική νόρμα ενός τυχαίου διανύσματος $R \in \mathbb{R}^n$ δίνεται από την $\|R\|_{\psi_2} = \sup_{r \in S^{n-1}} \|\langle R, r \rangle\|_{\psi_2}$ (βλ. [52, Ενότητες 2.4 και 2.5]). Η τυχαία μεταβλητή $\frac{Y_i}{1+x_i Y_i}$ έχει μέτρο φραγμένο από $(1-t)^{-1}$, επομένως η υποκανονική της νόρμα φράσσεται από $C(1-t)^{-1}$ για κάποια απόλυτη σταθερά C . Από το [52, Λήμμα 3.4.2], ένα τυχαίο διάνυσμα Z με ανεξάρτητες υποκανονικές συντεταγμένες που έχουν μέση τιμή 0, είναι επίσης υποκανονικό, με $\|Z\|_{\psi_2} \leq C \max_i \|Z_i\|_{\psi_2}$. Συνεπώς, το τυχαίο διάνυσμα $\left(\frac{Y_1}{1+x_1 Y_1}, \dots, \frac{Y_n}{1+x_n Y_n} \right)$ έχει υποκανονική νόρμα φραγμένη από $C(1-t)^{-1}$, το

οποίο σημαίνει ότι για κάθε $s > 0$

$$\mathbb{P}[|h(Y)| \geq s] \leq 2 \exp(-Cs^2(1-t)^2).$$

Έστω $s_0 \geq \frac{1}{\sqrt{C}}$. Τότε,

$$\begin{aligned} \int_A |h| d\nu &= \int_0^\infty \nu(\{|h| \geq t\} \cap A) dt \\ &\leq \int_0^\infty \min(\nu(A), 2e^{-Cs^2(1-t)^2}) ds \\ &\leq \nu(A) s_0 + \frac{1}{\sqrt{C}(1-t)^2} \int_{s_0}^\infty 2sC(1-t)^2 e^{-Cs^2(1-t)^2} ds \\ &\leq \nu(A) s_0 + \frac{1}{\sqrt{C}(1-t)^2} e^{-Cs_0^2(1-t)^2}. \end{aligned}$$

Παίρνοντας $s_0 = \sqrt{\frac{1}{C(1-t)^2} \log \frac{e}{\nu(A)}} \geq \frac{1}{\sqrt{C}}$ έχουμε ότι

$$\begin{aligned} \int_A |h| d\nu &\leq \nu(A) \frac{1}{\sqrt{C}(1-t)} \sqrt{\log \left(\frac{e}{\nu(A)} \right)} + \frac{1}{\sqrt{C}(1-t)^2} e^{-\log \left(\frac{e}{\nu(A)} \right)} \\ &\leq \frac{L}{(1-t)^2} \nu(A) \sqrt{\log \left(\frac{e}{\nu(A)} \right)} \end{aligned}$$

για κάποια σταθερά $L > 0$. Ειδικότερα, ισχύει και η

$$(2.6.8) \quad \int_A h d\nu \leq \frac{L}{(1-t)^2} \nu(A) \sqrt{\log \left(\frac{e}{\nu(A)} \right)}.$$

Τώρα επιλέγουμε $\alpha_i = \partial_i g(x) \left(\sum_{i=1}^n \partial_i g(x)^2 \right)^{-1/2}$, και παρατηρούμε ότι

$$\begin{aligned} \left(\int_A h d\nu \right)^2 &= \left(\int_A \sum_{i=1}^n \frac{\partial_i g(x)}{\sqrt{\sum_{i=1}^n \partial_i g(x)^2}} \frac{y_i}{1+x_i y_i} d\nu \right)^2 \\ &= \frac{1}{\sum_{i=1}^n \partial_i g(x)^2} \left(\sum_{i=1}^n \partial_i g(x) \int_A \frac{y_i}{1+x_i y_i} d\nu \right)^2 \\ &= \frac{1}{\sum_{i=1}^n \partial_i g(x)^2} \left(\sum_{i=1}^n \partial_i g(x)^2 d\nu \right)^2 = \sum_{i=1}^n \partial_i g(x)^2 = \|\nabla g\|_2^2. \end{aligned}$$

Συνδυάζοντας αυτήν την ισότητα με την (2.6.8) παίρνουμε το ζητούμενο. $\square$

Το δεύτερο λήμμα, που η αρχική ομοιόμορφη εκδοχή του οφείλεται στον Talagrand [50], ουσιαστικά φράσσει τα βάρη Fourier του δευτέρου επιπέδου από αυτά του πρώτου επιπέδου.

Είναι όμοιο με το [36, Λήμμα 6], αλλά για πραγματικές συναρτήσεις.

Λήμμα 2.6.3 (ανισότητα επιπέδου-2). *Υπάρχει μια συνεχής συνάρτηση $C : [0, 1) \rightarrow [0, \infty)$ τέτοια ώστε να ισχύει το εξής: Έστω $g : [-1, 1]^n \rightarrow [-1, 1]$ η αρμονική επέκταση μιας μονότονης συνάρτησης, και έστω $x \in (-1, 1)^n$ με $|x_i| = t$ για κάθε i . Τότε,*

$$(2.6.9) \quad \|\nabla^2 g(x)\|_{\text{HS}}^2 \leq C(t) \|\nabla g(x)\|_2^2 \cdot \log \left(\frac{C(t)}{\|\nabla g(x)\|_2^2} \right),$$

όπου $\nabla^2 g$ είναι η Εσσιανή $(\partial_i \partial_j g)_{i,j=1}^n$ της g , και $\|X\|_{\text{HS}} = \sqrt{\sum_{i,j} X_{ij}^2}$ είναι η νόρμα Hilbert-Schmidt ενός πίνακα.

Απόδειξη. Χρησιμοποιώντας την Πρόταση 2.6.1 για το $S = \{i, j\}$, το γεγονός ότι η g είναι αρμονική και επιπλέον $\partial_{ii} g = 0$ για κάθε $i = 1, \dots, n$, καθώς και το γεγονός ότι $|x_i| = t$, έχουμε

$$(2.6.10) \quad \begin{aligned} \|\nabla^2 g(x)\|_{\text{HS}}^2 &= \sum_{1 \leq i \neq j \leq n} (\partial_i \partial_j g(x))^2 \\ &= \sum_{1 \leq i \neq j \leq n} \left(\frac{16}{1-t^2} \widehat{g}_p(\{i, j\}) \right)^2 \\ &\leq 2C(t) \sum_{S \subseteq [n], |S|=2} \widehat{g}_p(S)^2. \end{aligned}$$

Το επόμενο λήμμα, το οποίο φράσσει το άθροισμα των τετραγώνων p -μεροληπτικών συντελεστών Fourier, προκύπτει άμεσα από το [36, Λήμμα 6].

Λήμμα 2.6.4. *Έστω $0 \leq t < 1$ και $p \in (0, 1)^n$ τέτοια ώστε $p_i \in \left\{ \frac{1+t}{2}, \frac{1-t}{2} \right\}$ για κάθε i . Για κάθε συνάρτηση $g : \{-1, 1\}^n \rightarrow \{-1, 1\}$, ορίζουμε*

$$\mathcal{W}(f) = p(1-p) \sum_{i=1}^n I_i^p(g)^2.$$

Υπάρχει συνάρτηση $C(t)$ τέτοια ώστε, για κάθε g ,

$$(2.6.11) \quad \sum_{S \subseteq [n], |S|=2} \widehat{g}_p(S)^2 \leq C(t) \mathcal{W}(g) \cdot \log \left(\frac{2}{\mathcal{W}(g)} \right).$$

Συνδυάζοντας τις (2.6.10) και (2.6.11), παίρνουμε

$$(2.6.12) \quad \|\nabla^2 g(x)\|_{\text{HS}}^2 \leq C(t) \mathcal{W}(g) \cdot \log \left(\frac{2}{\mathcal{W}(g)} \right).$$

Από την (2.6.6), για μονότονες συναρτήσεις η επιρροή του i -οστού ψηφίου ισούται με

$$I_i^p(g) = 2\sqrt{p_i} \sqrt{1-p_i} \widehat{g}_p(\{i\}),$$

άρα

$$\mathcal{W}(f) = p(1-p) \sum_{i=1}^n I_i^p(g)^2 = 4p^2(1-p)^2 \sum_{i=1}^n \widehat{g}_p(\{i\})^2.$$

Από την άλλη πλευρά, χρησιμοποιώντας την Πρόταση 2.6.1 για το $S = \{i\}$ βλέπουμε ότι

$$\begin{aligned} \|\nabla g(x)\|_2^2 &= \sum_{i=1}^n (\partial_i g(x))^2 \\ &= \frac{16}{1-t^2} \sum_{i=1}^n \widehat{g}_p(\{i\})^2 \\ &= \frac{4}{(1-t^2)p^2(1-p)^2} \mathcal{W}(g) := C'(t) \mathcal{W}(g). \end{aligned}$$

Εισάγοντας αυτήν την ισότητα στην (2.6.12), βλέπουμε ότι υπάρχει κάποια $C(t)$ τέτοια ώστε

$$\|\nabla^2 g(x)\|_{\text{HS}}^2 \leq C(t) \|\nabla g(x)\|_2^2 \log \left(\frac{C(t)}{\|\nabla g(x)\|_2^2} \right).$$

□

Σημείωση 2.6.5. Στα δύο αυτά λήμματα, η απαίτηση ότι $|x_i| = t$ για κάθε i δεν είναι κρίσιμη, και μπορεί να αντικατασταθεί από την $|x_i| \leq t$ για κάθε i .

ΚΕΦΑΛΑΙΟ 3

Εργαλεία από τον Στοχαστικό Λογισμό

3.1 Δεσμευμένη μέση τιμή και martingales

Ορισμός 3.1.1 (δεσμευμένη μέση τιμή). Έστω $(\Omega, \mathcal{F}, P)$ χώρος πιθανότητας, $\mathcal{G} \subseteq \mathcal{F}$ μια σ -άλγεβρα και $X : \Omega \rightarrow \mathbb{R}$ μια τυχαία μεταβλητή με $\mathbb{E}[|X|] < \infty$. Καλούμε **δεσμευμένη μέση τιμή της X ως προς την σ -άλγεβρα $\mathcal{G}$** κάθε τυχαία μεταβλητή $Y : \Omega \rightarrow \mathbb{R}$ που είναι $\mathcal{G}$ -μετρήσιμη και επίσης, για κάθε $A \in \mathcal{G}$ ικανοποιεί την

$$\int_A X dP = \int_A Y dP.$$

Προφανώς, η παραπάνω έννοια της δεσμευμένης μέσης τιμής δεν θα είχε νόημα αν δεν υπήρχε τέτοια τυχαία μεταβλητή. Το παρακάτω θεώρημα εξασφαλίζει την ύπαρξη της δεσμευμένης μέσης τιμής μιας τυχαίας μεταβλητής, η οποία για καλή μας τύχη είναι και μοναδική κατά μία έννοια.

Θεώρημα 3.1.2 (ύπαρξη και μοναδικότητα). Έστω $(\Omega, \mathcal{F}, P)$ χώρος πιθανότητας, $X : \Omega \rightarrow \mathbb{R}$ μια τυχαία μεταβλητή με $\mathbb{E}[|X|] < \infty$ και $\mathcal{G} \subseteq \mathcal{F}$ μια σ -άλγεβρα. Υπάρχει τυχαία μεταβλητή $Y : \Omega \rightarrow \mathbb{R}$ που είναι δεσμευμένη μέση τιμή της X ως προς την σ -άλγεβρα $\mathcal{G}$, και αν $Y' : \Omega \rightarrow \mathbb{R}$ είναι μια άλλη δεσμευμένη μέση τιμή της X ως προς την $\mathcal{G}$, τότε $P(Y = Y') = 1$.

Λόγω της μοναδικότητας της δεσμευμένης μέσης τιμής, μπορούμε τώρα να θεωρήσουμε τον συμβολισμό $\mathbb{E}[X | \mathcal{G}]$ για τη δεσμευμένη μέση τιμή της τυχαίας μεταβλητής X ως προς την σ -άλγεβρα $\mathcal{G}$. Επίσης, αν $X, Y : \Omega \rightarrow \mathbb{R}$ είναι δύο τυχαίες μεταβλητές στον χώρο πιθανότητας $(\Omega, \mathcal{F}, P)$ με $\sigma(Y) \subseteq \mathcal{F}$, τότε θα χρησιμοποιούμε τον συμβολισμό

$$\mathbb{E}[X | Y] := \mathbb{E}[X | \sigma(Y)]$$

όπου $\sigma(Y)$ είναι η σ -άλγεβρα που παράγεται από την τυχαία μεταβλητή Y .

Στην επόμενη πρόταση συνοψίζουμε κάποιες βασικές ιδιότητες της δεσμευμένης μέσης τιμής μιας τυχαίας μεταβλητής.

Πρόταση 3.1.3 (ιδιότητες δεσμευμένης μέσης τιμής). Έστω $(\Omega, \mathcal{F}, P)$ χώρος πιθανότητας, $X, Y : \Omega \rightarrow \mathbb{R}$ τυχαίες μεταβλητές με $\mathbb{E}[|X|] < \infty, \mathbb{E}[|Y|] < \infty$ και έστω $\mathcal{G} \subseteq \mathcal{F}$ μια σ -άλγεβρα. Τότε:

(α) **Θεώρημα διπλής μέσης τιμής:** Ισχύει ότι

$$\mathbb{E}[\mathbb{E}[X | \mathcal{G}]] = \mathbb{E}[X].$$

(β) Αν $a, b \in \mathbb{R}$ τότε

$$\mathbb{E}[aX + bY | \mathcal{G}] = a\mathbb{E}[X | \mathcal{G}] + b\mathbb{E}[Y | \mathcal{G}].$$

(γ) Αν $X \leq Y$, τότε

$$\mathbb{E}[X | \mathcal{G}] \leq \mathbb{E}[Y | \mathcal{G}].$$

(δ) Αν n X είναι $\mathcal{G}$ -μετρήσιμη, τότε

$$\mathbb{E}[X | \mathcal{G}] = X.$$

(στ) Αν $\mathcal{G}_1 \subseteq \mathcal{G}_2 \subseteq \mathcal{F}$ είναι σ -άλγεβρες, τότε

$$\mathbb{E}[\mathbb{E}[X | \mathcal{G}_1] | \mathcal{G}_2] = \mathbb{E}[X | \mathcal{G}_1]$$

και

$$\mathbb{E}[\mathbb{E}[X | \mathcal{G}_2] | \mathcal{G}_1] = \mathbb{E}[X | \mathcal{G}_1].$$

(ζ) Αν n Y είναι $\mathcal{G}$ -μετρήσιμη και $\mathbb{E}[|XY|] < \infty$, τότε

$$\mathbb{E}[XY | \mathcal{G}] = Y\mathbb{E}[X | \mathcal{G}].$$

(η) **Ανισότητα Jensen:** Αν f είναι μια κυρτή συνάρτηση σε ένα διάστημα $I \subseteq \mathbb{R}$ με $P(X \in I) = 1$ και $\mathbb{E}[|f(X)|] < \infty$, τότε ισχύει ότι

$$f(\mathbb{E}[X | \mathcal{G}]) \leq \mathbb{E}[f(X) | \mathcal{G}].$$

Ας κάνουμε τώρα ένα βήμα προς την έννοια των στοχαστικών διαδικασιών που αποτελούν το κύριο εργαλείο αυτής της διπλωματικής εργασίας.

Ορισμός 3.1.4 (διήθηση). Έστω $(\Omega, \mathcal{F}, P)$ ένας χώρος πιθανότητας. **Διήθηση** ως προς τον χώρο αυτόν θα καλούμε κάθε οικογένεια σ -αλγεβρών $(\mathcal{F}_t)_{t \geq 0}$ με την ιδιότητα ότι για κάθε $0 \leq t \leq s$ ισχύει ότι $\mathcal{F}_t \subseteq \mathcal{F}_s \subseteq \mathcal{F}$.

Ορισμός 3.1.5 (διαδικασία–προσαρμοσμένη διαδικασία). **Διαδικασία** σε έναν χώρο πιθανότητας $(\Omega, \mathcal{F}, P)$ καλούμε κάθε οικογένεια $(X_t)_{t \in I}$ τυχαίων μεταβλητών στον χώρο πιθανότητας

αυτόν, όπου I αυθαίρετο σύνολο δεικτών. Επίσης, στην περίπτωση όπου $I = [0, +\infty)$ και $(\mathcal{F}_t)_{t \geq 0}$ είναι μια διήθηση, τότε η διαδικασία $(X_t)_{t \geq 0}$ καλείται **προσαρμοσμένη** αν κάθε τυχαία μεταβλητή X_t είναι $\mathcal{F}_t$ -μετρήσιμη.

Αν $X = (X_t)_{t \geq 0}$ είναι μια διαδικασία και $\omega \in \Omega$ σταθερό, τότε η συνάρτηση $t \mapsto X_t(\omega)$ καλείται **μονοπάτι** της διαδικασίας X .

Ορισμός 3.1.6 (martingale–submartingale–supermartingale). Έστω $(\Omega, \mathcal{F}, P)$ είναι χώρος πιθανότητας και $(\mathcal{F}_t)_{t \geq 0}$ μια διήθηση στον χώρο αυτόν. Μια διαδικασία $X = (X_t)_{t \geq 0}$ καλείται **martingale** ως προς αυτή τη διήθηση αν ικανοποιεί τις εξής τρεις συνθήκες:

- Η X είναι προσαρμοσμένη ως προς τη διήθηση.
- Για κάθε $t \geq 0$ ισχύει ότι $\mathbb{E}[|X_t|] < \infty$.
- Για κάθε $0 \leq t \leq s$ ισχύει ότι $\mathbb{E}[X_s | \mathcal{F}_t] = X_t$.

Αν οι δύο πρώτες συνθήκες ικανοποιούνται και, αντί της τρίτης συνθήκης, για κάθε $0 \leq t \leq s$ ισχύει ότι $\mathbb{E}[X_s | \mathcal{F}_t] \geq X_t$ τότε η διαδικασία X καλείται **submartingale**, ενώ αν αντί για την τρίτη συνθήκη έχουμε ότι για κάθε $0 \leq t \leq s$ ισχύει $\mathbb{E}[X_s | \mathcal{F}_t] \leq X_t$, τότε η διαδικασία X καλείται **supermartingale**.

Μία απο τις βασικές ιδιότητες των martingales που χρησιμοποιούμε παρακάτω, είναι η εξής.

Πρόταση 3.1.7. Έστω $(\Omega, \mathcal{F}, P)$ χώρος πιθανότητας, $(\mathcal{F}_t)_{t \geq 0}$ διήθηση σε αυτόν τον χώρο πιθανότητας, και $(X_t)_{t \geq 0}$ martingale ως προς αυτή τη διήθηση. Αν $f : I \rightarrow \mathbb{R}$ είναι μια κυρτή συνάρτηση για την οποία ισχύει ότι $P(X_t \in I) = 1$ και $\mathbb{E}[|f(X_t)|] < \infty$ για κάθε $t \geq 0$, τότε η $f(X_t)$ είναι submartingale.

Μια άλλη βασική έννοια που θα χρησιμοποιήσουμε αρκετές φορές παρακάτω, είναι η έννοια των χρόνων διακοπής.

Ορισμός 3.1.8 (χρόνοι διακοπής). Έστω $(\Omega, \mathcal{F}, P)$ χώρος πιθανότητας και $(\mathcal{F}_t)_{t \geq 0}$ διήθηση σε αυτόν. Κάθε συνάρτηση $\tau : \Omega \rightarrow [0, \infty]$ με την ιδιότητα ότι $\{\tau \leq t\} \in \mathcal{F}_t$ για κάθε $t \geq 0$ καλείται **χρόνος διακοπής**.

Τώρα, μέσω των χρόνων διακοπής μπορούμε, ξεκινώντας από μια στοχαστική διαδικασία, να κατασκευάσουμε άλλες διαδικασίες οι οποίες ενδέχεται να έχουν πολύ καλές ιδιότητες, όπως την ιδιότητα του να είναι martingales. Συγκεκριμένα, αν τ είναι ένας χρόνος διακοπής και $X = (X_t)_{t \geq 0}$ είναι μια διαδικασία στον χώρο πιθανότητας $(\Omega, \mathcal{F}, P)$, τότε ορίζουμε τη διαδικασία $X^\tau := (X_{\tau \wedge t})_{t \geq 0}$ με

$$(3.1.1) \quad X^\tau(\omega) = X_{\tau(\omega) \wedge t}(\omega) = \begin{cases} X_t(\omega), & t \leq \tau(\omega) \\ X_{\tau(\omega)}, & \tau(\omega) < t \end{cases}$$

για κάθε $t \geq 0$ και $\omega \in \Omega$. Με άλλα λόγια, για κάθε $\omega \in \Omega$, αν $r = \tau(\omega)$ τότε η X^τ ακολουθεί την X στις τιμές της μέχρι την X_r και έπειτα σταθεροποιείται σε αυτήν.

Πρόταση 3.1.9. Έστω $(\Omega, \mathcal{F}, P)$ χώρος πιθανότητας, $(\mathcal{F}_t)_{t \geq 0}$ διήθηση σε αυτόν τον χώρο, και $\tau : \Omega \rightarrow [0, \infty]$ χρόνος διακοπής. Αν n $X = (X_t)_{t \geq 0}$ είναι *martingale* τότε και η διαδικασία X^τ είναι *martingale*.

Ένα από τα βασικά θεωρήματα του Στοχαστικού Λογισμού, το οποίο θα χρησιμοποιηθεί παρακάτω, είναι το θεώρημα επιλεκτικής διακοπής.

Θεώρημα 3.1.10 (θεώρημα επιλεκτικής διακοπής). Έστω $(\Omega, \mathcal{F}, P)$ χώρος πιθανότητας, $(\mathcal{F}_t)_{t \geq 0}$ διήθηση σε αυτόν, και $X = (X_t)_{t \geq 0}$ *martingale* ως προς αυτή τη διήθηση. Αν $\tau : \Omega \rightarrow [0, \infty]$ είναι φραγμένος χρόνος διακοπής, τότε

$$\mathbb{E}[X_\tau] = \mathbb{E}[X_0].$$

Θα χρησιμοποιήσουμε επίσης την κλασική ανισότητα του Doob.

Θεώρημα 3.1.11 (ανισότητα Doob για submartingales). Έστω $(\Omega, \mathcal{F}, P)$ χώρος πιθανότητας, $(\mathcal{F}_t)_{t \geq 0}$ διήθηση σε αυτόν, και $X = (X_t)_{t \geq 0}$ *submartingale* ως προς αυτή τη διήθηση. Για κάθε $t > 0$ και $\lambda > 0$ ισχύει ότι

$$P\left(\sup_{s \in [0, t]} X_s \geq \lambda\right) \leq \frac{1}{\lambda} \mathbb{E}[X_t^+].$$

3.2 Στοχαστικές διαδικασίες και τετραγωνική κύμανση

Μια στοχαστική διαδικασία Poisson (N_t) με ένταση $\lambda(t)$ είναι μια διαδικασία με ακέραιες τιμές για την οποία ισχύει ότι $N_0 = 0$ και για κάθε $0 \leq a < b$ η διαφορά $N_b - N_a$ έχει την κατανομή μιας τυχαίας μεταβλητής Poisson με ένταση $\int_a^b \lambda(t) dt$. Αν $\int_a^b \lambda(t) dt < \infty$ για κάθε $0 \leq a \leq b$, τότε τα μονοπάτια της στοχαστικής διαδικασίας Poisson (N_t) είναι δεξιά συνεχή σχεδόν βεβαίως. Το (τυχαίο) σύνολο των χρονικών στιγμών κατά τις οποίες το μονοπάτι είναι ασυνεχές το συμβολίζουμε με $\text{Jump}(N_t)$.

Αν η $\lambda(t)$ είναι τέτοια ώστε $\int_a^b \lambda(t) dt < \infty$ για κάθε $0 \leq a \leq b$ και (N_t) είναι μια στοχαστική διαδικασία Poisson με ένταση $\lambda(t)$, τότε αποδεικνύεται ότι $\text{Jump}(N_t) = \{t_1, t_2, \dots\}$, δηλαδή το σύνολο των σημείων ασυνέχειας της N_t είναι σχεδόν βεβαίως το πολύ αριθμήσιμο. Λέμε ότι μια στοχαστική διαδικασία (X_t) είναι τμηματικά ομαλή εάν είναι δεξιά συνεχής και ομαλή στο διάστημα $[t_i, t_{i+1})$ για κάθε $i = 1, 2, \dots$. Αυτός ο ορισμός μπορεί να επεκταθεί και στην περίπτωση όπου η $\lambda(t)$ είναι τέτοια ώστε $\int_0^b \lambda(t) dt = \infty$ αλλά $\int_a^b \lambda(t) dt < \infty$ για κάθε $0 < a < b$ (για παράδειγμα, αυτό μπορεί να συμβεί αν $\lambda(t) = 1/t$): σε αυτήν την περίπτωση το $\text{Jump}(N_t)$ έχει μοναδικό σημείο συσσώρευσης το 0 και τα διαστήματα μεταξύ των διαδοχικών χρονικών στιγμών των αλμάτων της N_t είναι καλά ορισμένα.

Μια σημαντική έννοια για την ανάλυση των στοχαστικών διαδικασιών είναι η τετραγωνική κύμανση. Η τετραγωνική κύμανση μιας στοχαστικής διαδικασίας (X_t) , την οποία θα συμβολίζουμε με $[X]_t$, είναι ένα μέτρο των αυξομειώσεων της (X_t) και ορίζεται από την

$$[X]_t := \lim_{\|P\| \rightarrow 0} \sum_{k=1}^n (X_{t_k} - X_{t_{k-1}})^2$$

αν βέβαια το όριο υπάρχει, όπου P είναι μια διαμέριση n σημείων του διαστήματος $[0, t]$ και ο συμβολισμός $\lim_{\|P\| \rightarrow 0}$ υποδεικνύει ότι το πλάτος $\|P\|$ της διαμέρισης P τείνει στο 0. Αξίζει να σημειωθεί ότι, εν γένει, δεν ισχύει ότι κάθε στοχαστική διαδικασία έχει πεπερασμένη τετραγωνική κύμανση. Αν όμως η (X_t) είναι τμηματικά ομαλή διαδικασία τότε αυτό ισχύει. Αποδεικνύεται μάλιστα ότι αν η (X_t) είναι τμηματικά ομαλή διαδικασία, τότε

$$(3.2.1) \quad [X]_t = \sum_{s \in \text{Jump}(X_t) \cap [0, t]} (\Delta X_s)^2$$

όπου $\Delta(X_s) = \lim_{\varepsilon \rightarrow 0^+} (X_{s+\varepsilon} - X_{s-\varepsilon})$ είναι το μέγεθος του άλματος την χρονική στιγμή s .

Η τετραγωνική κύμανση είναι ιδιαίτερα χρήσιμη στην περίπτωση των martingales, λόγω της σχέσης τους με τη διασπορά: Εάν η X_t είναι martingale, τότε

$$(3.2.2) \quad \text{Var}(X_t) = \mathbb{E} [[X]_t].$$

Λήμμα 3.2.1. Έστω $a \in [0, 1)$ και $c_1 \in [0, 1/3)$. Τότε, υπάρχει μια σταθερά c_2 , που εξαρτάται μόνο από τα a και c_1 , τέτοια ώστε να ισχύει η ακόλουθη πρόταση. Έστω $(X_t)_{t=0}^1$ ένα δεξιά συνεχές martingale που ικανοποιεί τις

$$(\alpha) \quad X_0 \in [-a, a].$$

$$(\beta) \quad X_1 \in \{-1, 1\}.$$

$$(\gamma) \quad X_t \in [-1, 1] \text{ για κάθε } t \in [0, 1].$$

Τότε,

$$\mathbb{P} [[X]_1 \geq c_1(1-a)^2] \geq c_2(1-a)^2.$$

Μπορούμε επίσης να επιλέξουμε τη σταθερά c_2 μεγαλύτερη από $\frac{c_1}{2(2-\log(1-a)^2-\log c_1)}$.

Απόδειξη. Έστω s ένας ακέραιος, τον οποίο θα επιλέξουμε κατάλληλα αργότερα. Θέτουμε $c_2 = \frac{c_1}{2(2-\log(1-a)^2-\log c_1)}$ και υποθέτουμε προς άτοπο ότι

$$\mathbb{P} [[X]_1 \geq c_1(1-a)^2] < c_2(1-a)^2.$$

Τότε,

$$\begin{aligned} \mathbb{E} [[X]_1] &\leq \mathbb{E} [[X]_1 \mathbf{1}_{\{[X]_1 \leq c_1(1-a)^2\}}] + \mathbb{E} [[X]_1 \mathbf{1}_{\{[X]_1 \in [c_1(1-a)^2, s]\}}] + \mathbb{E} [[X]_1 \mathbf{1}_{\{[X]_1 \geq s\}}] \\ &\leq c_1(1-a)^2 + c_2(1-a)^2 s + \int_s^\infty \mathbb{P} [[X]_t \geq t] dt. \end{aligned}$$

Για να φράξουμε τον τρίτο όρο, χρησιμοποιούμε το γεγονός ότι η τετραγωνική κύμανση είναι υπο-εκθετική, ως εξής. Θέτουμε $\tau_0 = 0$ και για κάθε $k \in \mathbb{N}$ ορίζουμε $\tau_k := \inf\{t > 0 \mid [X]_t \geq 2k\}$. Υπενθυμίζουμε ότι το infimum του κενού συνόλου είναι ∞ . Αφού η τετραγωνική κύμανση είναι πεπερασμένη σχεδόν βεβαίως, η τυχαία μεταβλητή $k^* = \sup\{k \in \mathbb{N} \cup \{0\} \mid \tau_k < \infty\}$ είναι

πεπερασμένη σχεδόν βεβαίως. Για κάθε μετρήσιμο χρόνο διακοπής $\tau \in [0, 1]$ έχουμε

$$\mathbb{E} [[X]_1 - [X]_\tau] \leq 1.$$

Εφαρμόζοντας την ανισότητα Markov παίρνουμε

$$\mathbb{P} [k^* \geq k + 1 \mid k^* \geq k] = \mathbb{P} [[X]_1 - [X]_{\tau_k} \geq 2] \leq \frac{\mathbb{E} [[X]_1 - [X]_\tau]}{2} \leq \frac{1}{2}.$$

Με επαγωγή βλέπουμε ότι

$$\mathbb{P} [[X]_1 \geq 2k] = \mathbb{P} [k^* \geq k] \leq 2^{-k}.$$

Συνεπώς, αν ο $s \geq 2$ είναι άρτιος, έχουμε ότι

$$\int_s^\infty \mathbb{P} [[X]_t \geq t] dt \leq 2 \sum_{k=s/2}^\infty 2^{-k} = 4 \cdot 2^{-s/2}.$$

Μπορούμε λοιπόν να φράξουμε τη μέση τιμή της τετραγωνικής κύμανσης ως εξής:

$$\mathbb{E} [[X]_1] \leq c_1(1-a)^2 + c_2(1-a)^2 s + 4 \cdot 2^{-s/2}.$$

Επιλέγουμε τον s να είναι ο μικρότερος άρτιος ακέραιος που είναι μεγαλύτερος από $2(2 - \log_2((1-a)^2) - \log_2 c_1)$. Τότε,

$$\begin{aligned} \mathbb{E} [[X]_1] &\leq c_1(1-a)^2 + 4c_2(1-a)^2(2 - \log_2((1-a)^2) - \log_2 c_1) + c_1(1-a)^2 \\ &= (1-a)^2(2c_1 + 4c_2(2 - \log_2((1-a)^2) - \log_2 c_1)). \end{aligned}$$

Από την επιλογή του $c_2 = \frac{c_1}{2(2 - \log_2((1-a)^2) - \log_2 c_1)}$ παίρνουμε

$$\mathbb{E} [[X]_1] \leq 3c_1(1-a)^2.$$

Εφόσον $c_1 < 1/3$, αυτό έρχεται σε αντίφαση με το γεγονός ότι

$$\mathbb{E} [[X]_1] = \text{Var}(X_1) \geq (1-a)^2$$

και έτσι ολοκληρώνεται η απόδειξη. □

3.3 Ανελίξεις στον συνεχή κύβο και σκιαγράφηση της μεθόδου

Το ερώτημα από το οποίο θα ξεκινήσουμε και θα μας οδηγήσει στην ανισότητα Talagrand, είναι με ποιον τρόπο μπορούμε να δημιουργήσουμε μια στοχαστική διαδικασία στον συνεχή κύβο $[-1, 1]^n$, η οποία τη χρονική στιγμή $t = 0$ βρίσκεται στη θέση 0 και την χρονική στιγμή $t = 1$ η θέση της είναι ομοιόμορφα κατανεμημένη στον διακριτό κύβο $\{-1, 1\}^n$.

Για διευκόλυνση της θέσης μας αυτή τη στιγμή, αναζητούμε μια διαδικασία που οι συντεταγμένες της είναι ανεξάρτητες και ισόνομες. Με άλλα λόγια, για κάθε t αναζητούμε

$B_t = (B_t^{(1)}, B_t^{(2)}, \dots, B_t^{(n)})$, όπου οι $B_t^{(i)}$ είναι ανεξάρτητες και ισόνομες με μια μονοδιάστατη τυχαία μεταβλητή X_t και θέλουμε να συμβαδίζουν με τις ιδιότητες που περιγράψαμε πιο πάνω. Έτσι, το πρόβλημα ανάγεται στην εύρεση μιας μονοδιάστατης τυχαίας μεταβλητής X_t με τις εξής ιδιότητες: $X_0 = 0$ και στη χρονική στιγμή $t = 1$ η X_1 ακολουθεί την κατανομή της ρίψης ενός δίκαιου νομίσματος, δηλαδή

$$\mathbb{P}(X_1 = 1) = \mathbb{P}(X_1 = -1) = \frac{1}{2}.$$

Το ερώτημα είναι: πώς συμπεριφέρεται η X_t στις ενδιάμεσες χρονικές στιγμές;

Γενικά, θα μπορούσε να γίνεται οτιδήποτε στο ενδιάμεσο: μπορούμε να την κάνουμε να κάθεται στο 0, δηλαδή να έχουμε $X_t = 0$ και να μην αλλάξει τίποτα μέχρι τη χρονική στιγμή $t = 1$ όπου θα εκτινάσσεται στο 1 ή -1 . Ένα άλλο παράδειγμα είναι να θεωρήσουμε ότι η X_t κινείται χαοτικά στο ενδιάμεσο χρονικό διάστημα.

Το πρόβλημα όμως στις παραπάνω δύο περιπτώσεις είναι ότι δεν μπορούμε να μελετήσουμε τη συμπεριφορά της X_t στις ενδιάμεσες χρονικές στιγμές. Ένας καλός τρόπος ώστε να μπορούμε να μελετάμε την X_t στις ενδιάμεσες χρονικές στιγμές t είναι να απαιτήσουμε να ισχύει ότι $|X_t| = t$ για κάθε t , το οποίο διαισθητικά σημαίνει ότι η $X = (X_t)$ επιλέγει μια κατεύθυνση κίνησης, προς το 1 ή το -1 , και τότε περπατάει για λίγο προς το 1 ή προς το -1 αντίστοιχα. Όμως, για κάθε δεδομένη χρονική στιγμή t , η X_t μπορεί να αλλάξει απόφαση ως προς την κατεύθυνση που θα κινηθεί και να πηδήξει από το t στο $-t$ ή αντίστροφα. Επομένως, όλη η τυχειότητα της X_t κρύβεται πίσω από το πρόσημό της και το πότε η X_t αποφασίζει να το αλλάξει.

Υπάρχουν όμως πολλοί τρόποι να αλλάξει αυτή η τυχειότητα και ένας από αυτούς τους τρόπους είναι η $X = (X_t)_{t \geq 0}$ να αποτελεί martingale. Με αυτόν τον τρόπο, εάν γνωρίζουμε πού βρίσκεται η X τη χρονική στιγμή s τότε για κάθε χρονική στιγμή $t > s$ η αναμενόμενη θέση της X θα είναι X_s , δηλαδή ουσιαστικά η X δεν κινείται στο πέρασμα του χρόνου.

Με όρους πιθανοτήτων ή, ειδικότερα, στοχαστικής ανάλυσης η X θα ικανοποιεί την

$$\mathbb{E}[X_t | X_s] = X_s \text{ για κάθε } t > s.$$

Φυσικά, δεν θα περιγράψαμε αυτή τη διαδικασία αν δεν γινόταν να επιτευχθεί. Μάλιστα, αυτή η στοχαστική διαδικασία, πέρα από ότι επιτυγχάνεται, είναι και μοναδική, λόγω του ότι η απαίτηση των δύο παραπάνω συνθηκών μας καθορίζει πλήρως την πιθανότητα η διαδικασία να κάνει άλμα σε οποιαδήποτε δοθείσα χρονική στιγμή: Υποθέτουμε ότι η X_t είναι θετική σε κάποια χρονική στιγμή t . Τότε, λόγω της ιδιότητας $|X_t| = t$ έχουμε ότι $X_t = t$. Τώρα, για κάθε χρονική στιγμή $t + \varepsilon$ έχουμε ότι είτε η X διατηρεί το ίδιο πρόσημο και άρα είναι ίση με $X_{t+\varepsilon} = t + \varepsilon$ ή αλλάζει πρόσημο και άρα $X_{t+\varepsilon} = -t - \varepsilon$. Εάν τώρα p είναι η πιθανότητα να αλλάξει πρόσημο η X την χρονική στιγμή $t + \varepsilon$, τότε η αναμενόμενη θέση της X την χρονική στιγμή $t + \varepsilon$ δεδομένου ότι $X_t = t$ δίνεται από τη σχέση

$$\mathbb{E}[X_{t+\varepsilon} | X_t = t] = (1 - p)(t + \varepsilon) - p(t + \varepsilon).$$

Αφού όμως η X είναι martingale, έπεται ότι

$$\mathbb{E}[X_{t+\varepsilon} | X_t = t] = t$$

και τελικά ότι

$$t = (1 - 2p)(t + \varepsilon)$$

οπότε λύνοντας ως προς p έχουμε ότι

$$p = \mathbb{P}(\operatorname{sgn}(X_t) \neq \operatorname{sgn}(X_{t+\varepsilon}) | X_t) = \frac{\varepsilon}{2(t + \varepsilon)}.$$

Αυτό δείχνει ότι η πιθανότητα άλματος της X σε ένα χρονικό διάστημα μικρού μήκους ε είναι $\frac{\varepsilon}{2(t+\varepsilon)}$ και άρα προσδιορίζεται πλήρως.

Θεώρημα 3.3.1. Υπάρχει μοναδική στοχαστική διαδικασία $X = (X_t)_{t \geq 0}$ τέτοια ώστε:

- (i) $|X_t| = t$ για κάθε $t \geq 0$.
- (ii) $H X$ είναι martingale.

Θεωρούμε τώρα τυχούσα συνάρτηση $f : \{-1, 1\}^n \rightarrow \mathbb{R}$ και θεωρώντας τη διαδικασία $(B_t)_{t \in [0,1]}$ που κατασκευάσαμε παραπάνω, μπορούμε να ορίσουμε τη νέα διαδικασία $f_t := f(B_t)$. Αφού $B_1 \sim \text{Uniform}(\{-1, 1\}^n)$ ένας φυσιολογικός τρόπος να υπολογίσουμε τη μέση τιμή της f είναι μέσω της σχέσης

$$\mathbb{E}[f] = \mathbb{E}[f(B_1)]$$

όπου το δεξί μέλος εξαρτάται από της τυχαιότητα της διαδικασίας B_t . Με όμοιο τρόπο, ο πιο φυσιολογικός τρόπος για να υπολογίσουμε τη διασπορά της f είναι μέσω της

$$\operatorname{Var}(f) = \operatorname{Var}(f(B_1))$$

όπου πάλι το δεξί μέλος εξαρτάται από την τυχαιότητα της διαδικασίας (B_t) . Τελείως ανάλογα, η i -οστή επιρροή της f ισούται με

$$I_i(f) = \mathbb{E}[\partial_i f(B_1)^2]$$

όπου $\partial_i f$ είναι η i -οστή διακριτή παράγωγος της f .

Περνάμε τώρα μια σε μια πιο τεχνική περιγραφή των επιχειρημάτων των Eldan και Gross, τα οποία θα παρουσιάσουμε λεπτομερώς στα επόμενα δύο κεφάλαια. Όπως είπαμε, οι αποδείξεις των αποτελεσμάτων βασίζονται στην κατασκευή ενός martingale

$$B_t = (B_t^{(1)}, \dots, B_t^{(n)}) \in \mathbb{R}^n$$

που ικανοποιεί τις $|B_t^{(i)}| = t$ και $B_1 \sim \text{Uniform}(\{-1, 1\}^n)$ (Πρόταση 4.1.1).

Αφού η B_1 είναι ομοιόμορφη στον διακριτό κύβο, η μέση τιμή και η διασπορά της f δίνονται από τις $\mathbb{E}[f] = \mathbb{E}[f(B_1)]$ και $\operatorname{Var}(f) = \operatorname{Var}(f(B_1))$, όπου οι μέσες τιμές στα δεξιά μέλη

είναι ως προς την τυχαιότητα της διαδικασίας B_t . Όμοια, η επιρροή της i -οστής συντεταγμένης δίνεται από την $I_i(f) = \mathbb{E}[\partial_i f(B_1)^2]$, όπου $\partial_i f$ είναι η μερική παράγωγος της f στην κατεύθυνση i , και η $\mathbb{E}[h_f^p]$ ισούται με $\mathbb{E}[\|\nabla f(B_1)\|_2^{2p}]$.

Η δύναμη της στοχαστικής διαδικασίας πηγάζει από το γεγονός ότι η συμπεριφορά της $f(B_1)$ γίνεται κατανοητή από την ανάλυση των διαδικασιών B_t , $f(B_t)$ και $\nabla f(B_t)$ για χρόνους μικρότερους από 1. Πράγματι, υπάρχει ένας φυσιολογικός τρόπος να επεκτείνουμε το πεδίο ορισμού μιας Boolean συνάρτησης στον συνεχή υπερκύβο $[-1, 1]^n$ έτσι ώστε οι διαδικασίες $f(B_t)$ και $\partial_i f(B_t)$ να γίνουν martingales. Τότε, μπορούμε να εκφράσουμε τη διασπορά της f ως

$$\text{Var}(f) = 2\mathbb{E}\left[\sum_{i=1}^n \int_0^1 t(\partial_i f(B_t))^2 dt\right]$$

(Λήμμα 4.1.3 και Πόρισμα 4.1.4). Κατόπιν, για να φράξουμε τη διασπορά αρκεί να φράξουμε το ολοκλήρωμα

$$\mathbb{E}\left[\sum_{i=1}^n \int_0^1 t(\partial_i f(B_t))^2 dt\right],$$

και για τον σκοπό αυτό μπορούμε να χρησιμοποιήσουμε εργαλεία από την πραγματική ανάλυση και τις στοχαστικές διαδικασίες. Συγκεκριμένα, δύο πολύ γνωστές ανισότητες – οι λεγόμενες ανισότητες Επιπέδου-1 και Επιπέδου-2 – μας δίνουν φράγματα για την ταχύτητα με την οποία οι διαδικασίες $\partial_i f(B_t)^2$ και το άθροισμά τους $\sum_{i=1}^n (\partial_i f(B_t))^2$ κινούνται συναρτήσει της τρέχουσας τιμής τους. Στο πλαίσιο των Gaussian διαδικασιών, παρόμοιες ιδέες που χρησιμοποιούν ανισότητες επιπέδων εμφανίζονται στο [16].

Στο σημείο αυτό αξίζει να σημειώσουμε μια σημαντική εννοιολογική διαφορά ανάμεσα στις υπάρχουσες τεχνικές που χρησιμοποιούν την υπερσυσταλτότητα του τελεστή θερμότητας και την τεχνική που εφαρμόζεται εδώ. Ενώ οι προηγούμενες αποδείξεις ξεκινούν από τη συνάρτηση f και αναλύουν τον τρόπο με τον οποίο μεταβάλλεται εφαρμόζοντας την ημιομάδα της θερμότητας, που αντιστοιχεί στο να κινούμαστε προς τα πίσω ως προς τον χρόνο t , η ανάλυση που εφαρμόζεται εδώ πηγαίνει προς τα εμπρός στον χρόνο. Μπορούμε να σκεφτόμαστε τη διαδικασία B_t ως έναν τρόπο να παίρνουμε δείγμα από τον διακριτό κύβο $\{-1, 1\}^n$ μέσω μιας συνεχούς διήθησης, προσθέτοντας «απειροστά ψηφία τυχαιότητας» καθώς προχωράει ο χρόνος. Η ανάλυση ξεκινάει από την $f(B_0) = \mathbb{E}[f]$ και εξετάζει τον τρόπο με τον οποίο εξελίσσονται τα martingales καθώς εκλεπτύνουμε την διήθηση, ή με άλλα λόγια, προσθέτουμε περισσότερη τυχαιότητα. Με μια πρώτη ματιά, αυτή η διαφορά μοιάζει να είναι απλώς παιδαγωγική, όμως η ισχύς αυτής της προσέγγισης βρίσκεται στο ότι η ανάλυση των μονοπατιών μας εφοδιάζει με νέα εργαλεία, όπως η χρήση χρόνων διακοπής και το θεώρημα βέλτιστης διακοπής, όπως και η δέσμευση στο παρελθόν.

Για να αποδείξουμε την εικασία του Talagrand, χρησιμοποιούμε την ανισότητα Επιπέδου-1 (Λήμμα 2.6.3) ώστε να φράξουμε το $\sum_{i=1}^n (\partial_i f(B_t))^2$ από μια δύναμη, που εξαρτάται από τον χρόνο, του αθροίσματος των τετραγώνων των επιρροών $\sum_{i=1}^n I_i(f)^2$. Ιδεολογικά, όταν αυτό το άθροισμα είναι μικρό, αυτό χοντρικά συνεπάγεται ότι η διαδικασία $\|\nabla f(B_t)\|_2^2$ πραγματοποιεί το μεγαλύτερο μέρος της κίνησής της πολύ κοντά στη χρονική στιγμή $t = 1$. Αυτή είναι ουσια-

στικά η σημασία του Θεωρήματος 2.5.4. Στη συνέχεια χρησιμοποιούμε αυτό το αποτέλεσμα για να δείξουμε ότι το μεγαλύτερο μέρος της τετραγωνικής κύμανσης της $f(B_t)$ προέρχεται από μονοπάτια στα οποία υπάρχει κάποια χρονική στιγμή t τέτοια ώστε η $\|\nabla f(B_t)\|_2$ να είναι μεγαλύτερη από $\alpha \left(\log \left(\frac{1}{\sum_{i=1}^n I_i^2(f)} \right) \right)^{1/2}$ (Λήμμα 4.3.2). Όμως, η τετραγωνική κύμανση είναι η ίδια μεγάλη με πιθανότητα που είναι ευθέως ανάλογη προς τη διασπορά (Λήμμα 3.2.1). Αυτό είναι ένα από τα σημεία όπου η ανάλυση μονοπατιών χρησιμοποιείται με κρίσιμο τρόπο. Χωρίς αυτήν, θα γνωρίζαμε μόνο ότι η τετραγωνική κύμανση έχει μεγάλη μέση τιμή, και αυτό δεν θα απέκλειε την περίπτωση ολόκληρη η συνεισφορά στη διασπορά να προέρχεται από κάποιο ενδεχόμενο με αμελητέα πιθανότητα. Εφόσον η $\|\nabla f(B_1)\|_2^{2p}$ είναι submartingale, αν υπάρχει κάποια χρονική στιγμή κατά την οποία η $\|\nabla f(B_1)\|_2^{2p}$ είναι μεγάλη, τότε ως προς τη μέση τιμή συνεχίζει να είναι μεγάλη. Έτσι, η $\mathbb{E} \left[\|\nabla f(B_1)\|_2^{2p} \right]$ είναι μεγαλύτερη από $\alpha \text{Var}(f) \left(\log \left(\frac{1}{\sum_{i=1}^n I_i^2(f)} \right) \right)^p$, και αυτό αποδεικνύει την αρχική εικασία του Talagrand (Θεώρημα 2.4.6) όταν $p = 1/2$. Με πρόσθετη φροντίδα, μπορεί κανείς να δείξει ότι η $f(B_t)$ είναι η ίδια μεγάλη για κάποια χρονική στιγμή προτού η κλίση γίνει μεγάλη, και αυτό αποδεικνύει το βελτιωμένο αποτέλεσμα (Θεώρημα 2.4.7).

Αξίζει εδώ να σημειώσουμε μια αναλογία ανάμεσα στην τεχνική που χρησιμοποιείται από τους Eldan και Gross και την τεχνική των Barthe και Maurey στο [1], όπου δίνουν μια στοχαστική απόδειξη της επέκτασης που είχε δώσει ο Bobkov για την ανισότητα (2.4.3). Χρησιμοποιούν ένα στοχαστικό επιχείρημα για να αποδείξουν μια μονοδιάστατη ανισότητα από την οποία έπεται η ανισότητα του Bobkov μέσω tensorization, και μία από τις κεντρικές συνιστώσες στην απόδειξή τους είναι να εξασφαλίσουν ότι μια διαδικασία που είναι ανάλογη με την $\|\nabla f(B_1)\|_2$ είναι submartingale (αυτό βασίζεται σε ιδέες που είχαν εισαχθεί από τους Capitaine, Hsu και Ledoux στο [12]).

Όμοια, για την απόδειξη του Θεωρήματος 2.4.5, χρησιμοποιούμε την ανισότητα Επιπέδου-1 (Λήμμα 2.6.2) για να φράξουμε καθεμία από τις $(\partial_i f(B_t))^2$ από μια δύναμη, που εξαρτάται από τον χρόνο, της επιρροής $I_i(f)$ (Λήμμα 5.1.1). Όταν οι επιρροές είναι μικρές, αυτό χοντρικά συνεπάγεται ότι το martingale $f(B_t)$ πραγματοποιεί το μεγαλύτερο μέρος της κίνησής του πολύ κοντά στη χρονική στιγμή $t = 1$. Κατόπιν, παίρνουμε το Θεώρημα 2.4.5 εισάγοντας αυτό το φράγμα στο ολοκλήρωμα (Πρόταση 5.1.2).

Η απόδειξη του Θεωρήματος 2.5.1 είναι πιο πολύπλοκη, και χρησιμοποιεί το γεγονός ότι η $f(B_t)$ είναι ταυτόχρονα διαδικασία αλμάτων και martingale: Για διαδικασίες αυτού του είδους, η διασπορά της $f(B_1)$ δίνεται από το άθροισμα των τετραγώνων των αλμάτων της $f(B_t)$ μέχρι την χρονική στιγμή 1:

$$\text{Var}(f) = \mathbb{E} \left[\sum_{s \in \text{Jump}(B_t)} (\Delta f(B_s))^2 \right] = 2\mathbb{E} \left[\sum_{i=1}^n \int_0^1 t (\partial_i f(B_t))^2 dt \right].$$

Το κύριο τεχνικό μέρος της απόδειξης (Πρόταση 5.1.3 και Λήμμα 5.3.1) δείχνει ότι αν οι $T(f) = \sum_{i=1}^n \frac{I_i(f)}{1 + \log(1/I_i(f))}$ και $\text{Var}(f)$ διαφέρουν μόνο κατά μία πολλαπλασιαστική σταθερά, τότε με μη αμελητέα πιθανότητα η διαδικασία $f(B_t)$ πρέπει να πραγματοποιήσει ένα σχετικά μεγάλο

άλμα κάπου στη διαδρομή. Χοντρικά, αυτό ισχύει διότι αν η διαδικασία $B_t^{(i)}$ κάνει άλμα την χρονική στιγμή t , τότε η συνάρτηση $f(B_t)$ κάνει κι αυτή άλμα, αλλάζοντας κατά μία ποσότητα ίση με $2t\partial_i f(B_t)$. Αν όλα τα άλματα ήταν μικρά, τότε η έκφραση $\sum_{s \in \text{Jump}(B_t)} (\Delta f(B_s))^2$ στο αριστερό μέλος της παραπάνω ισότητας (που ενδιαφέρεται μόνο για τα άλματα) θα πρέπει να είναι σημαντικά μικρότερη από το ολοκλήρωμα στο δεξιό μέλος (που ενδιαφέρεται μόνο για το μέγεθος των παραγώγων).

Τώρα, όταν η διαδικασία $f(B_t)$ κάνει ένα μεγάλο άλμα, αυτό αναγκαστικά σημαίνει ότι το μέτρο κάποιας από τις μερικές παραγώγους $\partial_i f$ είναι μεγάλο. Αφού η διαδικασία $\partial_i f(B_t)$ είναι επίσης martingale, αν είναι μεγάλη σε κάποια χρονική στιγμή, τότε συνεχίζει να είναι μεγάλη με σχετικά μεγάλη πιθανότητα. Όμως, την χρονική στιγμή $t = 1$, αφού η B_1 είναι ομοιόμορφη στον διακριτό κύβο, οι μόνες δυνατότητες για τις τιμές $\partial_i f(B_1)$ είναι -1 , 0 και 1 . Έτσι, είναι πιθανό να έχουμε $|\partial_i f(B_1)| = 1$. Αυτό αντιστοιχεί ακριβώς στο ότι το σημείο B_1 ανήκει στο σύνολο κορυφών, και έτσι προκύπτει ότι το σύνολο κορυφών είναι μεγάλο. Η διάκριση ανάμεσα στην εσωτερική και την εξωτερική κορυφή έπεται με παρόμοια επιχειρήματα, με χρήση συμμετρικοποίησης της B_t (Πρόταση 5.3.2).

ΚΕΦΑΛΑΙΟ 4

Το κύριο εργαλείο: Μια στοχαστική διαδικασία με άλματα

4.1 Κατασκευή και βασικές ιδιότητες

Η απόδειξη των Θεωρημάτων 2.4.5 και 2.5.1 βασίζεται στην κατασκευή μιας τμηματικά ομαλής στοχαστικής διαδικασίας $(B_t)_{t \geq 0}$ που είναι martingale και έχει τις ιδιότητες της επόμενης πρότασης. Μία από τις βασικές ιδιότητές της, είναι ότι μας επιτρέπει να εκφράσουμε τη διασπορά μιας συνάρτησης f συναρτήσει των μερικών παραγώγων της αρμονικής επεκτασής της, ως εξής:

$$\text{Var}(f) = 2\mathbb{E} \left[\sum_{i=1}^n \int_0^1 t (\partial_i f(B_t))^2 dt \right].$$

Πρόταση 4.1.1. Υπάρχει μια στοχαστική διαδικασία $(B_t)_{t \geq 0}$ η οποία είναι δεξιά συνεχές martingale και χαρακτηρίζεται από τις εξής ιδιότητες:

- $B_t \in \mathbb{R}^n$ και οι συνιστώσες $B_t^{(i)}$, $i \in [n]$ είναι ανεξάρτητες και ισόνομες.
- Κάθε $B_t^{(i)}$ είναι martingale.
- $|B_t^{(i)}| = t$ σχεδόν βεβαίως για κάθε $i \in [n]$ και κάθε $t \geq 0$.
- Για κάθε $t, h > 0$ ισχύει ότι

$$(4.1.1) \quad \mathbb{P} \left[\text{sign} B_{t+h}^{(i)} \neq \text{sign} B_t^{(i)} \mid B_t \right] = \frac{h}{2(t+h)}.$$

Απόδειξη. Έστω W_s μια τυπική κίνηση Brown και ας θεωρήσουμε την οικογένεια των χρόνων διακοπής

$$\tau(t) = \inf\{s > 0 : |W_s| > t\}.$$

Θέτοντας $X_t = W_{\tau(t)}$ για κάθε $t \geq 0$, έχουμε ότι $|X_t| = t$ και η X_t είναι martingale από το θεώρημα επιλεκτικής διακοπής. Παρατηρούμε ότι η X_t αποτυγχάνει να είναι δεξιά συνεχής

αν και μόνο αν $\text{sign}W_{\tau(t)} \neq \text{sign}W_{\tau(s)}$ για κάθε $s \neq t$ σε ένα ανοιχτό διάστημα γύρω από το t . Αυτό όμως το ενδεχόμενο έχει μηδενική πιθανότητα, και άρα υπάρχει ένα αντίγραφο της X_t το οποίο συμβολίζουμε επίσης με X_t και έχει μονοπάτια $t \rightarrow X_t(\omega)$ που είναι δεξιά συνεχή. Τότε, η ζητούμενη στοχαστική διαδικασία είναι η $B_t = (B_t^{(1)}, \dots, B_t^{(n)})$, όπου κάθε $B_t^{(i)}$ είναι ανεξάρτητο και ισόνομο αντίγραφο της X_t , για κάθε t και $i \in [n]$.

Για την απόδειξη της (4.1.1) έχουμε ότι αν θέσουμε $p = \mathbb{P}[\text{sign}X_{t+\varepsilon} \neq \text{sign}X_t \mid X_t]$ τότε από τον τρόπο κατασκευής της X_t έχουμε ότι

$$t \cdot \text{sign}X_t = X_t = \mathbb{E}[X_{t+h} \mid X_t] = p(-t-h)\text{sign}(X_t) + (1-p)(t+h)\text{sign}X_t.$$

Λύνοντας ως προς p βλέπουμε ότι

$$p = \frac{h}{2(t+h)}.$$

□

Όπως είναι ήδη προφανές, κάθε $B_t^{(i)}$ είναι μια κατά τμήματα ομαλή στοχαστική διαδικασία με ένταση $\lambda(t) = 1/(2t)$.

Έστω τώρα $J_i = \text{Jump}(B_t^{(i)})$ το σύνολο των σημείων ασυνέχειας της $B_t^{(i)}$. Όπως ήδη περιγράψαμε στην (2.6.1), αφού η αρμονική επέκταση μιας Boolean συνάρτησης είναι ένα άθροισμα μονωνύμων και το γινόμενο ανεξάρτητων martingales είναι επίσης martingale, λόγω και της ανεξαρτησίας των συντεταγμένων $B_t^{(i)}$ του τυχαίου διανύσματος B_t , συμπεραίνουμε το εξής.

Λήμμα 4.1.2. Για κάθε συνάρτηση $f : \{-1, 1\}^n \rightarrow \mathbb{R}$, η στοχαστική διαδικασία $f(B_t)$ είναι martingale.

Θα συμβολίζουμε αυτή τη στοχαστική διαδικασία με $f_t := f(B_t)$, και χωρίς βλάβη της γενικότητας θα γράφουμε $\partial_i f_t := \partial_i f(B_t)$ και $\nabla f_t = \nabla f(B_t)$. Αφού η B_t είναι δεξιά συνεχής, οι f_t είναι επίσης δεξιά συνεχείς και όταν αναφερόμαστε στα όρια από αριστερά στα σημεία ασυνέχειας της f_t , θα γράφουμε f_{t-} , $\partial_i f_{t-}$ και ∇f_{t-} , όπου $f_{t-} = \lim_{\varepsilon \searrow 0} f_{t-\varepsilon}$.

Αφού τώρα η f_t είναι μια τμηματικά ομαλή στοχαστική διαδικασία με άλματα, από την (3.2.1) η τετραγωνική της κύμανση είναι ίση με το άθροισμα των τετραγώνων των αλμάτων της. Τώρα, σχεδόν βεβαίως, η B_t μπορεί να κάνει άλμα μόνο σε μια συντεταγμένη κάθε φορά, και όταν η i -οστή συντεταγμένη $B_t^{(i)}$ κάνει άλμα τότε η τιμή της f_t μεταβάλλεται κατά $2t\partial_i f_t$, διότι η f είναι πολυγραμμική. Επομένως, η τετραγωνική κύμανση της f_t είναι ίση με

$$(4.1.2) \quad [f]_t = \sum_{i=1}^n \sum_{s \in J_i \cap [0, t]} (2s \cdot \partial_i f_s)^2.$$

Μια κρίσιμη ιδιότητα της διαδικασίας (B_t) είναι ότι η μέση τιμή των αλμάτων της συμπεριφέρεται με ομαλό τρόπο, όπως δείχνει το επόμενο λήμμα.

Λήμμα 4.1.3. Έστω $0 \leq t_1 < t_2 \leq 1$ και έστω g_t μια φραγμένη διαδικασία για την οποία ικανοποιείται τουλάχιστον μία από τις εξής συνθήκες:

- Η g_t είναι αριστερά συνεχής και μετρήσιμη ως προς τη διήθηση που παράγεται από τη διαδικασία $(B_s)_{0 \leq s \leq t}$.
- Υπάρχει μια συνεχής συνάρτηση $g : [-1, 1]^n \rightarrow \mathbb{R}$ τέτοια ώστε $g_t = g(B_t)$ για κάθε t .

Τότε, ισχύει ότι

$$(4.1.3) \quad \mathbb{E} \left[\sum_{t \in J_i \cap [t_1, t_2]} 4t^2 g_t \right] = 2 \mathbb{E} \left[\int_{t_1}^{t_2} t \cdot g_t dt \right].$$

Απόδειξη. Ας υποθέσουμε αρχικά ότι η g_t ικανοποιεί την πρώτη συνθήκη, δηλαδή είναι αριστερά συνεχής και μετρήσιμη ως προς τη διήθηση που παράγεται από τη διαδικασία $(B_s)_{0 \leq s \leq t}$. Για να αποδείξουμε την ισότητα (4.1.3), υποθέτουμε αρχικά ότι $t_1 > 0$, οπότε το πλήθος των αλμάτων της B_t στο διάστημα $[t_1, t_2]$ είναι σχεδόν βεβαίως πεπερασμένο. Για κάθε $N > 0$ χωρίζουμε το διάστημα $[t_1, t_2]$ σε N υποδιαστήματα με το ίδιο μήκος, θέτοντας $t_k^N = t_1 + \frac{k}{N}(t_2 - t_1)$ για $k = 0, 1, \dots, N$. Αφού τώρα, σχεδόν βεβαίως κανένα από τα άλματα της B_t δεν εμφανίζεται σε κάποιο t_k^N και υπάρχουν μόνο πεπερασμένα το πλήθος άλματα, έχουμε ότι υπάρχει σχεδόν βεβαίως πεπερασμένο N_0 τέτοιο ώστε για κάθε $N > N_0$, κάθε υποδιάστημα της διαμέρισης να περιέχει το πολύ ένα σημείο άλματος. Αφού η g_t είναι αριστερά συνεχής, έπεται ότι σχεδόν βεβαίως έχουμε

$$\sum_{t \in J_i \cap [t_1, t_2]} 4t^2 g_t = \lim_{N \rightarrow \infty} \left[\sum_{k=0}^{N-1} 4(t_k^N)^2 g_{t_k^N} \mathbb{1}_{J_i \cap [t_k^N, t_{k+1}^N] \neq \emptyset} \right].$$

Αφού όμως η g_t είναι φραγμένη, η ποσότητα

$$\sum_{k=0}^{N-1} 4(t_k^N)^2 g_{t_k^N} \mathbb{1}_{J_i \cap [t_k^N, t_{k+1}^N] \neq \emptyset}$$

είναι φραγμένη κατ' απόλυτη τιμή από μια σταθερά επί το πλήθος των αλμάτων της B_t στο διάστημα $[t_1, t_2]$, που είναι ολοκληρώσιμη συνάρτηση. Από το θεώρημα κυριαρχημένης σύγκλισης, παίρνουμε

$$\mathbb{E} \left[\sum_{t \in J_i \cap [t_1, t_2]} 4t^2 g_t \right] = \lim_{N \rightarrow \infty} \mathbb{E} \left[\sum_{k=0}^{N-1} 4(t_k^N)^2 g_{t_k^N} \mathbb{1}_{J_i \cap [t_k^N, t_{k+1}^N] \neq \emptyset} \right].$$

Αφού η $g_{t_k^N}$ είναι μετρήσιμη ως προς τη διήθηση που παράγει η $(B_s)_{0 \leq s \leq t_k^N}$, είναι ανεξάρτητη από το αν συμβαίνει ή όχι κάποιο άλμα στο διάστημα $[t_k^N, t_{k+1}^N]$, και έτσι μπορούμε να σπάσουμε τη μέση τιμή σε γινόμενο:

$$(4.1.4) \quad \mathbb{E} \left[\sum_{t \in J_i \cap [t_1, t_2]} 4t^2 g_t \right] = \lim_{N \rightarrow \infty} \sum_{k=0}^{N-1} \left(\mathbb{E}[4(t_k^N)^2 g_{t_k^N}] \right) \left(\mathbb{E}[\mathbb{1}_{J_i \cap [t_k^N, t_{k+1}^N] \neq \emptyset}] \right).$$

Ο πληθάριθμος του συνόλου $J_i = \text{Jump}(B_t^{(i)})$ είναι μια διαδικασία Poisson με ένταση $1/(2t)$ και

άρα το πλήθος των αλμάτων στο διάστημα $[t_k^N, t_{k+1}^N]$ κατανέμεται σαν $\text{Poisson}(\lambda)$, όπου

$$\lambda = \int_{t_k^N}^{t_{k+1}^N} \frac{1}{2t} dt = \frac{1}{2} \log \left(\frac{t_{k+1}^N}{t_k^N} \right).$$

Η πιθανότητα, τώρα, να υπάρχει τουλάχιστον ένα άλμα της $B_t^{(i)}$ στο διάστημα $[t_k^N, t_{k+1}^N]$ είναι ίση με

$$\begin{aligned} \mathbb{P}[J_i \cap [t_k^N, t_{k+1}^N] \neq \emptyset] &= 1 - e^{-\lambda} = 1 - \sqrt{\frac{t_k^N}{t_{k+1}^N}} = 1 - \sqrt{1 - \frac{(t_2 - t_1)/N}{t_{k+1}^N}} \\ &= \frac{(t_2 - t_1)/N}{2t_{k+1}^N} + O\left(\frac{1}{N^2}\right). \end{aligned}$$

Αντικαθιστώντας στην (4.1.4) παίρνουμε

$$\mathbb{E} \left[\sum_{t \in J_i \cap [t_1, t_2]} 4t^2 g_t \right] = \lim_{N \rightarrow \infty} \mathbb{E} \left[\sum_{k=0}^{N-1} 4(t_k^N)^2 g_{t_k^N} \left(\frac{(t_2 - t_1)/N}{2t_{k+1}^N} + O\left(\frac{1}{N^2}\right) \right) \right].$$

Όμως ο παράγοντας $O\left(\frac{1}{N^2}\right)$ είναι αμελητέος όταν $N \rightarrow \infty$, διότι το άθροισμα περιέχει μόνο N το πλήθος φραγμένους όρους. Επομένως, χρησιμοποιώντας και το θεώρημα φραγμένης σύγκλισης, έχουμε τελικά ότι

$$\begin{aligned} \lim_{N \rightarrow \infty} \mathbb{E} \left[\sum_{k=0}^{N-1} 4(t_k^N)^2 g_{t_k^N} \left(\frac{(t_2 - t_1)/N}{2t_{k+1}^N} + O\left(\frac{1}{N^2}\right) \right) \right] &= \lim_{N \rightarrow \infty} \mathbb{E} \left[\sum_{k=0}^{N-1} \left(2t_k^N g_{t_k^N} \right) \frac{t_2 - t_1}{N} \right] \\ &= \mathbb{E} \left[\lim_{N \rightarrow \infty} \sum_{k=0}^{N-1} \left(2t_k^N g_{t_k^N} \right) \frac{t_2 - t_1}{N} \right]. \end{aligned}$$

Αφού όμως η g_t είναι συνεχής σχεδόν παντού, από τον ορισμό του ολοκληρώματος Riemann, το όριο είναι ίσο με $2\mathbb{E} \left[\int_{t_1}^{t_2} t g_t dt \right]$ και τελικά παίρνουμε ότι

$$\mathbb{E} \left[\sum_{t \in J_i \cap [t_1, t_2]} 4t^2 g_t \right] = 2\mathbb{E} \left[\int_{t_1}^{t_2} t \cdot g_t dt \right]$$

για κάθε $t_1 > 0$. Παίρνοντας τώρα όριο καθώς το $t_1 \rightarrow 0$ έχουμε το ζητούμενο αποτέλεσμα και στην περίπτωση που $t_1 = 0$, λόγω της συνέχειας του δεξιού μέλους ως προς t_1 .

Αν υποθέσουμε ότι ικανοποιείται η δεύτερη συνθήκη, δηλαδή ότι $g_t = g(B_t)$ για κάποια συνεχή συνάρτηση $g : [-1, 1]^n \rightarrow \mathbb{R}$, τότε η απόδειξη είναι εντελώς ανάλογη. $\square$

Πόρισμα 4.1.4. Έστω $f : \{-1, 1\}^n \rightarrow \mathbb{R}$. Τότε, για κάθε $t_0 > 0$ ισχύει ότι

$$(4.1.5) \quad \text{Var}(f_{t_0}) = 2\mathbb{E} \left[\sum_{i=1}^n \int_0^{t_0} t (\partial_i f_t)^2 dt \right].$$

Απόδειξη. Αφού η f_{t_0} είναι martingale, από την (3.2.2) και την (4.1.2) έχουμε ότι

$$\text{Var}(f_{t_0}) = \mathbb{E} \left[\sum_{i=1}^n \sum_{t \in J_i \cap [0, t_0]} (2t \partial_i f_t)^2 \right].$$

Θέτοντας $g_t = (\partial_i f_t)^2$ για κάθε $i \in [n]$ στην (4.1.3) παίρνουμε το ζητούμενο. $\square$

Πόρισμα 4.1.5. Έστω $f : \{-1, 1\}^n \rightarrow \mathbb{R}$. Τότε,

$$\frac{d}{dt} \mathbb{E}[f_t^2] = 2t \mathbb{E} \left[\sum_{i=1}^n (\partial_i f_t)^2 \right] = 2t \mathbb{E} [\|\nabla f_t\|_2^2].$$

Απόδειξη. Αφού η f_t είναι martingale, έχουμε ότι

$$\frac{d}{dt} \mathbb{E}[f_t^2] = \frac{d}{dt} (\mathbb{E}[f_t^2] - \mathbb{E}[f_0^2]) = \frac{d}{dt} \mathbb{E}[(f_t - f_0)^2] = \frac{d}{dt} \text{Var}(f_t).$$

Παραγωγίζοντας ως προς t στην (4.1.5) και χρησιμοποιώντας το θεμελιώδες θεώρημα του Απειροστικού Λογισμού στο δεξί μέλος, παίρνουμε το ζητούμενο αποτέλεσμα. $\square$

4.2 Η διαδικασία της επιρροής

Είναι γνωστό (βλέπε, για παράδειγμα, [25, Ενότητα 4.3]) ότι αν η f έχει ανάπτυγμα Fourier $f(x) = \sum_S \widehat{f}(S) w_S(x)$, τότε η ευστάθεια θορύβου της, η οποία ορίστηκε στην (2.5.2), δίνεται από την

$$S_\varepsilon(f) = \sum_{S \neq \emptyset} \widehat{f}(S)^2 (1 - \varepsilon)^{|S|}.$$

Από την άλλη πλευρά, αν θυμηθούμε ότι $f_t = f(B_t)$, με έναν σύντομο υπολογισμό διαπιστώνουμε ότι

$$\text{Var}(f_t) = \sum_{S \neq \emptyset} \widehat{f}(S)^2 t^{2|S|}.$$

Έπεται ότι $S_\varepsilon(f) = \text{Var}(f_{\sqrt{1-\varepsilon}})$ και η ανισότητα (2.5.3) στη διατύπωση του Θεωρήματος 2.5.4 γίνεται

$$\text{Var}(f_{\sqrt{1-\varepsilon}}) \leq C \text{Var}(f) \left(\sum_{i=1}^n I_i(f)^2 \right)^{c\varepsilon}.$$

Χρησιμοποιώντας και την (4.1.5) ξαναγράφουμε αυτή την ανισότητα στη μορφή

$$(4.2.1) \quad \mathbb{E} \left[\sum_{i=1}^n \int_0^{\sqrt{1-\varepsilon}} t (\partial_i f_t)^2 dt \right] \leq C \text{Var}(f) \left(\sum_{i=1}^n I_i(f)^2 \right)^{c\varepsilon}.$$

Θα ασχοληθούμε στο εξής με αυτή την ανισότητα αντί για την αρχική ανισότητα (2.5.3).

Για κάθε δείκτη i , συμβολίζουμε με $f^{(i)}$ την αρμονική επέκταση της $|\partial_i f|$, και ορίζουμε

$$(4.2.2) \quad f_t^{(i)} = f^{(i)}(B_t).$$

Αν η f είναι μονότονη, τότε $f^{(i)} = \partial_i f$, αφού οι παράγωγοι είναι θετικές, όμως γενικά

$$(4.2.3) \quad f^{(i)}(x) \geq |\partial_i f(x)| \text{ για κάθε } x \in [-1, 1]^n$$

λόγω κυρτότητας. Ειδικότερα, εισάγοντας την (4.2.3) στο Πόρισμα 4.1.4 έχουμε ότι

$$(4.2.4) \quad \text{Var}(f_{t_0}) \leq 2 \sum_{i=1}^n \int_0^{t_0} t \mathbb{E} \left[\left(f_t^{(i)} \right)^2 \right] dt.$$

Ονομάζουμε τη διαδικασία $f_t^{(i)}$ «διαδικασία επιρροής» λόγω του τρόπου με τον οποίο η μέση τιμή του τετραγώνου της συνδέεται με την επιρροή της f : Λόγω της (2.1.2), έχουμε ότι

$$(4.2.5) \quad f_0^{(i)} = \mathbb{E} \left[f^{(i)} \right] = \mathbb{E} \left[|\partial_i f| \right] = I_i(f).$$

Έτσι, τη χρονική στιγμή 0 ισχύει ότι $\mathbb{E} \left[\left(f_0^{(i)} \right)^2 \right] = \left(f_0^{(i)} \right)^2 = I_i(f)^2$, ενώ τη χρονική στιγμή 1, εφόσον $f^{(i)}(y)^2 = f^{(i)}(y)$ για $y \in \{-1, 1\}^n$, ισχύει ότι $\mathbb{E} \left[\left(f_1^{(i)} \right)^2 \right] = \mathbb{E} \left[f_1^{(i)} \right] = \mathbb{E} \left[f^{(i)} \right] = I_i(f)$. Η μέση τιμή $\mathbb{E} \left[\left(f_t^{(i)} \right)^2 \right]$ αυξάνει από την τιμή $I_i(f)^2$ ως την $I_i(f)$ καθώς το t αυξάνει από το 0 στο 1. Συμβολίζουμε αυτή τη μέση τιμή με $\psi_i(t) := \mathbb{E} \left[\left(f_t^{(i)} \right)^2 \right]$. Τότε, η (4.2.4) γίνεται

$$\text{Var}(f(B_{t_0})) \leq 2 \sum_{i=1}^n \int_0^{t_0} t \psi_i(t) dt.$$

Μπορούμε να χειριστούμε το ολοκλήρωμα $\mathbb{E} \left[\int_0^1 t \psi_i(t) dt \right]$ πιο εύκολα χρησιμοποιώντας μια αλλαγή χρόνου μέσω της οποίας η $\psi_i(t)$ γίνεται λογαριθμικά κυρτή. Στη συνέχεια, μπορούμε να την φράξουμε από μια δύναμη της επιρροής. Γι' αυτόν τον σκοπό, για κάθε $s \in (0, \infty)$, θέτουμε

$$(4.2.6) \quad \varphi_i(s) := \psi_i(e^{-s}) = \mathbb{E} \left[\left(f_{e^{-s}}^{(i)} \right)^2 \right].$$

Λήμμα 4.2.1. Έστω g η αρμονική επέκταση μιας Boolean συνάρτησης, και έστω $h(s) = \mathbb{E} \left[g(B_{e^{-s}})^2 \right]$. Τότε, η $h(s)$ είναι λογαριθμικά κυρτή συνάρτηση του s .

Απόδειξη. Αναπτύσσοντας την g ως πολυώνυμο Fourier, παίρνουμε

$$(4.2.7) \quad \begin{aligned} h(s) &= \mathbb{E} \left[g(B_{e^{-s}})^2 \right] = \mathbb{E} \left[\left(\sum_{S \subseteq [n]} \widehat{g}(S) \prod_{i \in S} (B_{e^{-s}}^{(i)}) \right)^2 \right] \\ &= \sum_{S \subseteq [n]} \widehat{g}(S)^2 \prod_{i \in S} \mathbb{E} \left[\left(B_{e^{-s}}^{(i)} \right)^2 \right] \\ &= \sum_{S \subseteq [n]} \widehat{g}(S)^2 e^{-2s|S|}. \end{aligned}$$

Η τελευταία ποσότητα είναι θετικός γραμμικός συνδυασμός των λογαριθμικά κυρτών συναρτήσεων $e^{-2s|S|}$, άρα είναι κι αυτή λογαριθμικά κυρτή (βλ. [11, Ενότητα 3.5.2]). $\square$

Τέλος, θα χρειαστούμε το ακόλουθο απλό τεχνικό λήμμα.

Λήμμα 4.2.2. Έστω $g : [0, \infty) \rightarrow [0, \infty)$ διαφορίσιμη συνάρτηση που ικανοποιεί την

$$(4.2.8) \quad g'(t) \leq C \cdot g(t) \log \left(\frac{K}{g(t)} \right),$$

όπου C, K είναι κάποιες θετικές σταθερές. Υποθέτουμε ότι $g(0) \leq K/2$. Τότε, υπάρχει κάποια χρονική στιγμή t_0 , η οποία εξαρτάται μόνο από τις σταθερές C και K , τέτοια ώστε, για κάθε $t \in [0, t_0]$,

$$g(t) \leq K^{Ct} g(0)^{1-Ct}.$$

Απόδειξη. Θέτουμε $t_1 = \inf \{t : g(t) \geq K\}$ και $L = \max \left\{ x \log \left(\frac{K}{x} \right) : x \in [0, K] \right\}$. Σημειώνουμε ότι το L εξαρτάται μόνο από το K . Τότε, για κάθε $t \leq t_1$ έχουμε ότι

$$g'(t) \leq C \cdot L.$$

Ολοκληρώνοντας, συμπεραίνουμε ότι για κάθε $t \leq t_1$

$$g(t) \leq g(0) + tCL \leq \frac{K}{2} + tCL.$$

Ειδικότερα, $t_1 \geq \frac{K}{2CL}$, αλλιώς θα είχαμε $g(t_1) < K$, το οποίο έρχεται σε αντίφαση με τον ορισμό του t_1 και τη συνέχεια της g . Θέτοντας $t_0 = \frac{K}{4CL}$, έχουμε ότι $g(t) < K$ για κάθε $t \in [0, t_0]$. Αυτό μας εξασφαλίζει ότι η $\log \left(\frac{K}{g(t)} \right)$ είναι θετική σε αυτό το διάστημα, το οποίο σημαίνει ότι μπορούμε αναδιατάσσοντας την ανισότητα (4.2.8) να πάρουμε την

$$-\frac{g'(t)}{g(t) \log \left(\frac{g(t)}{K} \right)} \leq C$$

για κάθε $t \in [0, t_0]$. Ένας απλός υπολογισμός δείχνει ότι το αριστερό μέλος είναι η παράγωγος της $-\log \log (K/g)$. Ολοκληρώνοντας από το 0 ως το t , παίρνουμε

$$\log \log \left(\frac{K}{g(0)} \right) - \log \log \left(\frac{K}{g(t)} \right) \leq Ct.$$

Ξαναγράφοντας αυτή την ανισότητα στη μορφή

$$\log \log \left(\frac{K}{g(t)} \right) \geq \log \log \left(\frac{K}{g(0)} \right) - Ct$$

και συνθέτοντας με την εκθετική συνάρτηση δύο φορές, παίρνουμε την ανισότητα

$$g(t) \leq \left(\frac{1}{K} \right)^{e^{-Ct}-1} g(0)^{e^{-Ct}}.$$

Από το ανάπτυγμα Taylor της e^{-Ct} , μπορούμε να γράψουμε $e^{-Ct} := 1 - Ct + R(Ct)$, όπου $R(x) = O(x^2)$ και $R(x) > 0$ για $x > 0$. Έπεται ότι

$$\begin{aligned} g(t) &\leq g(0)^{1-Ct+R(Ct)} \left(\frac{1}{K}\right)^{-Ct+R(Ct)} = g(0)^{1-Ct} K^{Ct} \left(\frac{g(0)}{K}\right)^{R(Ct)} \\ &\leq g(0)^{1-Ct} K^{Ct} 2^{-R(Ct)} \leq g(0)^{1-Ct} K^{Ct}, \end{aligned}$$

που είναι ο ισχυρισμός του λήμματος. □

4.3 Απόδειξη της βελτιωμένης εικασίας Talagrand

Θα αποδείξουμε το Θεώρημα 2.4.7 υποθέτοντας ότι ισχύει η βελτιωμένη ανισότητα Keller-Kindler (4.2.1). Χωρίς περιορισμό της γενικότητας υποθέτουμε ότι $\mathbb{E}[f] \leq 0$.

Η συνάρτηση $g_f : \{-1, 1\}^n \rightarrow [0, 1]$ που χρησιμοποιείται στο Θεώρημα 2.4.7 είναι η

$$g_f(y) = \mathbb{E} \left[\sup_{0 \leq s \leq 1} \frac{1 + f_s}{2} \mid B_1 = y \right].$$

Από την ανισότητα του Doob (βλέπε, για παράδειγμα, [15, Θεώρημα 4.4.4]), γνωρίζουμε ότι αν $(X_t)_{t=0}^1$ είναι ένα μη αρνητικό martingale, τότε

$$\mathbb{E} \left[\left(\sup_{0 \leq s \leq 1} X_s \right)^2 \right] \leq 4 \mathbb{E} [X_1^2].$$

Χρησιμοποιώντας αυτή την ανισότητα για το μη αρνητικό martingale $X_t = \frac{1+f_t}{2}$, παίρνουμε

$$\begin{aligned} \mathbb{E} [g_f^2] &= \mathbb{E} \left[\mathbb{E} \left[\sup_{0 \leq s \leq 1} \frac{1 + f_s}{2} \mid B_1 = y \right]^2 \right] \leq \mathbb{E} \left[\mathbb{E} \left[\sup_{0 \leq s \leq 1} \left(\frac{1 + f_s}{2} \right)^2 \mid B_1 = y \right] \right] \\ &= \mathbb{E} \left[\sup_{0 \leq s \leq 1} \left(\frac{1 + f_s}{2} \right)^2 \right] \leq 2(1 + \mathbb{E}[f]) = 2 \frac{1 - (\mathbb{E}[f])^2}{1 - \mathbb{E}[f]} \leq 2 \text{Var}(f), \end{aligned}$$

όπως απαιτεί το θεώρημα.

Περνάμε τώρα στην απόδειξη της ανισότητας (2.4.6) χρησιμοποιώντας την g_f . Γι' αυτόν τον σκοπό, μπορούμε να συνδέσουμε το γινόμενο $h_f^p(y) g_f(y)$ με τις στοχαστικές κατασκευές της προηγούμενης ενότητας. Από τον ορισμό της διακριτής παραγώγου, για κάθε $y \in \{-1, 1\}^n$, έχουμε $\partial_i f(y) = 0$ αν $f(y) = f(\sigma_i(y))$, και $\partial_i f(y) = \pm 1$ αν $f(y) \neq f(\sigma_i(y))$, άρα

$$h_f(y) = \sum_{i=1}^n \partial_i f(y)^2 = \|\nabla f(y)\|_2^2.$$

Έτσι βλέπουμε ότι $h_f^p(y) = \|\nabla f(y)\|_2^{2p}$, και χρησιμοποιώντας αυτή τη σχέση ορίζουμε τη

στοχαστική διαδικασία

$$(4.3.1) \quad \Psi_t = \|\nabla f_t\|_2^{2p} \sup_{0 \leq s \leq t} \frac{1 + f_s}{2},$$

και σημειώνουμε ότι

$$\mathbb{E}[\Psi_1] = \mathbb{E}[h_f^p g_f].$$

Στόχος μας είναι λοιπόν να φράξουμε την $\mathbb{E}[\Psi_1]$ από κάτω.

Από το γεγονός ότι η συνάρτηση $\|\cdot\|_2^{2p} : \mathbb{R}^n \rightarrow \mathbb{R}$ είναι κυρτή και την ανισότητα Jensen, βλέπουμε ότι η διαδικασία $\|\nabla f_t\|_2^{2p}$ είναι submartingale. Κατόπιν, ένας απλός υπολογισμός δείχνει ότι η $\mathbb{E}[\Psi_t]$ είναι αύξουσα συνάρτηση του t . Για το υπόλοιπο αυτής της παραγράφου υποθέτουμε λοιπόν ότι

$$(4.3.2) \quad \mathbb{E}[\Psi_t] \leq \text{Var}(f) \cdot \left(\log \left(2 + \frac{e}{\sum_i I_i(f)^2} \right) \right)^p \text{ για κάθε } 0 \leq t \leq 1,$$

αλλιώς δεν έχουμε τίποτα να δείξουμε.

Στην επόμενη πρόταση θα δείξουμε ότι υπάρχει χρόνος διακοπής τ_α τέτοιος ώστε με μεγάλη πιθανότητα η Ψ_{τ_α} να είναι μεγάλη. Σταθεροποιούμε $\alpha > 0$ που η τιμή του θα επιλεγεί αργότερα, και ορίζουμε

$$(4.3.3) \quad \tau_\alpha := \inf \left\{ 0 \leq t \leq 1 \mid \Psi_t > \frac{1}{8} \alpha^{2p} \left(\log \left(2 + \frac{e}{\sum_i I_i(f)^2} \right) \right)^p \right\} \wedge 1.$$

Πρόταση 4.3.1. Υποθέτουμε ότι ισχύει η (4.3.2). Τότε υπάρχουν απόλυτες σταθερές $C > 0$ και $\alpha > 0$ ώστε

$$\mathbb{P}[\tau_\alpha < 1] \geq C \text{Var}(f).$$

Υποθέτοντας ότι αυτή η πρόταση έχει αποδειχθεί, ολοκληρώνουμε την απόδειξη της βελτιωμένης εικασίας Talagrand ως εξής:

Απόδειξη του Θεωρήματος 2.4.7. Δεσμεύοντας ως προς το ενδεχόμενο $\tau_\alpha < 1$, έχουμε

$$\begin{aligned} \mathbb{E}[\Psi_1] &\geq \mathbb{E}[\Psi_{\tau_\alpha}] \geq \mathbb{E}[\Psi_{\tau_\alpha} \mid \tau_\alpha < 1] \mathbb{P}[\tau_\alpha < 1] \\ &\geq \frac{1}{8} \alpha^{2p} \left(\log \left(2 + \frac{e}{\sum_i I_i(f)^2} \right) \right)^p \cdot C \text{Var}(f). \end{aligned}$$

□

Το υπόλοιπο αυτής της ενότητας είναι αφιερωμένο στην απόδειξη της Πρότασης 4.3.1. Η βασική ιδέα είναι να δούμε πώς συνεισφέρουν τα διαφορετικά μονοπάτια στην τετραγωνική κύμανση της f_t . Το μεγαλύτερο μέρος της τετραγωνικής κύμανσης προκύπτει από μονοπάτια στα οποία η νόρμα της κλίσης $\|\nabla f_t\|_2$ είναι μεγάλη. Αφού η τετραγωνική κύμανση έχει σχετικά μεγάλη πιθανότητα να είναι μεγάλη, η $\|\nabla f_t\|_2$ πρέπει κι αυτή να είναι μεγάλη με σχετικά μεγάλη

πιθανότητα. Αυτό το επιχείρημα λαμβάνει υπόψη του τη συνεισφορά της κλίσης στην Ψ_t . Για να δούμε τη συνεισφορά του supremum, δείχνουμε ότι με αρκετά μεγάλη πιθανότητα είτε η f_t κάνει ένα μεγάλο άλμα (το οποίο έχει ως αποτέλεσμα τόσο το $\sup f_t$ όσο και η νόρμα της κλίσης να είναι ταυτόχρονα μεγάλα) ή υπάρχει κάποια χρονική στιγμή όπου η θέση της f_t είναι φραγμένη μακριά από τα άκρα $\{-1, 1\}$, κάτι που επιτρέπει στην κλίση της να είναι μεγάλη αργότερα.

Απόδειξη της Πρότασης 4.3.1. Ορίζουμε

$$(4.3.4) \quad \vartheta := \inf\{t \geq 0 \mid f_t > 0\} \wedge 1.$$

Αφού η f_t είναι martingale, έχουμε ότι

$$f_0 = \mathbb{E}[f_1] = 2\mathbb{P}[f_1 = 1] - 1,$$

και αφού $\{f_1 = 1\} \subseteq \{\vartheta < 1\}$, έπεται ότι

$$(4.3.5) \quad \mathbb{P}[\vartheta < 1] \geq \mathbb{P}[f_1 = 1] = \frac{1 + f_0}{2} = \frac{1 - f_0^2}{2(1 - f_0)} = \frac{\text{Var}(f)}{2(1 - f_0)} \geq \frac{1}{4} \text{Var}(f).$$

Δεσμεύοντας στο $\vartheta < 1$ έχουμε ότι

$$\begin{aligned} \mathbb{P}[\tau_\alpha < 1] &\geq \mathbb{P}[\tau_\alpha < 1 \mid \vartheta < 1] \mathbb{P}[\vartheta < 1] \\ &\geq \frac{1}{4} \text{Var}(f) \cdot \mathbb{P}[\tau_\alpha < 1 \mid \vartheta < 1]. \end{aligned}$$

Μένει να δείξουμε ότι για αρκετά μικρό (αλλά σταθερό) α η πιθανότητα $\mathbb{P}[\tau_\alpha < 1 \mid \vartheta < 1]$ είναι μεγαλύτερη από κάποια σταθερά.

Έχουμε υποθέσει ότι $\mathbb{E}[f] \leq 0$, άρα η διαδικασία f_t ξεκινάει με $f_0 \leq 0$. Υπάρχουν δύο διαφορετικοί τρόποι για να περάσει η f_t πάνω από την τιμή 0. Θα μπορούσε να περάσει πάνω από το 0 με συνεχή τρόπο ή να κάνει άλμα από κάποια τιμή μικρότερη από 0 σε κάποια τιμή μεγαλύτερη από 0. Συνεχίζουμε την απόδειξη διακρίνοντας δύο περιπτώσεις, οι οποίες εξαρτώνται από την πιθανότητα να κάνει η f_t ένα πολύ μεγάλο άλμα πάνω από το 0 τη χρονική στιγμή ϑ .

4.3.1 Περίπτωση 1n-Μικρά άλματα

Υποθέτουμε αρχικά ότι $\mathbb{P}[f_\vartheta \in [0, 1/2] \mid \vartheta < 1] \geq 1/2$, δηλαδή με σταθερή πιθανότητα δεν γίνεται κανένα μεγάλο άλμα την χρονική στιγμή ϑ . Οι επόμενες δύο προτάσεις αποδεικνύουν ότι με μεγάλη πιθανότητα η τετραγωνική κύμανση από τη χρονική στιγμή ϑ και μετά πρέπει να είναι μεγάλη και ως εκ τούτου το μεγαλύτερο μέρος της τετραγωνικής κύμανσης αποκτάται από τους χρόνους όπου, προτού η στοχαστική διαδικασία κάνει άλμα, η κλίση $\nabla f_{t-} = \lim_{\varepsilon \searrow 0} \nabla f_{t-\varepsilon}$

είναι μεγάλη. Για να αποδείξουμε όλα αυτά, θα χρειαστούμε κάποιους ορισμούς. Έστω

$$(4.3.6) \quad F_{\alpha,t} = \left\{ \|\nabla f_{t-}\|_2 > a \sqrt{\log \left(2 + \frac{e}{\sum_i I_i(f)^2} \right)} \right\}$$

το ενδεχόμενο η νόρμα της κλίσης να είναι μεγάλη ακριβώς πριν από τη χρονική στιγμή t , και έστω

$$(4.3.7) \quad E_t = \left\{ \sup_{0 \leq s < t} f_s \geq 0 \right\}$$

το ενδεχόμενο η ποσότητα $\sup f_s$ να είναι μεγάλη γνήσια πριν από την χρονική στιγμή t . Έστω επίσης

$$(4.3.8) \quad V_\alpha = \sum_{i=1}^n \sum_{t \in J_i \cap [0,1]} (2t \partial_i f_t)^2 \mathbb{1}_{F_{\alpha,t}^c} \mathbb{1}_{E_t}$$

η τετραγωνική κύμανση που έχει αποκτηθεί στις χρονικές στιγμές όπου το supremum είναι μεγάλο αλλά η κλίση είναι μικρή, και έστω

$$(4.3.9) \quad V^{t_1 \rightarrow t_2} = \sum_{i=1}^n \sum_{t \in J_i \cap (t_1, t_2]} (2t \partial_i f_t)^2$$

η μεταβολή της τετραγωνικής κύμανσης από τη χρονική στιγμή t_1 έως και τη χρονική στιγμή t_2 .

Λήμμα 4.3.2. Έστω $0 < \alpha < 1/e$ και ας υποθέσουμε ότι η (4.3.2) ισχύει. Τότε υπάρχει συνάρτηση $\rho : [0, 1] \rightarrow \mathbb{R}$ για την οποία ισχύει ότι $\lim_{x \rightarrow 0} \rho(x) = 0$, τέτοια ώστε

$$(4.3.10) \quad E[V_\alpha] \leq \rho(\alpha) \text{Var}(f).$$

Απόδειξη. Πρώτα απ' όλα θα εκφράσουμε την ποσότητα

$$\mathbb{E}[V_\alpha] = \mathbb{E} \left[\sum_{i=1}^n \sum_{t \in J_i \cap [0,1]} (2t \partial_i f_t)^2 \mathbb{1}_{F_{\alpha,t}^c} \mathbb{1}_{E_t} \right]$$

που ορίστηκε στην (4.3.8) σαν ολοκλήρωμα αντί για άθροισμα πάνω απ' όλα τα άλματα της f . Αφού η $\partial_i f_t$ είναι ανεξάρτητη από τη συντεταγμένη i , για κάθε $t \in J_i$ ισχύει ότι $\partial_i f_t = \partial_i f_{t-}$ και άρα

$$\mathbb{E}[V_\alpha] = \mathbb{E} \left[\sum_{i=1}^n \sum_{t \in J_i \cap [0,1]} (2t \partial_i f_{t-})^2 \mathbb{1}_{F_{\alpha,t}^c} \mathbb{1}_{E_t} \right].$$

Η στοχαστική διαδικασία $g_t = (\partial_i f_{t-})^2 \mathbb{1}_{F_{\alpha,t}^c} \mathbb{1}_{E_t}$ είναι μετρήσιμη ως προς τη διήθηση που πα-

ράγεται απο την $(B_s)_{0 \leq s < t}$ και αριστερά συνεχής. Επομένως, απο το Λήμμα 4.1.3 έχουμε ότι

$$\begin{aligned}
 (4.3.11) \quad \mathbb{E}[V_\alpha] &= \mathbb{E} \left[\sum_{i=1}^n \sum_{t \in J_i \cap [0,1]} 4t^2 g_t \right] \\
 &= 2 \int_0^1 t \mathbb{E} \left[\|f_{t-}\|_2^2 \mathbb{1}_{F_{\alpha,t}^c} \mathbb{1}_{E_t} \right] dt \\
 &= 2 \int_0^1 t \mathbb{E} \left[\|f_t\|_2^2 \mathbb{1}_{F_{\alpha,t}^c} \mathbb{1}_{E_t} \right] dt,
 \end{aligned}$$

όπου η τελευταία ισότητα ισχύει επειδή η ποσότητα $\|f_{t-}\|_2^2 \mathbb{1}_{F_{\alpha,t}^c} \mathbb{1}_{E_t}$ διαφέρει απο την ποσότητα $\|f_t\|_2^2 \mathbb{1}_{F_{\alpha,t}^c} \mathbb{1}_{E_t}$ μόνο σε πεπερασμένα το πλήθος σημεία ασυνέχειας.

Έστω τώρα $\delta' > 0$ το οποίο ορίζουμε ως εξής

$$\delta' = \begin{cases} \frac{c^{-1} \log(1/\alpha)}{\log\left(2 + \frac{e}{\sum_i I_i(f)^2}\right)} & \text{αν } \sum_i I_i(f)^2 \leq \frac{1}{2} \\ 1 & \text{αλλιώς} \end{cases},$$

όπου c είναι η απόλυτη σταθερά του Θεωρήματος 2.5.4, και έστω

$$(4.3.12) \quad \delta = \min\{\delta', 1\}.$$

Θεωρούμε το ολοκλήρωμα

$$\begin{aligned}
 (4.3.13) \quad \int_0^1 t \mathbb{E} \left[\|\nabla f_t\|_2^2 \mathbb{1}_{F_{\alpha,t}^c} \mathbb{1}_{E_t} \right] dt &= \int_0^{1-\delta} t \mathbb{E} \left[\|\nabla f_t\|_2^2 \mathbb{1}_{F_{\alpha,t}^c} \mathbb{1}_{E_t} \right] dt \\
 &\quad + \int_{1-\delta}^1 t \mathbb{E} \left[\|\nabla f_t\|_2^2 \mathbb{1}_{F_{\alpha,t}^c} \mathbb{1}_{E_t} \right] dt.
 \end{aligned}$$

Το πρώτο ολοκλήρωμα του δεξιού μέλους είναι ίσο με μηδέν αν $\delta = 1$. Διαφορετικά, αναγκαστικά θα πρέπει να ισχύει ότι $\sum_i I_i(f)^2 \leq 1/2$, και σε αυτή την περίπτωση μπορούμε να φράξουμε το ολοκλήρωμα απο πάνω χρησιμοποιώντας την (4.2.1): αφού $1 - \delta \leq \sqrt{1 - \delta}$ για κάθε $\delta \in [0, 1]$, έχουμε ότι

$$\begin{aligned}
 \int_0^{1-\delta} t \mathbb{E} \left[\|\nabla f_t\|_2^2 \mathbb{1}_{F_{\alpha,t}^c} \mathbb{1}_{E_t} \right] dt &\leq \mathbb{E} \left[\int_0^{1-\delta} t \|\nabla f_t\|_2^2 dt \right] \\
 &\leq \mathbb{E} \left[\int_0^{\sqrt{1-\delta}} t \|\nabla f_t\|_2^2 dt \right] \\
 &\leq C_1 \text{Var}(f) \left(\sum_i I_i(f)^2 \right)^{\varepsilon \delta} \\
 &\leq C_1 \text{Var}(f) \alpha^{\log(\sum_i I_i(f)^2) \log\left(\frac{\sum_i I_i(f)^2}{2 \sum_i I_i(f)^2 + e}\right)^{-1}}
 \end{aligned}$$

όπου $C_1 > 0$ είναι κάποια θετική σταθερά. Αφού όμως $\sum_i I_i(f)^2 \leq 1/2$ έπεται ότι ο εκθέτης

$\log(\sum_i I_i(f)^2) \log\left(\frac{\sum_i I_i(f)^2}{2\sum_i I_i(f)^{2+e}}\right)^{-1}$ φράσσεται από πάνω από 1/3. Αφού τώρα $\alpha < 1$, έχουμε ότι τελικά

$$(4.3.14) \quad \int_0^{1-\delta} t \mathbb{E} \left[\|\nabla f_t\|_2^2 \mathbb{1}_{F_{\alpha,t}^c} \mathbb{1}_{E_t} \right] dt \leq C_1 \alpha^{1/3} \text{Var}(f).$$

Μπορούμε να φράξουμε το δεύτερο ολοκλήρωμα του δεξιού μέλους χρησιμοποιώντας τον ορισμό των $F_{\alpha,t}$ και E_t . Από τον ορισμό του $F_{\alpha,t}$ έχουμε ότι

$$(4.3.15) \quad \begin{aligned} \mathbb{E} \left[\|\nabla f_t\|_2^{2-2p} \|\nabla f_t\|_2^{2p} \mathbb{1}_{F_{\alpha,t}^c} \mathbb{1}_{E_t} \right] &= \mathbb{E} \left[\left(\|\nabla f_t\|_2^{2-2p} \mathbb{1}_{F_{\alpha,t}^c} \right) \|\nabla f_t\|_2^{2p} \mathbb{1}_{E_t} \right] \\ &\leq \alpha^{2-2p} \left(\log \left(2 + \frac{e}{\sum_i I_i(f)^2} \right) \right)^{1-p} \mathbb{E} \left[\|\nabla f_t\|_2^{2p} \mathbb{1}_{E_t} \right]. \end{aligned}$$

Επίσης, από τον ορισμό του E_t ,

$$\begin{aligned} \mathbb{E} \left[\|\nabla f_t\|_2^{2p} \mathbb{1}_{E_t} \right] &\leq 2 \mathbb{E} \left[\|\nabla f_t\|_2^{2p} \sup_{0 \leq s < t} \frac{1+f_s}{2} \mathbb{1}_{E_t} \right] \\ &\leq 2 \mathbb{E} \left[\|\nabla f_t\|_2^{2p} \sup_{0 \leq s < t} \frac{1+f_s}{2} \right] \\ &\leq 2 \mathbb{E}[\Psi_t] \leq 2 \text{Var}(f) \left(\log \left(2 + \frac{e}{\sum_i I_i(f)^2} \right) \right)^p, \end{aligned}$$

όπου στην τελευταία ανισότητα χρησιμοποιήσαμε την (4.3.2). Εισάγοντας την τελευταία ανισότητα στην (4.3.15) βλέπουμε ότι το δεύτερο ολοκλήρωμα της (4.3.13) μπορεί να φραχτεί από πάνω ως εξής:

$$\begin{aligned} \int_{1-\delta}^1 t \mathbb{E} \left[\|\nabla f_t\|_2^2 \mathbb{1}_{F_{\alpha,t}^c} \mathbb{1}_{E_t} \right] dt &= \int_{1-\delta}^1 t \mathbb{E} \left[\|\nabla f_t\|_2^{2-2p} \|\nabla f_t\|_2^{2p} \mathbb{1}_{F_{\alpha,t}^c} \mathbb{1}_{E_t} \right] dt \\ &\leq \delta \alpha^{2-2p} \text{Var}(f) \log \left(2 + \frac{e}{\sum_i I_i(f)^2} \right). \end{aligned}$$

Από την $\alpha < 1/e$ έχουμε ότι υπάρχει σταθερά $C_2 > 0$ τέτοια ώστε, σε κάθε περίπτωση, να ισχύει ότι

$$\delta < C_2 \frac{e^{-1} \log(1/\alpha)}{\log \left(2 + \frac{e}{\sum_i I_i(f)^2} \right)}.$$

Συνεπώς,

$$(4.3.16) \quad \int_{1-\delta}^1 t \mathbb{E} \left[\|\nabla f_t\|_2^2 \mathbb{1}_{F_{\alpha,t}^c} \mathbb{1}_{E_t} \right] dt \leq C_2 c^{-1} \alpha^{2-2p} \log(1/\alpha) \text{Var}(f).$$

Τελικά, συνδυάζοντας τις (4.3.14) και (4.3.16) έχουμε ότι υπάρχει μια απόλυτη σταθερά $C :=$

$C_1 + C_2 c^{-1} > 0$ για την οποία ισχύει ότι

$$(4.3.17) \quad \int_0^1 t \mathbb{E} \left[\|\nabla f_t\|_2^2 \mathbb{1}_{F_{\alpha,t}^c} \mathbb{1}_{E_t} \right] dt \leq C(\alpha^{1/3} + \alpha^{2-2p}) \log(1/\alpha) \text{Var}(f).$$

Εισάγοντας αυτή την ανισότητα στην (4.3.11) ολοκληρώνουμε την απόδειξη θεωρώντας τη συνάρτηση $\rho(x) = C(x^{1/3} + x^{2-2p}) \log(1/x)$ που ικανοποιεί τις ιδιότητες της εκφώνησης. $\square$

Η επόμενη πρόταση μας δείχνει ότι για να φτάσει ένα martingale σε ένα σημείο μακριά από την αρχική του θέση, πρέπει να έχει μεγάλη τετραγωνική κύμανση. Η απόδειξη, ωστόσο, χρησιμοποιεί τις ιδιότητες των αλμάτων του συγκεκριμένου martingale.

Πρόταση 4.3.3. Έστω $a \in [0, 1)$ και $0 \leq \vartheta \leq 1$ ένας B_t -μετρήσιμος χρόνος διακοπής για τον οποίο ισχύει ότι

$$\mathbb{P}[f_\vartheta \in [-a, a] \mid \vartheta < 1] \geq q$$

για κάποιο $q \in [0, 1]$. Τότε,

$$\mathbb{P} \left[V^{\vartheta \rightarrow 1} \geq \frac{1}{5} q(1-a)^2 \mid \vartheta < 1 \right] \geq \frac{1}{9} q(1-a)^2.$$

Απόδειξη. Έστω $x > 0$ ένας θετικός αριθμός ο οποίος θα επιλεγεί παρακάτω, και έστω

$$\sigma = \inf\{t > \vartheta : V^{\vartheta \rightarrow t} \geq x^2\} \wedge 1$$

η πρώτη χρονική στιγμή κατά την οποία η τετραγωνική κύμανση ξεπερνάει το x^2 . Αφού η τετραγωνική κύμανση αυξάνεται μόνο όταν η f_t κάνει άλμα και η πιθανότητα αυτή να έχουμε άλμα τη χρονική στιγμή $t = 1$ είναι μηδενική, το ενδεχόμενο $\{V^{\vartheta \rightarrow 1} \geq x^2 \mid \vartheta < 1\}$ συμπίπτει με το $\{\sigma < 1 \mid \vartheta < 1\}$, το οποίο με την σειρά του συνεπάγεται ότι $|f_\sigma| < 1$:

$$(4.3.18) \quad \mathbb{P} \left[V^{\vartheta \rightarrow 1} \geq x^2 \mid \vartheta < 1 \right] \geq \mathbb{P} [|f_\sigma| < 1 \mid \vartheta < 1].$$

Επομένως, αναγώμαστε στο να φράξουμε την πιθανότητα $|f_\sigma| < 1$. Αυτό μπορεί να γίνει κοιτάζοντας τη δεύτερη ροπή $\mathbb{E}[(f_\sigma - \mathbb{F}_\vartheta)^2]$: Από τη μία πλευρά, αφού $f_\vartheta \in [-a, a]$ με πιθανότητα τουλάχιστον q όταν $\vartheta < 1$, έχουμε ότι

$$\begin{aligned} \mathbb{E}[(f_\sigma - f_\vartheta)^2 \mid \vartheta < 1] &\geq \mathbb{E}[(f_\sigma - f_\vartheta)^2 \mid \{\vartheta < 1\} \cap \{|f_\sigma| = 1\}] \mathbb{P}[|f_\sigma| = 1 \mid \vartheta < 1] \\ &\geq q(1-a)^2 \mathbb{P}[|f_\sigma| = 1 \mid \vartheta < 1], \end{aligned}$$

που με την σειρά του μας δίνει ότι

$$(4.3.19) \quad \mathbb{P}[|f_\sigma| < 1 \mid \vartheta < 1] \geq 1 - \frac{\mathbb{E}[(f_\sigma - f_\vartheta)^2 \mid \vartheta < 1]}{q(1-a)^2}.$$

Από την άλλη πλευρά, η $\mathbb{E}[(f_\sigma - f_\vartheta)^2 \mid \vartheta < 1]$ φράσσεται αν λάβουμε υπόψιν μας το μέγεθος

των αλμάτων της f_t : Αφού η σ -άλγεβρα που παράγεται από το ενδεχόμενο $\vartheta < 1$ περιέχεται σε αυτήν που παράγεται από την B_ϑ ,

$$\mathbb{E}[(f_\sigma - f_\vartheta)^2 \mid \vartheta < 1] = \mathbb{E}[\mathbb{E}[(f_\sigma - f_\vartheta)^2 \mid B_\vartheta] \mid \vartheta < 1].$$

Αν όμως δεσμεύσουμε στην B_ϑ , η διαδικασία f_t είναι martingale για κάθε $t \in [\vartheta, \sigma]$, και άρα από την (2.1.2) έχουμε ότι

$$\begin{aligned} \mathbb{E}[\mathbb{E}[(f_\sigma - f_\vartheta)^2 \mid B_\vartheta] \mid \vartheta < 1] &= \mathbb{E}[\mathbb{E}[(f_\sigma - f_\vartheta)^2 \mid B_\vartheta] \mid \vartheta < 1] \\ &= \mathbb{E}[(f_\sigma - f_\vartheta)^2 \mid \vartheta < 1] \\ &= \mathbb{E}[V^{\vartheta \rightarrow \sigma} \mid \vartheta < 1] \\ &= \mathbb{E}[V^{\vartheta \rightarrow \sigma^-} + (\Delta f_\sigma)^2 \mid \vartheta < 1]. \end{aligned}$$

Τώρα, από τον ορισμό τώρα του σ έχουμε ότι $V^{\vartheta \rightarrow \sigma^-} \leq x^2$, και αφού όλα τα άλματα είναι φραγμένα από 2 παίρνουμε

$$\begin{aligned} (4.3.20) \quad \mathbb{E}[(f_\sigma - f_\vartheta)^2 \mid \vartheta < 1] &\leq x^2 + \mathbb{E}[(\Delta f_\sigma)^2 \mid \vartheta < 1] \\ &\leq x^2 + x^2 \mathbb{P}[\Delta f_\sigma < x \mid \vartheta < 1] + 4 \mathbb{P}[\Delta f_\sigma \geq x \mid \vartheta < 1] \\ &= x^2 + x^2 (1 - \mathbb{P}[\Delta f_\sigma \geq x \mid \vartheta < 1]) + 4 \mathbb{P}[\Delta f_\sigma \geq x \mid \vartheta < 1] \\ &= 2x^2 + \mathbb{P}[\Delta f_\sigma \geq x \mid \vartheta > 1] (4 - x^2) \end{aligned}$$

Συνδυάζοντας την (4.3.20) με τις (4.3.18) και (4.3.19), παίρνουμε ότι

$$\mathbb{P}[V^{\vartheta \rightarrow 1} \geq x^2 \mid \vartheta < 1] \geq 1 - \frac{2x^2 + \mathbb{P}[\Delta f_\sigma \geq x \mid \vartheta < 1] (4 - x^2)}{q(1-a)^2}.$$

Αφού τώρα $\{\Delta f_\sigma \geq x \mid \vartheta < 1\} \subset \{V^{\vartheta \rightarrow 1} \geq x^2 \mid \vartheta < 1\}$, παίρνουμε ότι

$$\mathbb{P}[V^{\vartheta \rightarrow 1} \geq x^2 \mid \vartheta < 1] \geq 1 - \frac{2x^2 + \mathbb{P}[V^{\vartheta \rightarrow 1} \geq x^2 \mid \vartheta < 1] (4 - x^2)}{q(1-a)^2}.$$

Λύνοντας ως προς την πιθανότητα $\mathbb{P}[V^{\vartheta \rightarrow 1} \geq x^2 \mid \vartheta > 1]$, έχουμε ότι

$$\mathbb{P}[V^{\vartheta \rightarrow 1} \geq x^2 \mid \vartheta < 1] \geq 1 - \frac{4 + x^2}{q(1-a)^2 + 4 - x^2}$$

και για $x^2 = \frac{4q(1-a)^2}{16+q(1-a)^2} \geq \frac{1}{5}q(1-a)^2$ συμπεραίνουμε ότι

$$\mathbb{P}\left[V^{\vartheta \rightarrow 1} \geq \frac{1}{5}q(1-a)^2 \mid \vartheta < 1\right] \geq \frac{1}{9}q(1-a)^2$$

όπως θέλαμε. □

Με αυτές τις δύο προτάσεις είμαστε τώρα έτοιμοι να αποδείξουμε ότι

$$\mathbb{P}[\tau_\alpha < 1 \mid \vartheta < 1] \geq C \operatorname{Var}(f).$$

Από τη μία, λόγω του Λήμματος 4.3.2, η μέση τιμή αύξησης της τετραγωνικής κύμανσης από τη χρονική στιγμή ϑ μέχρι τη χρονική στιγμή 1 πρέπει να είναι μικρή: Έστω

$$A_{\alpha,\vartheta} = \bigcap_{t \in (\vartheta, 1]} F_{\alpha,t}^c$$

το ενδεχόμενο η κλίση να είναι μικρή σε όλο το χρονικό διάστημα ϑ έως 1. Τότε, έχουμε ότι

$$\begin{aligned} \mathbb{E}[V^{\vartheta \rightarrow 1} 1_{A_{\alpha,\vartheta}} \mid \vartheta < 1] &\leq \mathbb{E}\left[\sum_i \sum_{t \in J_i \cap (\vartheta, 1]} (2\partial_i f_t)^2 1_{F_{\alpha,t}^c} \mid \vartheta < 1\right] \\ &\leq \mathbb{E}\left[\sum_i \sum_{t \in J_i \cap (\vartheta, 1]} (2\partial_i f_t)^2 1_{F_{\alpha,t}^c} 1_{E_t} \mid \vartheta < 1\right] \\ &= \frac{\mathbb{E}[\sum_i \sum_{t \in J_i \cap (\vartheta, 1]} (2\partial_i f_t)^2 1_{F_{\alpha,t}^c} 1_{E_t}]}{\mathbb{P}[\vartheta < 1]} \\ &= \frac{\mathbb{E}[V_\alpha]}{\mathbb{P}[\vartheta < 1]} \leq 4\rho(\alpha), \end{aligned}$$

όπου στην τελευταία ανισότητα χρησιμοποιήσαμε τις (4.3.10) και (4.3.5). Από την άλλη, χρησιμοποιώντας την Πρόταση 4.3.3 για $q = 1/2$ και $a = 1/2$ βλέπουμε ότι η συνολική αύξηση της τετραγωνικής κύμανσης είναι μεγάλη με πιθανότητα:

$$\mathbb{P}\left[V^{\vartheta \rightarrow 1} > \frac{1}{40} \mid \vartheta < 1\right] \geq \frac{1}{72}.$$

Έχουμε επομένως ότι

$$\begin{aligned} \mathbb{E}[V^{\vartheta \rightarrow 1} 1_{A_{\alpha,\vartheta}} \mid \vartheta < 1] &\geq \frac{1}{40} \mathbb{P}\left[\left\{V^{\vartheta \rightarrow 1} > \frac{1}{40}\right\} \cap A_{\alpha,\vartheta} \mid \vartheta < 1\right] \\ &\geq \frac{1}{40} \left(\mathbb{P}\left[V^{\vartheta \rightarrow 1} > \frac{1}{40} \mid \vartheta < 1\right] - \mathbb{P}[A_{\alpha,\vartheta}^c \mid \vartheta < 1]\right). \end{aligned}$$

Λύνοντας τώρα ως προς $\mathbb{P}[A_{\alpha,\vartheta}^c \mid \vartheta < 1]$ παίρνουμε ότι

$$\begin{aligned} \mathbb{P}[A_{\alpha,\vartheta}^c \mid \vartheta < 1] &\geq \mathbb{P}\left[V^{\vartheta \rightarrow 1} > \frac{1}{40} \mid \vartheta < 1\right] \geq \frac{1}{72} - 40\mathbb{P}\left[V^{\vartheta \rightarrow 1} > \frac{1}{40}\right] \\ &\geq \frac{1}{72} - C\rho(\alpha). \end{aligned}$$

Αφού όμως $\lim_{\alpha \rightarrow 0} \rho(\alpha) = 0$, η τελευταία ποσότητα είναι μεγαλύτερη από κάποια σταθερά c αν το α είναι αρκετά μικρό. Όμως το ενδεχόμενο $A_{\alpha,\vartheta}^c$ σημαίνει ότι για κάποια χρονική στιγμή

$t_* > \vartheta$ το μέτρο της κλίσης $\|\nabla f_{t_*}\|_2$ ήταν μεγαλύτερο από $\alpha \sqrt{\log\left(2 + \frac{e}{\sum_i I_i(f)}\right)}$ ενώ, παράλληλα, όταν $\vartheta < 1$ έχουμε ότι $\sup_{0 \leq s \leq \vartheta} \frac{1+f_s}{2} \geq 1/2$. Συνδυάζοντας τα παραπάνω παίρνουμε ότι

$$\Psi_{t_*} \geq \frac{1}{2} \alpha^{2p} \left(\log\left(2 + \frac{e}{\sum_i I_i(f)}\right) \right)^p.$$

Έτσι, τελικά έχουμε ότι

$$\mathbb{P}[\tau_\alpha < 1 \mid \vartheta < 1] \geq \mathbb{P}[A_{\alpha, \vartheta}^c \mid \vartheta < 1] > c,$$

όπως επιθυμούσαμε.

4.3.2 Περίπτωση 2η-Μεγάλα άλματα

Υποθέτουμε τώρα ότι $\mathbb{P}[f_\vartheta \in [0, 1/2] \mid \vartheta < 1] > 1/2$, το οποίο σημαίνει ότι με μεγάλη πιθανότητα η f_t κάνει ένα μεγάλο άλμα τη χρονική στιγμή ϑ (που ορίστηκε στην (4.3.4)) σε κάποια τιμή μεγαλύτερη του $1/2$:

$$(4.3.21) \quad \mathbb{P}\left[\Delta f_\vartheta \geq \frac{1}{2} \mid \vartheta < 1\right] \geq \frac{1}{2}.$$

Το επόμενο λήμμα, το οποίο είναι αντίστοιχο με το Λήμμα 4.3.2, δείχνει ότι το μεγαλύτερο μέρος της τετραγωνικής κύμανσης αποκτάται από τις χρονικές στιγμές κατά τις οποίες, αμέσως μετά από κάποιο άλμα της στοχαστικής διαδικασίας, η κλίση ήταν μεγάλη. Για σταθερό $\alpha > 0$, την τιμή του οποίου θα επιλέξουμε παρακάτω, ορίζουμε

$$(4.3.22) \quad H_{\alpha, t} = \left\{ \|\nabla f_t\|_2 \geq \alpha \sqrt{\log\left(2 + \frac{e}{\sum_i I_i(f)^2}\right)} \right\}$$

το ενδεχόμενο η νόρμα της κλίσης να είναι μεγάλη ακριβώς τη χρονική στιγμή t . Επίσης, ορίζουμε

$$(4.3.23) \quad U_\alpha = \sum_{i=1}^n \sum_{t \in J_i \cap [0, 1]} (2t \partial_i f_t)^2 \mathbf{1}_{H_{\alpha, t}^c} \mathbf{1}_{f_t \geq 0}$$

να είναι η τετραγωνική κύμανση που έχει αποκτηθεί στις χρονικές στιγμές όπου η τιμή της f είναι μεγάλη αλλά η κλίση είναι μικρή.

Πρόταση 4.3.4. Έστω $0 < \alpha < 1/e$ και έστω ότι ισχύει η (4.3.2). Τότε, υπάρχει μια συνάρτηση $\rho : [0, 1] \rightarrow \mathbb{R}$ με $\lim_{x \rightarrow 0} \rho(x) = 0$, για την οποία ισχύει ότι

$$(4.3.24) \quad \mathbb{E}[U_\alpha] \leq \rho(\alpha) \text{Var}(f).$$

Απόδειξη. Για κάθε $i \in [n]$ θεωρούμε τη συνάρτηση $s_i : \mathbb{R}^n \rightarrow \mathbb{R}^n$ που αλλάζει το πρόσημο της i -οστής συντεταγμένης του x . Παρατηρούμε ότι αν η $B_t^{(i)}$ κάνει άλμα τη χρονική στιγμή t ,

τότε $B_{t-} = s_i(B_t)$. Έτσι, σε κάθε τέτοια χρονική στιγμή έχουμε

$$f(B_{t-}) = f(s_i(B_t)) \quad \text{και} \quad \|\nabla f(B_{t-})\|_2 = \|\nabla f(s_i(B_t))\|_2.$$

Θέτουμε

$$\beta = \alpha \sqrt{\log \left(2 + \frac{e}{\sum_i I_i(f)^2} \right)}.$$

Αφού οι B_t και $s_i(B_t)$ έχουν την ίδια κατανομή για κάθε $i \in [n]$, και αφού η $\partial_i f(x)$ δεν εξαρτάται από την i -οστή συντεταγμένη του x , έχουμε ότι

$$\begin{aligned} \mathbb{E}[U_\alpha] &= \mathbb{E} \left[\sum_{i=1}^n \sum_{t \in J_i \cap [0,1]} (2t \partial_i f_t)^2 \mathbf{1}_{\|\nabla f(B_t)\|_2 \leq \beta} \mathbf{1}_{f_{B_t} \geq 0} \right] \\ &= \mathbb{E} \left[\sum_{i=1}^n \sum_{t \in J_i \cap [0,1]} (2t \partial_i f_t)^2 \mathbf{1}_{\|\nabla f(s_i(B_t))\|_2 \leq \beta} \mathbf{1}_{f_{s_i(B_t)} \geq 0} \right] \\ &= \mathbb{E} \left[\sum_{i=1}^n \sum_{t \in J_i \cap [0,1]} (2t \partial_i f_t)^2 \mathbf{1}_{\|\nabla f_{t-}\|_2 \leq \beta} \mathbf{1}_{f_{t-} \geq 0} \right] \\ &\leq \mathbb{E} \left[\sum_{i=1}^n \sum_{t \in J_i \cap [0,1]} (2t \partial_i f_t)^2 \mathbf{1}_{\|\nabla f_{t-}\|_2 \leq \beta} \mathbf{1}_{\sup_{0 \leq s < t} f_s \geq 0} \right] \\ &= \mathbb{E}[V_\alpha], \end{aligned}$$

για την V_α που ορίστηκε στην (4.3.8). Τώρα, ο ισχυρισμός του λήμματος προκύπτει από το Λήμμα 4.3.2. $\square$

Με αυτό το λήμμα στα χέρια μας, μπορούμε να αποδείξουμε ότι

$$\mathbb{P}[\tau_\alpha < 1] \geq C \operatorname{Var}(f).$$

Θεωρούμε το ενδεχόμενο να ισχύουν ταυτόχρονα οι $\|\nabla f_\theta\|_2 \leq \alpha \sqrt{\log \left(2 + \frac{e}{\sum_i I_i(f)^2} \right)}$ και $\Delta f_\theta \geq 1/2$ τη χρονική στιγμή θ . Αφού το θ είναι η πρώτη χρονική στιγμή όπου $f_t \geq 0$, και αφού με πιθανότητα 1 δεν υπάρχει άλμα τη χρονική στιγμή 1, αυτό το ενδεχόμενο συμβάλλει τουλάχιστον $1/4$ στο U_α , και άρα

$$\begin{aligned} \mathbb{P} \left[U_\alpha \geq \frac{1}{4} \right] &\geq \mathbb{P} \left[\left\{ \|\nabla f_\theta\|_2 \leq \alpha \sqrt{\log \left(2 + \frac{e}{\sum_i I_i(f)^2} \right)} \right\} \cap \left\{ \Delta f_\theta \geq \frac{1}{2} \right\} \right] \\ &= \mathbb{P} \left[\left\{ \|\nabla f_\theta\|_2 \leq \alpha \sqrt{\log \left(2 + \frac{e}{\sum_i I_i(f)^2} \right)} \right\} \mid \Delta f_\theta \geq \frac{1}{2} \right] \mathbb{P} \left[\Delta f_\theta \geq \frac{1}{2} \right] \\ &\geq \mathbb{P} \left[\left\{ \|\nabla f_\theta\|_2 \leq \alpha \sqrt{\log \left(2 + \frac{e}{\sum_i I_i(f)^2} \right)} \right\} \mid \Delta f_\theta \geq \frac{1}{2} \right] \cdot \frac{1}{8} \operatorname{Var}(f), \end{aligned}$$

όπου στην τελευταία ανισότητα χρησιμοποιήσαμε τις (4.3.21) και (4.3.5). Απο την άλλη πλευρά, από την ανισότητα Markov και το Λήμμα 4.3.4, αυτή η πιθανότητα φράσσεται απο πάνω απο

$$\mathbb{P}\left[U_\alpha \geq \frac{1}{4}\right] \leq \frac{\mathbb{E}[U_\alpha]}{1/4} \leq 4\rho(\alpha)\text{Var}(f).$$

Συνδυάζοντας τις δύο αυτές ανισότητες, παίρνουμε τελικά ότι

$$\begin{aligned} \mathbb{P}[H_{\alpha,\vartheta} \mid \Delta f_\vartheta \geq 1/2] &= \mathbb{P}\left[\|\nabla f_\vartheta\|_2 \leq \alpha \sqrt{\log\left(2 + \frac{e}{\sum_i I_i(f)^2}\right)} \mid \Delta f_\vartheta \geq \frac{1}{2}\right] \\ &\leq 4\rho(\alpha). \end{aligned}$$

Παίρνοντας αρκετά μικρό α ώστε το δεξί μέλος της παραπάνω ανισότητας να είναι μικρότερο από $1/2$, έχουμε ότι

$$\mathbb{P}\left[H_{\alpha,\vartheta} \mid \Delta f_\vartheta \geq \frac{1}{2}\right] > \frac{1}{2}.$$

Γι' αυτήν την τιμή του α , κάτω απο το ενδεχόμενο $H_{\alpha,\vartheta} \cap \{\vartheta < 1\}$, η Ψ_ϑ (που ορίστηκε στην (4.3.1)) είναι μεγάλη:

$$\begin{aligned} \Psi_\vartheta &= \|\nabla f_\vartheta\|_2^{2p} \sup_{0 \leq s \leq \vartheta} \frac{1+f_s}{2} \geq \|\nabla f_\vartheta\|_2^{2p} \frac{1+f_\vartheta}{2} \\ &\geq \frac{1}{2} \alpha^{2p} \left(\log\left(2 + \frac{e}{\sum_i I_i(f)^2}\right)\right)^p, \end{aligned}$$

και άρα,

$$\begin{aligned} \mathbb{P}[\tau_\alpha < 1 \mid \vartheta < 1] &\geq \mathbb{P}\left[\tau_\alpha < 1 \mid \vartheta < 1 \text{ και } \Delta f_\vartheta \geq \frac{1}{2}\right] \mathbb{P}\left[\Delta f_\vartheta \geq \frac{1}{2} \mid \vartheta < 1\right] \\ &\geq \mathbb{P}\left[H_{\alpha,\vartheta} \mid \Delta f_\vartheta \geq \frac{1}{2}\right] \cdot \frac{1}{2} \geq \frac{1}{4} \end{aligned}$$

όπως θέλαμε. Για τη μεσαία ανισότητα χρησιμοποιήσαμε το γεγονός ότι το ενδεχόμενο $\{\Delta f_\vartheta \geq \frac{1}{2}\}$ περιέχει το $\{\vartheta < 1\}$ έως ένα σύνολο μέτρου 0 (που ισχύει γιατί με πιθανότητα 1 δεν έχουμε άλμα τη χρονική στιγμή 1) και για την τελευταία ανισότητα χρησιμοποιήσαμε την (4.3.21). $\square$

ΚΕΦΑΛΑΙΟ 5

Η ανισότητα επιρροών του Talagrand και η ευστάθειά της

5.1 Η ανισότητα επιρροών του Talagrand

Οι αποδείξεις των Θεωρημάτων 2.4.5 και 2.5.1 κινούνται στο ίδιο πνεύμα με την απόδειξη του Θεωρήματος 2.4.6, και απαιτούν πάλι να δώσουμε φράγματα για το κέρδος στην τετραγωνική κύμανση. Όμως, χρειάζεται πρόσθετη φροντίδα για να φράξουμε το μέγεθος καθεμιάς από τις διαδικασίες επιρροής $f_t^{(i)}$.

Ορίζουμε πρώτα διάφορες ποσότητες που θα παίξουν κεντρικό ρόλο στις αποδείξεις. Για σταθερό $0 < \alpha \leq 1$ που η τιμή του θα επιλεγεί αργότερα, θέτουμε

$$(5.1.1) \quad F_\alpha = \left\{ \exists t \in [0, 1], \exists i \in [n] \mid t \in J_i \text{ και } f_t^{(i)} \geq \alpha \right\}$$

το ενδεχόμενο να υπάρχει κάποια συντεταγμένη που να είχε μεγάλη παράγωγο τη στιγμή που έκανε άλμα. Ορίζουμε επίσης

$$\mathcal{Q}_\alpha^{(i)} = 2 \int_0^1 t \left(f_t^{(i)} \right)^2 \mathbb{1}_{\{f_t^{(i)} < \alpha\}} dt$$

και

$$(5.1.2) \quad \mathcal{Q}_\alpha = \sum_{i=1}^n \mathcal{Q}_\alpha^{(i)},$$

και θέτουμε

$$V_\alpha^{(i)} = \sum_{t \in J_i \cap [0, 1]} (2t \partial_i f_t)^2 \mathbb{1}_{\{f_t^{(i)} < \alpha\}}$$

και

$$(5.1.3) \quad V_\alpha = \sum_{i=1}^n V_\alpha^{(i)}.$$

Σημειώνουμε ότι αν $\alpha = 1$ τότε $Q_1 = V_1$ είναι απλώς η τετραγωνική κύμανση της f_t . Μπορούμε να σκεφτόμαστε την V_α ως την τετραγωνική κύμανση της διαδικασίας f_t , όπου όμως μεγάλα άλματα (δηλαδή, αυτά που είναι μεγαλύτερα από $t\alpha$) έχουν εξαιρεθεί. Η Q_α είναι το ολοκληρωτικό ανάλογο της V_α . Τέλος, ορίζουμε

$$(5.1.4) \quad \rho(x) = x \left(\log \frac{1}{x} + 2 \right).$$

Αντί να χρησιμοποιήσουμε το Θεώρημα 2.5.4 για να φράξουμε τις επιρροές, χρησιμοποιούμε το ακόλουθο λήμμα.

Λήμμα 5.1.1. Υπάρχει απόλυτη σταθερά $\gamma > 1$ τέτοια ώστε

$$(5.1.5) \quad \varphi_i(s) \leq \gamma I_i(f)^{1+s/(2\gamma)}$$

για κάθε $0 \leq s \leq \gamma$, όπου $\varphi_i(s) = \mathbb{E} \left[\left(f_{e^{-s}}^{(i)} \right)^2 \right]$ ορίζεται όπως στην (4.2.6).

Μπορούμε να αποδείξουμε αυτό το λήμμα μέσω της αρχής της υπερσυσταλτότητας (βλέπε, για παράδειγμα, [13] και [44, Πρόταση 9.25]). Θα δώσουμε όμως μια διαφορετική απόδειξη που βασίζεται στην ανάλυση της στοχαστικής διαδικασίας f_t . Ο λόγος είναι ότι μπορούμε να αναπτύξουμε αυτήν την ανάλυση περαιτέρω και να πάρουμε τα αποτελέσματα ευστάθειας. Διαισθητικά, αν θυμηθούμε και την εξίσωση (4.1.5), το λήμμα ισχυρίζεται ότι η ουσιαστική συνεισφορά της τετραγωνικής κύμανσης στη διασπορά της συνάρτησης συμβαίνει πολύ κοντά στη χρονική στιγμή 1.

Απόδειξη του Λήμματος 5.1.1. Έστω $\gamma > 1$ το οποίο θα επιλεγεί αργότερα. Αρχίζουμε αποδεικνύοντας ότι υπάρχει μια σταθερά $c'_\gamma > 0$ τέτοια ώστε

$$(5.1.6) \quad \varphi_i(\gamma) \leq \gamma \varphi_i(0)^{1+c'_\gamma}.$$

Θυμηθείτε ότι $\psi_i(t) = \varphi_i(\log 1/t)$. Εφαρμόζοντας το Πρόταση 4.1.5 για τη συνάρτηση $f_t^{(i)}$ που ορίστηκε στην (4.2.2), βλέπουμε ότι η ψ_i ικανοποιεί την

$$(5.1.7) \quad \frac{d\psi_i}{dt} = 2t \mathbb{E} \left[\|\nabla f_t^{(i)}\|_2^2 \right].$$

Μπορούμε να φράξουμε το δεξιό μέλος της (5.1.7) χρησιμοποιώντας το Λήμμα 2.6.2: Επιλέγοντας $g = f^{(i)}$ και $x = B_t$ στην εξίσωση (2.6.7) και αντικαθιστώντας αυτήν στην εξίσωση (5.1.7), έχουμε ότι

$$\frac{d\psi_i}{dt} \leq 2t C(t) \mathbb{E} \left[\left(f_t^{(i)} \right)^2 \log \left(\frac{e}{\left(f_t^{(i)} \right)^2} \right) \right].$$

Αφού η $C(t)$ είναι συνεχής, υπάρχει $L > 0$ τέτοιος ώστε $C(t) \leq L$ για κάθε $t \in [0, 1/2]$, άρα

$$(5.1.8) \quad \begin{aligned} \frac{d\psi_i}{dt} &\leq L \mathbb{E} \left[\left(f_t^{(i)} \right)^2 \log \left(\frac{e}{\left(f_t^{(i)} \right)^2} \right) \right] \\ (\text{ανισότητα Jensen}) &\leq L \mathbb{E} \left[\left(f_t^{(i)} \right)^2 \right] \log \left(\frac{e}{\mathbb{E} \left[\left(f_t^{(i)} \right)^2 \right]} \right) \\ &= L \psi_i(t) \log \left(\frac{e}{\psi_i(t)} \right). \end{aligned}$$

Εφαρμόζοντας το Λήμμα 4.2.2 με $C = L$ και $K = e$ βλέπουμε ότι υπάρχει κάποια χρονική στιγμή $t_0 \leq 1/2$ και μια σταθερά K_1 τέτοια ώστε, για κάθε $t \in [0, t_0]$,

$$\psi_i(t) \leq K_1 \psi_i(0) e^{1-Lt} = K_1 I_i(f)^{2-2Lt},$$

όπου στην τελευταία ισότητα χρησιμοποιήσαμε την εξίσωση (4.2.5) και το γεγονός ότι $\psi_i(0) = (f_0^{(i)})^2$. Εφόσον $\varphi_i(s) = \psi_i(e^{-s})$, αν $\gamma \geq \log(1/t_0)$ τότε

$$\begin{aligned} \varphi_i(\gamma) &\leq K_1 I_i(f)^{2-2Le^{-\gamma}} \\ &= K_1 I_i(f)^{1+(1-2Le^{-\gamma})}. \end{aligned}$$

Θέτοντας $c'_\gamma = 1 - 2Le^{-\gamma}$ και επιλέγοντας το γ μεγαλύτερο από K_1 παίρνουμε το ζητούμενο: η εξίσωση 5.1.6 ισχύει διότι $\varphi_i(0) = \mathbb{E} \left(f_1^{(i)} \right)^2 = \mathbb{E} \left[\left(f^{(i)} \right)^2 \right] = I_i(f)$ από την (2.1.2).

Χρησιμοποιώντας την (5.1.6) μαζί με την λογαριθμική κυρτότητα που εξασφαλίζει το Λήμμα 4.2.1, για κάθε $0 \leq s \leq \gamma$ μπορούμε να φράξουμε την $\varphi_i(s)$ ως εξής:

$$\begin{aligned} \varphi_i(s) &= \varphi_i \left(\left(1 - \frac{s}{\gamma} \right) \cdot 0 + \frac{s}{\gamma} \cdot \gamma \right) \\ &\leq \varphi_i(0)^{1-s/\gamma} \varphi_i(\gamma)^{s/\gamma} \\ &\leq \varphi_i(0)^{1-s/\gamma} \left(\gamma \varphi_i(0)^{1+c'_\gamma} \right)^{s/\gamma} \\ &= \gamma I_i(f)^{1+c_\gamma s} \end{aligned}$$

όπως θέλαμε, με $c_\gamma = c'_\gamma/\gamma = (1 - 2Le^{-\gamma})/\gamma$. Κατόπιν, το λήμμα προκύπτει αν πάρουμε το γ αρκετά μεγάλο ώστε να ικανοποιούνται οι $\gamma \geq \log(2L)$ και $c'_\gamma \geq 1/2$. $\square$

Οι επόμενες δύο προτάσεις είναι κατά κάποιον τρόπο ανάλογες με τα Λήμματα 4.3.2 και 3.2.1.

Πρόταση 5.1.2. Έστω $0 < \alpha \leq 1$. Τότε,

$$\mathbb{E}[Q_\alpha] \leq 4\gamma^2 \rho(\alpha) T(f),$$

όπου γ είναι η απόλυτη σταθερά από το Λήμμα 5.1.1 και η $T(f)$ όπως ορίστηκε στην (2.5.1).

Απόδειξη. Εφόσον $Q_\alpha = \sum_{i=1}^n Q_\alpha^{(i)}$ όπως στην (5.1.2), αρκεί να δείξουμε ότι

$$\mathbb{E}[Q_\alpha^{(i)}] \leq 4\gamma^2 \rho(\alpha) \frac{I_i(f)}{1 + \log(1/I_i(f))}.$$

Ορίζοντας $\tilde{\varphi}_i(s) = \mathbb{E}\left[\left(f_{e^{-s}}^{(i)}\right)^2 \mathbb{1}_{\{f_{e^{-s}}^{(i)} < \alpha\}}\right]$, με μια αλλαγή μεταβλητής παίρνουμε

$$(5.1.9) \quad \mathbb{E}[Q_\alpha^{(i)}] \leq 2 \int_0^\infty e^{-2s} \tilde{\varphi}_i(s) ds.$$

Θέτουμε $\tau = \frac{\log(1/\alpha)}{2 + \frac{1}{2\gamma} \log(1/I_i(f))}$. Υποθέτουμε αρχικά ότι $\tau \leq \frac{1}{2}\gamma$. Τότε, το ολοκλήρωμα στην (5.1.9) χωρίζεται σε τρία μέρη:

$$(5.1.10) \quad \mathbb{E}[Q_\alpha^{(i)}] \leq 2 \int_0^\tau e^{-2s} \tilde{\varphi}_i(s) ds + 2 \int_\tau^\gamma e^{-2s} \tilde{\varphi}_i(s) ds + 2 \int_\gamma^\infty e^{-2s} \tilde{\varphi}_i(s) ds.$$

Για το πρώτο ολοκλήρωμα στο δεξιό μέλος, γράφουμε

$$\tilde{\varphi}_i(s) \leq \alpha \mathbb{E}\left[\left|f_{e^{-s}}^{(i)}\right|\right] = \alpha \mathbb{E}\left[f_{e^{-s}}^{(i)}\right] = \alpha f_0^{(i)} = \alpha I_i(f).$$

Έτσι,

$$(5.1.11) \quad \int_0^\tau e^{-2s} \tilde{\varphi}_i(s) ds \leq \alpha \tau I_i(f)$$

$$(\text{από την επιλογή του } \tau) \leq 2\gamma \frac{I_i(f)}{4\gamma + \log(1/I_i(f))} \alpha \log(1/\alpha)$$

$$(\text{αφού } \gamma > 1) \leq 2\gamma \frac{I_i(f)}{1 + \log(1/I_i(f))} \alpha \log(1/\alpha).$$

Για το δεύτερο και το τρίτο ολοκλήρωμα, χρησιμοποιούμε το γεγονός ότι $\tilde{\varphi}_i(s) \leq \varphi_i(s)$ για κάθε s . Από το Λήμμα 5.1.1, για $s \in [\tau, \gamma]$ έχουμε ότι $\tilde{\varphi}_i(s) \leq \gamma I_i(f)^{1+s/(2\gamma)}$. Επομένως, το δεύτερο ολοκλήρωμα φράσσεται από

$$(5.1.12) \quad \int_\tau^\gamma e^{-2s} \tilde{\varphi}_i(s) ds \leq \gamma \int_\tau^\gamma e^{-2s} I_i(f)^{1+s/(2\gamma)} ds$$

$$\leq \gamma \int_\tau^\infty e^{-2s} I_i(f)^{1+s/(2\gamma)} ds$$

$$= \gamma I_i(f) \int_\tau^\infty e^{s(\frac{1}{2\gamma} \log I_i(f) - 2)} ds$$

$$\leq \gamma \frac{I_i(f)}{2 + \frac{1}{2\gamma} \log(1/I_i(f))} e^{\tau(\frac{1}{2\gamma} \log I_i(f) - 2)}$$

$$\leq 2\gamma^2 \frac{I_i(f)}{1 + \log(1/I_i(f))} \alpha.$$

Για το τρίτο ολοκλήρωμα, εφόσον $\gamma > 1$ και $\tau \leq \frac{1}{2}\gamma$, βλέπουμε εύκολα ότι $\int_{\gamma}^{\infty} e^{-2s} ds \leq \int_{\tau}^{\gamma} e^{-2s} ds$. Χρησιμοποιώντας το γεγονός ότι η $\varphi_i(s)$ είναι φθίνουσα συνάρτηση του s (κάτι που προκύπτει από την (4.2.7)), βλέπουμε ότι

$$(5.1.13) \quad \int_{\gamma}^{\infty} e^{-2s} \tilde{\varphi}_i(s) ds \leq \int_{\tau}^{\gamma} e^{-2s} \tilde{\varphi}_i(s) ds.$$

Συνδυάζοντας τις (5.1.11), (5.1.12) και (5.1.13), στην περίπτωση που $\tau < \frac{1}{2}\gamma$ παίρνουμε ότι

$$\begin{aligned} \mathbb{E}[Q_{\alpha}^{(i)}] &\leq 2 \left(2\gamma \frac{I_i(f)}{1 + \log(1/I_i(f))} \alpha \log(1/\alpha) + (2+2)\gamma^2 \frac{I_i(f)}{1 + \log(1/I_i(f))} \alpha \right) \\ &= 4\gamma^2 \rho(\alpha) \frac{I_i(f)}{1 + \log(1/I_i(f))}. \end{aligned}$$

Τώρα υποθέτουμε ότι $\tau \geq \frac{1}{2}\gamma$. Τότε, το ολοκλήρωμα στην (5.1.9) χωρίζεται σε δύο μέρη:

$$\mathbb{E}[Q_{\alpha}^{(i)}] \leq 2 \int_0^{\tau} e^{-2s} \tilde{\varphi}_i(s) ds + 2 \int_{\tau}^{\infty} e^{-2s} \tilde{\varphi}_i(s) ds.$$

Πάλι, εφόσον η $\varphi_i(s)$ είναι φθίνουσα συνάρτηση του s και $\tau \geq \frac{1}{2}\gamma > \frac{1}{2}$, το δεύτερο ολοκλήρωμα είναι μικρότερο από το πρώτο, και έτσι, από την (5.1.11), βλέπουμε ότι

$$\begin{aligned} \mathbb{E}[Q_{\alpha}^{(i)}] &\leq 2 \cdot 2\gamma^2 \frac{I_i(f)}{1 + \log(1/I_i(f))} \alpha \log(1/\alpha) \\ &\leq 4\gamma^2 \rho(\alpha) \frac{I_i(f)}{1 + \log(1/I_i(f))} \end{aligned}$$

και σε αυτή την περίπτωση. □

Πρόταση 5.1.3. Αν $0 < \alpha < 1/8$ και $\mathbb{P}[F_{\alpha}] < 2^{-14} \text{Var}(f)$, τότε

$$(5.1.14) \quad \mathbb{P}\left[V_{\alpha} \geq \frac{1}{16}\right] > 2^{-14} \text{Var}(f).$$

Απόδειξη. Χωρίς περιορισμό της γενικότητας υποθέτουμε ότι $f_0 = \mathbb{E}[f] \leq 0$ (αν όχι, χρησιμοποιούμε την $-f$ αντί για την f , διότι η διασπορά και οι πιθανότητες $\mathbb{P}[V_{\alpha} \geq x]$ είναι οι ίδιες για τις δύο συναρτήσεις). Ορίζουμε $\tau = \inf\{0 \leq t \leq 1 \mid f_t \in (0, 2\alpha)\} \wedge 1$. Δεσμεύοντας στο ενδεχόμενο $\{\tau < 1\}$, για κάθε $x > 0$ έχουμε ότι

$$(5.1.15) \quad \mathbb{P}[[f]_1 \geq 1/16] \geq \mathbb{P}\left[[f]_1 \geq 1/16 \mid \tau < 1\right] \mathbb{P}[\tau < 1].$$

Ξεκινάμε φράσσοντας την πιθανότητα $\mathbb{P}[\tau < 1]$. Ορίζουμε $A = \{\exists t \in [0, 1] \text{ τ.ω. } f_t > 0\}$, και παρατηρούμε ότι $\{\tau < 1\} \subseteq A$. Με δεδομένο το ενδεχόμενο $A \setminus \{\tau < 1\}$, η διαδικασία f_t δεν επισκέφτηκε ποτέ το διάστημα $(0, 2\alpha)$ και όμως, σε κάποια χρονική στιγμή, έπιασε μια τιμή μεγαλύτερη από 0, άρα αναγκαστικά είχε ένα άλμα ασυνέχειας μεγέθους τουλάχιστον 2α . Όμως, ένα άλμα που συμβαίνει τη χρονική στιγμή t λόγω μιας ασυνέχειας στην $B_t^{(i)}$ έχει μέγεθος

$2t|\partial_i f_t|$, άρα $2t|\partial_i f_t| \geq 2\alpha$, απ' όπου έπεται ότι $f_t^{(i)} \geq |\partial_i f_t| \geq \alpha$. Συνεπώς, $A \cap \{\tau = 1\} \subseteq A \cap F_\alpha$, και άρα $A \cap F_\alpha^c \subseteq A \cap \{\tau < 1\} = \{\tau < 1\}$. Έπεται ότι

$$\mathbb{P}[\tau < 1] \geq \mathbb{P}[A \setminus F_\alpha] \geq \mathbb{P}[A] - \mathbb{P}[F_\alpha].$$

Για να φράξουμε την $\mathbb{P}[A]$, παρατηρούμε ότι $\{f_1 = 1\} \subseteq A$. Από την ιδιότητα martingale της f_t ,

$$f_0 = \mathbb{E}[f_1] = 2\mathbb{P}[f_1 = 1] - 1,$$

άρα

$$\begin{aligned} \mathbb{P}[A] &\geq \mathbb{P}[f_1 = 1] = \frac{1 + f_0}{2} = \frac{1 - f_0^2}{2(1 - f_0)} \\ &= \frac{\text{Var}(f)}{2(1 - f_0)} \geq \frac{1}{4} \text{Var}(f). \end{aligned}$$

Συνδυάζοντας το παραπάνω με την υπόθεση ότι $\mathbb{P}[F_\alpha] < 2^{-14} \text{Var}(f)$, παίρνουμε

$$(5.1.16) \quad \mathbb{P}[\tau < 1] \geq \frac{1}{8} \text{Var}(f).$$

Στη συνέχεια, φράσσουμε την πιθανότητα $\mathbb{P}[[f]_1 \geq 1/16 \mid \tau < 1]$, συνδέοντας την τετραγωνική κύμανση με τη διασπορά της f_t .

Ορίζουμε $X_t = f_{\tau+t(1-\tau)}$. Δεσμευμένη στο ενδεχόμενο $\tau < 1$, η X_t είναι πάλι martingale, με $X_0 \in (0, 2\alpha) \subseteq [-1/2, 1/2]$, $X_1 \in \{-1, 1\}$, και $X_t \in [-1, 1]$ για κάθε $t \in [0, 1]$. Εφαρμόζοντας το Λήμμα 3.2.1 με $a = 1/2$ και $c_1 = 1/4$, παίρνουμε

$$\mathbb{P}\left[[f]_1 \geq \frac{1}{16} \mid \tau < 1\right] \geq \frac{1}{2^{10}}.$$

Συνδυάζοντας με τις (5.1.15) και (5.1.16) παίρνουμε

$$\mathbb{P}[[f]_1 \geq 1/16] \geq \frac{1}{2^{13}} \text{Var}(f).$$

Από τον ορισμό των F_α και V_α , με δεδομένο το ενδεχόμενο F_α^c έχουμε ότι $V_\alpha = [f]_1$, δηλαδή ολόκληρη η τετραγωνική κύμανση προέρχεται από μικρά άλματα. Έπεται ότι

$$\begin{aligned} \mathbb{P}\left[V_\alpha \geq \frac{1}{16}\right] &\geq \mathbb{P}\left[\left\{[f]_1 \geq \frac{1}{16}\right\} \cap F_\alpha^c\right] \\ &\geq \mathbb{P}\left[[f]_1 \geq \frac{1}{16}\right] - \mathbb{P}[F_\alpha] \\ &\geq \frac{1}{2^{13}} \text{Var}(f) - \frac{1}{2^{14}} \text{Var}(f) = \frac{1}{2^{14}} \text{Var}(f). \end{aligned}$$

□

5.2 Απόδειξη του Θεωρήματος 2.4.5

Απόδειξη του Θεωρήματος 2.4.5. Έστω γ η σταθερά από τη διατύπωση του Λήμματος 5.1.1. Από την Πρόταση 5.1.2, για κάθε $0 < \alpha \leq 1$, έχουμε

$$\mathbb{E}[Q_\alpha] \leq 4\gamma^2 \rho(\alpha) T(f).$$

Επιλέγοντας $\alpha = 1$ παίρνουμε $Q_1 = 2 \sum_{i=1}^n \int_0^1 t (f_t^{(i)})^2 dt$, διότι οι παράγωγοι φράσσονται από 1. Η μέση τιμή αυτής της παράστασης, όπως είδαμε στην (4.2.4), είναι μεγαλύτερη από $\text{Var}(f)$. Έτσι, συμπεραίνουμε ότι

$$\text{Var}(f) \leq \mathbb{E}[Q_1] \leq 4\gamma^2 \rho(1) T(f) = 8\gamma^2 T(f).$$

□

5.3 Απόδειξη του Θεωρήματος 2.5.1

Χρησιμοποιώντας τις Προτάσεις 5.1.2 και 5.1.3, μπορούμε να πάρουμε το ακόλουθο λήμμα.

Λήμμα 5.3.1. Έστω γ η σταθερά από το Λήμμα 5.1.1, έστω F_α το ενδεχόμενο που ορίστηκε στην (5.1.1), και ας υποθέσουμε ότι ο $0 < \alpha < 1/16$ είναι αρκετά μικρός ώστε να ισχύει ότι $4\gamma^2 \rho(\alpha) T(f) \leq \frac{1}{2^{14}} \text{Var}(f)$. Τότε,

$$\mathbb{P}[F_\alpha] \geq \frac{1}{2^{14}} \text{Var}(f).$$

Απόδειξη. Υποθέτουμε προς άτοπο ότι $\mathbb{P}[F_\alpha] < \frac{1}{2^{14}} \text{Var}(f)$. Αφού οι $\partial_i f_x$ και $f_t^{(i)}(x)$ δεν εξαρτώνται από την i -οστή συντεταγμένη του x , μπορούμε να γράψουμε

$$\mathbb{E}[V_\alpha^{(i)}] = \mathbb{E}\left[\sum_{t \in J_i \cap [0,1]} (2t \partial_i f_t)^2 \mathbb{1}_{\{f_t^{(i)} < \alpha\}}\right] = \mathbb{E}\left[\sum_{t \in J_i \cap [0,1]} (2t \partial_i f_t)^2 \mathbb{1}_{\{f_t^{(i)} < \alpha\}}\right].$$

Εφαρμόζοντας το Λήμμα 4.1.3 με $g_t = (\partial_i f_t)^2 \mathbb{1}_{\{f_t^{(i)} < \alpha\}}$, παίρνουμε

$$\begin{aligned} \mathbb{E}[V_\alpha^{(i)}] &= \mathbb{E}\left[\sum_{t \in J_i \cap [0,1]} (2t \partial_i f_t)^2 \mathbb{1}_{\{f_t^{(i)} < \alpha\}}\right] \\ &\leq 2\mathbb{E}\left[\int_0^1 t (\partial_i f_t)^2 \mathbb{1}_{\{f_t^{(i)} < \alpha\}} dt\right] \\ &= 2\mathbb{E}\left[\int_0^1 t (\partial_i f_t)^2 \mathbb{1}_{\{f_t^{(i)} < \alpha\}} dt\right] \\ &= \mathbb{E}[Q_\alpha^{(i)}]. \end{aligned}$$

Λόγω της Πρότασης 5.1.2, αυτό σημαίνει ότι

$$\mathbb{E}[V_\alpha] \leq 4\gamma^2 \rho(\alpha) T(f).$$

Από την άλλη πλευρά, από την Πρόταση 5.1.3 (την (5.1.14)) και την ανισότητα Markov,

$$\mathbb{E}[V_\alpha] \geq \mathbb{P}\left[V_\alpha \geq \frac{1}{16}\right] \cdot \frac{1}{16} > \frac{1}{2^{18}} \text{Var}(f),$$

το οποίο έρχεται σε αντίφαση με την υπόθεση ότι $4\gamma^2\rho(\alpha)T(f) \leq 2^{-18}\text{Var}(f)$. $\square$

Ο βασικός ισχυρισμός που χρειάζεται για την απόδειξη του Θεωρήματος 2.5.1 συνδέει το σύνολο κορυφών με την πιθανότητα να παρουσιάζει μεγάλο άλμα η συνάρτηση.

Πρόταση 5.3.2. Για $0 < \alpha \leq 1$, έστω F_α το ενδεχόμενο που ορίζεται στην (5.1.1). Τότε,

$$(5.3.1) \quad \mu(\partial^\pm f) \geq \frac{1}{2}\alpha\mathbb{P}[F_\alpha].$$

Για να αποδείξουμε αυτήν την πρόταση, θα κατασκευάσουμε μια τροποποίηση $\tilde{B}_t$ της B_t , την οποία μπορούμε να σκεφτόμαστε σαν μια «διστακτική» εκδοχή της B_t . Για κάθε συντεταγμένη i , έστω $\tilde{J}_i$ το σύνολο αλμάτων μιας σημειακής διαδικασίας Poisson στο $(0, 1]$ με ένταση $1/(2t)$, ανεξάρτητη από την B_t (και ειδικότερα, ανεξάρτητη από την διαδικασία αλμάτων $J_i = \text{Jump}(B_t^{(i)})$). Ορίζουμε $\tilde{B}_t = (\tilde{B}_t^{(1)}, \dots, \tilde{B}_t^{(n)})$ τη διαδικασία με την ιδιότητα ότι για κάθε i ,

$$\tilde{B}_t^{(i)} = \begin{cases} 0 & t \in J_i \cup \tilde{J}_i \\ B_t^{(i)} & \text{αλλιώς.} \end{cases}$$

Περιγραφικά, υπάρχουν διάφοροι τρόποι για να σκεφτόμαστε την $\tilde{B}_t^{(i)}$:

- (i) Η διαδικασία $\tilde{B}_t^{(i)}$ μπορεί να ιδωθεί σαν μια «διστακτική» παραλλαγή της $B_t^{(i)}$: κάνει άλμα με διπλάσια ένταση (αφού το σύνολο των ασυνεχειών της είναι η ένωση δύο διαδικασιών Poisson με ένταση $1/(2t)$), αλλά τις μισές φορές επιστρέφει στο αρχικό πρόσημο αντί να το αντιστρέφει. Αναφερόμαστε σε αυτή την ερμηνεία ως το «τυπικό ζευγάρισμα» της $\tilde{B}_t^{(i)}$ με την $B_t^{(i)}$: Η διαδικασία $\tilde{B}_t^{(i)}$ είναι ένα αντίγραφο της $B_t^{(i)}$, αλλά με πρόσθετα ανεξάρτητα διστακτικά άλματα.
- (ii) Η διαδικασία $\tilde{B}_t^{(i)}$ είναι ίση με 0 σε ένα διακριτό σύνολο χρονικών στιγμών που ακολουθεί το νόμο μιας σημειακής διαδικασίας Poisson με ένταση $1/t$ (αυτή είναι η ένωση $J_i \cup \tilde{J}_i$). Ανάμεσα σε δύο διαδοχικές ρίζες επιλέγει τυχαία την τιμή t ή $-t$, καθεμία με πιθανότητα $1/2$.

Όμοια με τον συμβολισμό που χρησιμοποιούμε για την B_t , γράφουμε $\tilde{f}_t = f(\tilde{B}_t)$, και ανάλογα $\partial_i \tilde{f}_t$, $\nabla \tilde{f}_t$ και $\tilde{f}_t^{(i)}$.

Λήμμα 5.3.3. Η διαδικασία $\tilde{B}_t$ είναι martingale.

Απόδειξη. Έστω $0 \leq s < t \leq 1$. Για $s = 0$, εφόσον $\tilde{B}_0^{(i)} = 0$ πάντοτε, έχουμε τετριμμένα ότι $\mathbb{E}[\tilde{B}_t^{(i)}] = 0$, άρα υποθέτουμε ότι $s > 0$.

Αν $\tilde{B}_s^{(i)} = 0$, τότε $s \in J_i \cup \tilde{J}_i$. Αφού είναι ανεξάρτητες σημειακές διαδικασίες Poisson, σχεδόν βεβαίως έχουμε $J_i \cap \tilde{J}_i = \emptyset$, και $\mathbb{P}\left[s \in J_i \mid \tilde{B}_s^{(i)}\right] = \mathbb{P}\left[s \in \tilde{J}_i \mid \tilde{B}_s^{(i)}\right] = 1/2$. Έτσι, αφού $\tilde{B}_t^{(i)} = B_t^{(i)}$ σχεδόν βεβαίως, έχουμε ότι κάτω από το ενδεχόμενο $\tilde{B}_s^{(i)} = 0$,

$$\mathbb{E}\left[\tilde{B}_t^{(i)} \mid \tilde{B}_s^{(i)}\right] = \frac{1}{2}\mathbb{E}\left[B_t^{(i)} \mid B_s^{(i)}, s \in J_i\right] + \frac{1}{2}\mathbb{E}\left[B_t^{(i)} \mid B_s^{(i)}, s \in \tilde{J}_i\right].$$

Από τον ορισμό της διαδικασίας B_t είναι φανερό ότι, όταν $\tilde{B}_s^{(i)} = 0$,

$$\mathbb{E}\left[B_t^{(i)} \mid B_s^{(i)}, s \in J_i\right] = -\mathbb{E}\left[B_t^{(i)} \mid B_s^{(i)}, s \in \tilde{J}_i\right],$$

συνεπώς

$$\mathbb{E}\left[\tilde{B}_t^{(i)} \mid \tilde{B}_s^{(i)}\right] = 0 = \tilde{B}_s^{(i)}.$$

Τέλος, κάτω από το ενδεχόμενο $\tilde{B}_s^{(i)} \neq 0$, αφού $\tilde{B}_t^{(i)} \neq 0$ σχεδόν βεβαίως, από την (4.1.1) έχουμε ότι

$$\mathbb{P}\left[\text{sign}\tilde{B}_t^{(i)} \neq \text{sign}\tilde{B}_s^{(i)} \mid \tilde{B}_s^{(i)}\right] = \frac{t-s}{2t}.$$

Έτσι,

$$\begin{aligned}\mathbb{E}\left[\tilde{B}_t^{(i)} \mid \tilde{B}_s^{(i)}\right] &= \text{sign}\tilde{B}_s^{(i)} \cdot t \cdot \frac{t-s}{2t} + \text{sign}\tilde{B}_s^{(i)} \cdot (-t) \cdot \frac{t-s}{2t} \\ &= \text{sign}\tilde{B}_s^{(i)} \cdot s \\ &= \tilde{B}_s^{(i)}.\end{aligned}$$

□

Απόδειξη της Πρότασης 5.3.2. Για να διακρίνουμε ανάμεσα στα σύνορα κορυφών, θα χρησιμοποιήσουμε τη διαδικασία διστακτικών αλμάτων $\tilde{f}_t$ που ορίσαμε παραπάνω. Αποδεικνύουμε την (5.1.1) για το εσωτερικό σύνορο κορυφών ∂^+ . Η απόδειξη για το ∂^- έπεται αν θεωρήσουμε την $-f$. Θέτουμε $\tau = \inf\{t > 0 \mid \exists i \in [n] \text{ τ.ω. } \tilde{B}_t^{(i)} = 0 \text{ και } f_t^{(i)} \geq \alpha\} \wedge 1$. Σημειώνουμε ότι για κάθε $t_0 > 0$, έχουμε σχεδόν βεβαίως ότι $\tilde{B}_t^{(i)} = 0$ για πεπερασμένες μόνο τιμές του $t \in [t_0, 1]$. Έτσι, αν $0 < \tau < 1$, τότε το infimum στον ορισμό του τ πιάνεται ως minimum, και υπάρχει i_0 τέτοιος ώστε $\tilde{B}_\tau^{(i_0)} = 0$ και $f_\tau^{(i_0)} \geq \alpha$. Μάλιστα, αυτό ισχύει και αν $\tau = 0$: Σε αυτή την περίπτωση, υπάρχει μια ακολουθία χρονικών στιγμών $t_k \rightarrow 0$ και δεικτών i_k ώστε $f_{t_k}^{(i_k)} \geq \alpha$ και $\tilde{B}_{t_k}^{(i_k)} = 0$. Αφού υπάρχουν μόνο πεπερασμένοι το πλήθος διακεκριμένοι δείκτες i_k , υπάρχει μια υποακολουθία k_ℓ τέτοια ώστε όλοι οι i_{k_ℓ} να είναι ο ίδιος δείκτης i_0 , και ο ισχυρισμός προκύπτει από τη συνέχεια της $f^{(i_0)}$ και το γεγονός ότι $\tilde{B}_0 = 0$.

Όταν συμβαίνει το F_α , αναγκαστικά έχουμε $\tau < 1$, αφού $\tilde{B}_t^{(i)} = 0$ όταν η $B_t^{(i)}$ είναι ασυνεχής.

Αφού η $\tilde{B}_1$ είναι ομοιόμορφη στον υπερκύβο,

$$\begin{aligned}\mu(\partial^+ f) &= \mathbb{P}[\tilde{B}_1 \in \partial^+ f] \geq \mathbb{P}[\tilde{B}_1 \in \partial^+ f \mid \tau < 1] \mathbb{P}[\tau < 1] \\ &\geq \mathbb{P}[\tilde{B}_1 \in \partial^+ f \mid \tau < 1] \mathbb{P}[F_\alpha],\end{aligned}$$

και έτσι αρκεί να δείξουμε ότι

$$(5.3.2) \quad \mathbb{P}[\tilde{B}_1 \in \partial^+ f \mid \tau < 1] \geq \frac{1}{2}\alpha.$$

Αν υποθέσουμε ότι $\tau < 1$, συμβολίζουμε με i_0 μία συντεταγμένη για την οποία $\tilde{B}_\tau^{(i_0)} = 0$ και $f_\tau^{(i_0)} \geq \alpha$. Αν $\tau = 0$ τότε $\tilde{B}_\tau = B_\tau$. Αλλιώς, σχεδόν βεβαίως η i_0 είναι η μοναδική συντεταγμένη της $\tilde{B}_\tau$ που είναι 0, άρα $\tilde{B}_\tau^{(j)} = B_\tau^{(j)}$ για κάθε $j \neq i_0$ σχεδόν βεβαίως. Αφού η συνάρτηση $f^{(i)}(x)$ δεν εξαρτάται από την i -οστή συντεταγμένη, συμπεραίνουμε ότι $\tilde{f}_\tau^{(i_0)} = f_\tau^{(i_0)} \geq \alpha$ σχεδόν βεβαίως. Έτσι, εάν $\tau < 1$, από την ιδιότητα martingale της $\tilde{f}_t^{(i_0)}$, έχουμε

$$(5.3.3) \quad \mathbb{P}[\tilde{f}_1^{(i_0)} = 1 \mid \tilde{B}_\tau] \geq \alpha.$$

Όμοια, χρησιμοποιώντας την ιδιότητα martingale της $\tilde{B}_t^{(i_0)}$, έχουμε ότι $\mathbb{E}[\tilde{B}_1^{(i_0)} \mid \tilde{B}_\tau] = \tilde{B}_\tau^{(i_0)} = 0$, και έτσι

$$\mathbb{P}[\tilde{B}_1^{(i_0)} = 1 \mid \tilde{B}_\tau] = \mathbb{P}[\tilde{B}_1^{(i_0)} = -1 \mid \tilde{B}_\tau] = \frac{1}{2}.$$

Αφού η $\partial_{i_0}\tilde{f}_t$ είναι ανεξάρτητη από την $\tilde{B}_t^{(i_0)}$, τελικά παίρνουμε

$$\begin{aligned}\mathbb{P}[\tilde{B}_1 \in \partial^+ f \mid \tilde{B}_\tau] &= \mathbb{P}[\tilde{B}_1^{(i_0)} = 1 \wedge \partial_{i_0}\tilde{f}_1 = 1 \mid \tilde{B}_\tau] + \mathbb{P}[\tilde{B}_1^{(i_0)} = -1 \wedge \partial_{i_0}\tilde{f}_1 = -1 \mid \tilde{B}_\tau] \\ (\text{ανεξαρτησία}) &= \frac{1}{2}\mathbb{P}[\partial_{i_0}\tilde{f}_1 = 1 \mid \tilde{B}_\tau] + \frac{1}{2}\mathbb{P}[\partial_{i_0}\tilde{f}_1 = -1 \mid \tilde{B}_\tau] \\ &= \frac{1}{2}\mathbb{P}[\tilde{f}_1^{(i_0)} = 1 \mid \tilde{B}_\tau] \\ &\geq \frac{1}{2}\alpha\end{aligned}$$

από την (5.3.3). □

Απόδειξη του Θεωρήματος 2.5.1. Έστω γ η σταθερά από τη διατύπωση του Λήμματος 5.1.1. Έστω α που ικανοποιεί την $\alpha \log \frac{1}{\alpha} = \frac{1}{2^{30}\gamma^2} r_{\text{Tal}}$. Τότε, η συνθήκη $4\gamma^2 \rho(\alpha) T(f) \leq 2^{-18} \text{Var}(f)$ ικανοποιείται στο Λήμμα 5.3.1, και έχουμε ότι $\mathbb{P}[F_\alpha] \geq 2^{-14} \text{Var}(f)$. Συνδυάζοντας με την Πρόταση 5.3.2, παίρνουμε

$$(5.3.4) \quad \mu(\partial^\pm f) \geq \frac{1}{2}\alpha \mathbb{P}[F_\alpha] \geq \frac{1}{2^{14}}\alpha \text{Var}(f).$$

Μένει να δώσουμε ένα κάτω φράγμα για τον α . Γι' αυτόν τον σκοπό, παρατηρούμε ότι $\alpha \log \frac{1}{\alpha} \leq \sqrt{\alpha}$ για κάθε $\alpha \in [0, 1]$, άρα $\alpha \geq \frac{1}{2^{60}\gamma^4} r_{\text{Tal}}^2$. Έπεται ότι $\log \frac{1}{\alpha} \leq \log \left(\frac{2^{60}\gamma^4}{r_{\text{Tal}}^2} \right)$, άρα υπάρχει

μια σταθερά C'_B , την οποία μπορούμε να πάρουμε μεγαλύτερη από την C_T , τέτοια ώστε

$$\alpha = \frac{r_{\text{Tal}}}{2^{30}\gamma^2 \log \frac{1}{\alpha}} \geq \frac{r_{\text{Tal}}}{C'_B \log \frac{C'_B}{r_{\text{Tal}}}}.$$

Εισάγοντας αυτή την εκτίμηση στην (5.3.4) παίρνουμε το ζητούμενο. $\square$

5.4 Ευαισθησία θορύβου και επιρροές

Σε αυτή την ενότητα αποδεικνύουμε το Θεώρημα 2.5.4. Όπως εξηγήσαμε στην (4.2.1), στόχος μας είναι να δείξουμε ότι

$$(5.4.1) \quad S_\varepsilon(f) = \mathbb{E} \left[\sum_{i=1}^n \int_0^{\sqrt{1-\varepsilon}} t (\partial_i f_t)^2 dt \right] \leq C \text{Var}(f) \left(\sum_{i=1}^n I_i(f)^2 \right)^{c\varepsilon}.$$

Αποδεικνύουμε αρχικά ότι μπορούμε να υποθέσουμε ότι η f είναι μονότονη. Για κάθε $i = 1, \dots, n$, ορίζουμε έναν τελεστή κ_i θέτοντας

$$(\kappa_i f)(y) = \begin{cases} \max\{f(y), f(\sigma_i(y))\} & , \text{ αν } y_i = 1, \\ \min\{f(y), f(\sigma_i(y))\} & , \text{ αν } y_i = 0. \end{cases}$$

Το επόμενο λήμμα συνδέει τις επιρροές και τις ευαισθησίες των $\kappa_i f$ και f .

Λήμμα 5.4.1 ([5, Λήμμα 2.7]). *Η συνάρτηση $\kappa_1 \kappa_2 \dots \kappa_n f$ είναι μονότονη, και για κάθε ζεύγος δεικτών i, j ισχύει ότι $I_i(\kappa_j f) \leq I_i(f)$ και $S_\varepsilon(\kappa_i f) \geq S_\varepsilon(f)$.*

Έτσι, αν η εξίσωση (5.4.1) ισχύει για την $\tilde{f} = \kappa_1 \dots \kappa_n f$, τότε ισχύει και για την f , διότι $S_\varepsilon(f) \leq S_\varepsilon(\tilde{f})$ και $\sum_{i=1}^n I_i(f)^2 \geq \sum_{i=1}^n I_i(\tilde{f})^2$. Αρκεί λοιπόν να επαληθεύσουμε την (5.4.1) για μονότονες συναρτήσεις.

Για να αποδείξουμε την (5.4.1), μπορούμε επίσης για οποιαδήποτε δοσμένη απόλυτη σταθερά $K < 1$ να υποθέσουμε ότι

$$(5.4.2) \quad \sum_{i=1}^n I_i(f)^2 \leq K.$$

Πράγματι, αν $\sum_{i=1}^n I_i(f)^2 \geq K$ για κάποιον K , τότε αφού η f είναι μονότονη, έχουμε

$$\text{Var}(f) = \sum_{S \subseteq [n], S \neq \emptyset} \widehat{f}(S)^2 \geq \sum_{i=1}^n \widehat{f}(\{i\})^2 = \sum_{i=1}^n I_i(f)^2 \geq K$$

από την (2.1.3), άρα $\text{Var}(f) \left(\sum_i I_i(f)^2 \right) \geq K^2$. Τότε, η (5.4.1) ισχύει τετραμμένα με $C = 1/K^2$ και $c = 1$, διότι $S_\varepsilon \leq 1$ για κάθε ε .

Όμοια, μπορούμε για οποιαδήποτε δοσμένη απόλυτη σταθερά $K < 1$ να υποθέσουμε ότι

$$(5.4.3) \quad \text{Var}(f) \leq K.$$

Αλλιώς, το Θεώρημα 2.5.4 θα ήταν ισοδύναμο (έως σταθερές) με το αρχικό θεώρημα που αποδεικνύεται στο [36].

Παρατήρηση 5.4.2. Από την απόδειξη που θα δώσουμε, μπορεί επίσης κάποιος να ανακτήσει το αρχικό θεώρημα που αποδείχθηκε στο [36], κάνουμε όμως αυτή την υπόθεση διότι απλουστεύει κάποιες εκτιμήσεις.

Ορίζουμε

$$(5.4.4) \quad R(t) = \mathbb{E} \left[\sum_i (\partial_i f_t)^2 \right] = \mathbb{E} \left[\|\nabla f_t\|_2^2 \right].$$

Την χρονική στιγμή 0, έχουμε

$$R(0) = \sum_{i=1}^n (\partial_i f_0)^2 = \sum_{i=1}^n \left(\widehat{\partial_i f}(\emptyset) \right)^2 = \sum_{i=1}^n \widehat{f}(\{x_i\})^2 = \sum_{i=1}^n I_i(f)^2.$$

Η συνάρτηση $R(t)$ είναι μονότονη ως προς t : Αφού η $\partial_i f_t$ είναι martingale, η $(\partial_i f_t)^2$ είναι submartingale και έτσι η $\mathbb{E}[(\partial_i f_t)^2]$ είναι αύξουσα.

Εφαρμόζοντας το Πρόγραμμα 4.1.5 για την $\partial_i f$, για κάθε δείκτη i έχουμε

$$\frac{d}{dt} \mathbb{E}[(\partial_i f_t)^2] = 2t \mathbb{E} \left[\sum_{j=1}^n (\partial_j \partial_i f_t)^2 \right].$$

Έπεται ότι

$$(5.4.5) \quad \frac{d}{dt} R(t) = 2t \mathbb{E} \left[\sum_{i=1}^n \sum_{j=1}^n (\partial_i \partial_j f_t)^2 \right] \leq 2 \mathbb{E} \left[\|\nabla^2 f_t\|_{\text{HS}}^2 \right],$$

όπου $\|X\|_{\text{HS}}$ είναι η νόρμα Hilbert-Schmidt ενός πίνακα. Από το Λήμμα 2.6.3, υπάρχει μια θετική συνεχής συνάρτηση $C(t)$ τέτοια ώστε

$$\begin{aligned} \frac{d}{dt} R(t) &\leq 2C(t) \mathbb{E} \left[\|\nabla f_t\|_2^2 \log \left(\frac{C(t)}{\|\nabla f_t\|_2^2} \right) \right] \\ (\text{ανισότητα Jensen}) &\leq 2C(t) \mathbb{E} \left[\|\nabla f_t\|_2^2 \right] \log \left(\frac{C(t)}{\mathbb{E}[\|\nabla f_t\|_2^2]} \right) = 2C(t) R(t) \log \left(\frac{C(t)}{R(t)} \right). \end{aligned}$$

Η $C(t)$ είναι συνεχής, άρα φραγμένη στο $[0, 1/2]$, επομένως υπάρχει μια σταθερά $c > 0$ τέτοια

ώστε για κάθε $t \in [0, 1/2]$,

$$(5.4.6) \quad \frac{d}{dt} R(t) \leq c R(t) \log \left(\frac{c}{R(t)} \right).$$

Από την (5.4.2), μπορούμε να υποθέσουμε ότι $R(0) \leq c/2$. Χρησιμοποιώντας την (5.4.6) μαζί με το Λήμμα 4.2.2, βλέπουμε ότι υπάρχουν σταθερές $C, L > 0$ και μια χρονική στιγμή t_0 , που όλα τους εξαρτώνται μόνο από την c , τέτοια ώστε για κάθε $t \in [0, t_0]$,

$$R(t) \leq L \cdot R(0) e^{-Ct}.$$

Ειδικότερα, υπάρχει μια σταθερά $K > 0$ τέτοια ώστε

$$(5.4.7) \quad R(e^{-K}) \leq L \cdot R(0)^{5/6},$$

και αφού η R είναι αύξουσα, μπορούμε να υποθέσουμε ότι $K > 1$. Ορίζουμε

$$(5.4.8) \quad G(s) = R(e^{-s}).$$

Από το Λήμμα 4.2.1, κάθε προσθετός της μορφής $\mathbb{E}[(\partial_i f_{e^{-s}})^2]$ είναι λογαριθμικά κυρτή συνάρτηση, άρα και η $G(s)$ είναι λογαριθμικά κυρτή συνάρτηση του s ως θετικός γραμμικός συνδυασμός λογαριθμικά κυρτών συναρτήσεων.

Λήμμα 5.4.3. Έστω $K \geq 1$ και έστω $G(s)$ μια λογαριθμικά κυρτή φθίνουσα συνάρτηση. Θέτουμε $v = \int_0^K e^{-2s} G(s) ds$ και υποθέτουμε ότι $v > G(K)$. Τότε, για κάθε $r < K$,

$$\int_0^r e^{-2s} G(s) ds \geq v \cdot \left(1 - \left(\frac{G(K)}{v} \right)^{r/K} \right).$$

Απόδειξη. Θεωρούμε τη συνάρτηση

$$h_\ell(s) = \frac{v\ell}{1 - e^{-\ell K}} e^{-(\ell-2)s},$$

όπου ℓ είναι η μεγαλύτερη λύση της εξίσωσης $h_\ell(K) = G(K)$. Από την επιλογή της h_ℓ , έχουμε

$$\int_0^K e^{-2s} h_\ell(s) ds = v = \int_0^K e^{-2s} G(s) ds.$$

Εφόσον η $e^{-2s} h_\ell(s)$ είναι λογαριθμικά γραμμική στο $[0, K]$, η συνάρτηση $e^{-2s} G(s)$ είναι λογαριθμικά κυρτή στο $[0, K]$, έχουν το ίδιο ολοκλήρωμα στο $[0, K]$ και $G(K) = h_\ell(K)$, πρέπει να ισχύει κάποιο από τα παρακάτω:

(i) $h_\ell(s) = G(s)$

(ii) Οι συναρτήσεις τέμνονται το πολύ μία φορά στο διάστημα $[0, K]$ σε κάποιο σημείο s_0 τέτοιο ώστε $G(s) \geq h_\ell(s)$ για κάθε $s < s_0$.

Σε κάθε μία από τις δύο περιπτώσεις, για κάθε $r \in [0, K]$, έχουμε ότι

$$(5.4.9) \quad \int_0^r e^{-2s} G(s) ds \geq \int_0^r e^{-2s} h_\ell(s) ds = \frac{v}{1 - e^{-K\ell}} (1 - e^{-r\ell}) \geq v(1 - e^{-r\ell}).$$

Από την άλλη πλευρά, επιλέξαμε το ℓ έτσι ώστε να ισχύει $h_\ell(K) = G(K)$, άρα

$$\frac{\ell}{1 - e^{-\ell K}} e^{-(\ell-2)K} = \frac{h_\ell(K)}{v} = \frac{G(K)}{v} < 1,$$

όπου η τελευταία ανισότητα ισχύει από την υπόθεση για το v . Η συνάρτηση $\frac{x}{1 - e^{-xK}} e^{-(x-2)K}$ είναι φθίνουσα συνάρτηση του x στο διάστημα $[2, \infty)$, αλλά είναι μεγαλύτερη από 1 στο $x = 2$. Συνεπώς, αφού ο ℓ είναι ο μεγαλύτερος αριθμός για τον οποίο $h_\ell(K) = G(K)$, πρέπει να ισχύει ότι $\ell > 2$. Έτσι, έχουμε

$$\ell e^{-(\ell-2)K} = \frac{G(K)(1 - e^{-\ell K})}{v} \leq \frac{G(K)}{v},$$

και αναδιατάσσοντας αυτή την ανισότητα, αφού $\ell > 2$,

$$e^{-\ell K} \leq \frac{G(K)}{v} \frac{e^{-2K}}{\ell} \leq \frac{G(K)}{v}.$$

Έπεται ότι

$$\ell \geq \frac{1}{K} \log \left(\frac{v}{G(K)} \right).$$

Εισάγοντας αυτή την ανισότητα στο δεξί μέλος της (5.4.9) παίρνουμε

$$\begin{aligned} \int_0^r e^{-2s} G(s) ds &\geq v(1 - e^{-r\ell}) \\ &\geq v \left(1 - e^{-\frac{r}{K} \log \left(\frac{v}{G(K)} \right)} \right) = v \left(1 - \left(\frac{G(K)}{v} \right)^{\frac{r}{K}} \right). \end{aligned}$$

□

Περνάμε τώρα στην απόδειξη του θεωρήματος.

Απόδειξη του Θεωρήματος 2.5.4. Από το Πόρισμα 4.1.4 και την (5.4.4),

$$\text{Var}(f) = \text{Var}(f_1) = 2\mathbb{E} \left[\sum_{i=1}^n \int_0^1 t (\partial_i f_i)^2 dt \right] = 2 \int_0^1 t R(t) dt,$$

και με μια αλλαγή μεταβλητής (θυμηθείτε την (5.4.8)) η ισότητα αυτή γίνεται

$$\text{Var}(f) = 2 \int_0^\infty e^{-2s} G(s) ds.$$

Ορίζουμε $v = \int_0^K e^{-2s} G(s) ds$, όπου K είναι η σταθερά από την εξίσωση (5.4.7). Παρατηρούμε

ότι αφού $K \geq 1$ και η G είναι φθίνουσα,

$$(5.4.10) \quad \text{Var}(f) - \nu = \int_K^\infty e^{-2s} G(s) ds \leq G(K) \leq L \cdot R(0)^{5/6}.$$

Αναδιατάσσοντας αυτή την ανισότητα, παίρνουμε

$$(5.4.11) \quad \frac{\nu}{G(K)} \geq \frac{\text{Var}(f) - LR(0)^{5/6}}{LR(0)^{5/6}}.$$

Θέτουμε $g(x) = \frac{1+f(x)}{2}$, και χωρίς περιορισμό της γενικότητας υποθέτουμε ότι $\mathbb{E}[f] = f(0) \leq 0$, οπότε $\mathbb{E}[g] = g(0) \leq \frac{1}{2}$ (αν $f(0) > 0$, μπορούμε να θεωρήσουμε την $g(x) = (1 - f(x))/2$). Αυτό συνεπάγεται ότι

$$\text{Var}(g) = g(0)(1 - g(0)) \geq \frac{1}{2}g(0).$$

Εφαρμόζοντας το Λήμμα 2.6.2 για την g και με $t = 0$, βλέπουμε ότι υπάρχει μια σταθερά C τέτοια ώστε

$$(5.4.12) \quad \begin{aligned} R(0) &= \mathbb{E} \left[\sum_{i=1}^n (\partial_i f(0))^2 \right] = \|\nabla f(0)\|_2^2 = 4\|\nabla g(0)\|_2^2 \\ &\leq Cg(0)^2 \log \left(\frac{e}{g(0)} \right) \\ &\leq C'(\text{Var}(g))^2 \log \left(\frac{C'}{\text{Var}(g)} \right) \leq C''\text{Var}(f)^2 \log \left(\frac{4C''}{\text{Var}(f)} \right). \end{aligned}$$

Από την (5.4.3), μπορούμε να υποθέσουμε ότι η $\text{Var}(f)$ είναι αρκετά μικρή, οπότε η (5.4.12) συνεπάγεται ότι

$$(5.4.13) \quad R(0)^{2/3} \leq \text{Var}(f).$$

Εισάγοντας αυτή τη σχέση στην (5.4.11), παίρνουμε

$$\frac{\nu}{G(K)} \geq \frac{R(0)^{2/3} - LR(0)^{5/6}}{LR(0)^{5/6}}.$$

Λόγω της (5.4.2), μπορούμε να υποθέσουμε ότι το $R(0)$ είναι μικρότερο από οποιαδήποτε δοσμένη σταθερά. Αφού οι σταθερές L και K δεν εξαρτώνται από την R , μπορούμε να επιλέξουμε το $R(0)$ αρκετά μικρό ώστε να ισχύει η $LR(0)^{5/6} \leq \frac{1}{2}R(0)^{2/3}$, άρα

$$(5.4.14) \quad \frac{\nu}{G(K)} \geq \frac{1}{2L}R(0)^{-1/6}.$$

Για αρκετά μικρό $R(0)$, η (5.4.14) συνεπάγεται ότι $\nu > G(K)$, άρα μπορούμε να χρησιμοποι-

ήσουμε το Λήμμα 5.4.3. Από τις εξισώσεις (5.4.10) και (5.4.11), έχουμε

$$(5.4.15) \quad \int_0^r e^{-2s} G(s) ds \geq v \left(1 - \left(\frac{G(K)}{v} \right)^{r/K} \right) \\ \geq \left(\text{Var}(f) - LR(0)^{5/6} \right) \left(1 - \left(2LR(0)^{1/6} \right)^{r/K} \right).$$

Αυτό μας επιτρέπει να αποδείξουμε την (5.4.1): χρησιμοποιώντας τις (5.4.14), (5.4.13) και το γεγονός ότι $\varepsilon < 1$ και $K > 1$, βλέπουμε ότι

$$\begin{aligned} S_\varepsilon(f) &= \mathbb{E} \left[\sum_{i=1}^n \int_0^{\sqrt{1-\varepsilon}} t (\partial_i f_t)^2 dt \right] \\ &= \int_0^{\sqrt{1-\varepsilon}} t R(t) dt \\ &\leq \int_0^{e^{-\varepsilon/2}} t R(t) dt \\ &= \text{Var}(f) - \int_0^{-\varepsilon/2} e^{-2s} G(s) ds \\ &\leq LR(0)^{5/6} + \text{Var}(f) \left(2LR(0)^{1/6} \right)^{\varepsilon/(2K)} \\ &\leq L \text{Var}(f) R(0)^{1/6} + \text{Var}(f) \left(2LR(0)^{1/6} \right)^{\varepsilon/(2K)} \\ &\leq C \text{Var}(f) R(0)^{\varepsilon/(12K)} \end{aligned}$$

για κάποια απόλυτη σταθερά C .

□

ΚΕΦΑΛΑΙΟ 6

Η μέθοδος των τυχαίων περιορισμών

6.1 Εισαγωγή

Σε αυτό το κεφάλαιο συζητάμε την πρόσφατη εργασία των Eldan, Kindler, Lifshitz και Minzer [21] στην οποία αποδεικνύονται ισοπεριμετρικές ανισότητες στον διακριτό κύβο $\{0,1\}^n$. Όλες αυτές οι ανισότητες αφορούν διάφορες έννοιες συνόρων και τις σχέσεις μεταξύ τους για Boolean συναρτήσεις. Για μια συνάρτηση $f : \{0,1\}^n \rightarrow \{0,1\}$ και ένα σημείο $x \in \{0,1\}^n$, ορίζουμε την ευαισθησία της f στο x , και τη συμβολίζουμε με $h_f(x)$, να είναι το πλήθος των συντεταγμένων $i \in [n]$ για τις οποίες $f(x \oplus e_i) \neq f(x)$. Θα δώσουμε απλούστερες αποδείξεις για τα παρακάτω αποτελέσματα, τα οποία οφείλονται στους Talagrand [48] και Eldan-Gross [19], και τα οποία παρουσιάσαμε στα προηγούμενα κεφάλαια..

Θεώρημα 6.1.1. Υπάρχει απόλυτη σταθερά $c > 0$ τέτοια ώστε για κάθε συνάρτηση $f : \{0,1\}^n \rightarrow \{0,1\}$ να ισχύει ότι

$$\mathbb{E} \left[\sqrt{h_f} \right] \geq c \operatorname{Var}(f) \sqrt{\log \left(\frac{1}{\operatorname{Var}(f)} \right)}.$$

Θεώρημα 6.1.2. Υπάρχει απόλυτη σταθερά $c > 0$ τέτοια ώστε για κάθε συνάρτηση $f : \{0,1\}^n \rightarrow \{0,1\}$ να ισχύει ότι

$$\mathbb{E} \left[\sqrt{h_f} \right] \geq c \operatorname{Var}(f) \sqrt{\log \left(1 + 1 / \left(\sum_{i=1}^n I_i[f]^2 \right) \right)}.$$

Η προσέγγιση που θα παρουσιάσουμε βασίζεται σε μια απλή παρατήρηση, η οποία κατόπιν συνδυάζεται με γνωστά αποτελέσματα συγκέντρωσης από την ανάλυση Fourier.

Για δοθείσα $f : \{0,1\}^n \rightarrow \{0,1\}$, θεωρούμε τους συντελεστές Fourier

$$\widehat{f}(S) = \mathbb{E} \left[f(x) \prod_{i \in S} (-1)^{x_i} \right]$$

καθώς και την κλίση $\nabla f : \{0, 1\}^n \rightarrow \{-1, 0, 1\}^n$ που ορίζεται από την

$$\nabla f(x) = (f(x \oplus e_i) - f(x))_{i=1, \dots, n}.$$

Αρχικά παρατηρούμε ότι η τετραγωνική ρίζα της ευαισθησίας της f στο x είναι ίση με $\|\nabla f(x)\|_2$, και επίσης κοιτάζοντας την απόλυτη τιμή της i -οστής συντεταγμένης της ∇f , δηλαδή τη συνάρτηση $x \mapsto |f(x \oplus e_i) - f(x)|$, βλέπουμε ότι η μέση τιμή της είναι τουλάχιστον ίση με την απόλυτη τιμή του συντελεστή Fourier $|\widehat{f}(\{i\})|$. Αυτό υποδεικνύει ότι αν η f έχει μεγάλους συντελεστές Fourier σε μονοσύνολα τότε το μέτρο της κλίσης πρέπει να είναι μεγάλο, άρα η μέση τιμή της τετραγωνικής ρίζας της ευαισθησίας πρέπει επίσης να είναι μεγάλη. Πράγματι, με απλή εφαρμογή της τριγωνικής ανισότητας παίρνουμε

$$\mathbb{E} [\|\nabla f(x)\|_2] \geq \|\mathbb{E} [|\nabla f(x)|]\|_2 \geq \|(|\widehat{f}(\{1\})|, \dots, |\widehat{f}(\{n\})|)\|_2,$$

που ισούται με την $\sqrt{\sum_{i=1}^n \widehat{f}(\{i\})^2}$. Καθώς το άθροισμα των τετραγώνων όλων των συντελεστών Fourier είναι το πολύ 1, βλέπουμε έτσι ότι

$$\mathbb{E} [\|\nabla f(x)\|_2] \geq \sum_{i=1}^n \widehat{f}(\{i\})^2.$$

Με λόγια, με μια απλή εφαρμογή της τριγωνικής ανισότητας έχουμε καταφέρει να δώσουμε ένα κάτω φράγμα για την ποσότητα που μας ενδιαφέρει μέσω του βάρους Fourier επιπέδου 1 της συνάρτησης f . Αυτή είναι ήδη μια ενδιαφέρουσα ισοπεριμετρική ανισότητα, αν και αρκετά ασθενής.

Στη συνέχεια δείχνουμε πώς να εκμεταλλευτούμε αυτό το βασικό αποτέλεσμα για να φτάσουμε στα Θεωρήματα 6.1.1 και 6.1.2 μέσω τυχαίων περιορισμών. Για να αποκτήσουμε κάποια διαίσθηση, ας υποθέσουμε ότι γνωρίζαμε πως η συνάρτηση f έχει σημαντικό βάρος στο επίπεδο d (όπου $d > 1$). Το ερώτημα είναι αν μπορούμε να χρησιμοποιήσουμε την παραπάνω λογική για να δώσουμε ενδιαφέροντα κάτω φράγματα για την

$$\mathbb{E} [\|\nabla f(x)\|_2].$$

Η απάντηση είναι ναι, και ένας τρόπος για να το επιτύχουμε είναι μέσω τυχαίων περιορισμών. Αν η f έχει σημαντικό βάρος στο επίπεδο d , επιλέγουμε ένα σύνολο συντεταγμένων $I \subseteq [n]$ στο οποίο συμπεριλαμβάνουμε κάθε συντεταγμένη με πιθανότητα $p = 1/d$ και στη συνέχεια περιορίζουμε τις συντεταγμένες που ανήκουν στο $\bar{I} = [n] \setminus I$ σε μια ομοιόμορφα επιλεγμένη τιμή $z \in \{0, 1\}^{\bar{I}}$, τότε περιμένουμε το βάρος της f να πέσει από το επίπεδο d στο επίπεδο $pd = 1$ περίπου. Σε αυτό το σημείο θα μπορούσαμε να χρησιμοποιήσουμε την παραπάνω λογική για την περιορισμένη συνάρτηση και να πάρουμε ένα κάτω φράγμα για τη μέση τιμή της ρίζας της ευαισθησίας της. Το ερώτημα που τίθεται είναι πώς σχετίζεται αυτή η τελευταία ποσότητα με την $\mathbb{E} [\|\nabla f(x)\|_2]$.

Αν το σκεφτούμε λίγο, για ένα σταθερό σημείο x , αν επιλέξουμε το I τυχαία όπως παραπάνω και σταθεροποιήσουμε όλες τις συντεταγμένες που δεν ανήκουν στο I , τότε θα επιβιώσει μόνο ένα ποσοστό (περίπου) $1/d$ των συντεταγμένων που είναι ευαίσθητες στο x , άρα περιμένουμε ότι η ευαισθησία θα πέσει κατά έναν παράγοντα $1/d$ ως αποτέλεσμα αυτού του περιορισμού. Πράγματι, μπορούμε να αποδείξουμε ότι κάτω από τυχαίους περιορισμούς, η ποσότητα $\mathbb{E}[\|\nabla f(x)\|_2]$ πέφτει κατά έναν παράγοντα $1/\sqrt{d}$. Έτσι, συνδυάζοντας αυτήν την παρατήρηση με τον προηγούμενο συλλογισμό περιμένουμε ένα κάτω φράγμα της μορφής

$$\mathbb{E}[\|\nabla f(x)\|_2] \geq \Omega\left(\sqrt{d} \sum_{d \leq |S| \leq 2d} \widehat{f}(S)^2\right),$$

και πράγματι αυτό είναι το φράγμα που αποδεικνύουμε.

Κατόπιν, τα Θεωρήματα 6.1.1 και 6.1.2 προκύπτουν από γνωστές εκτιμήσεις ουρών της ανάλυσης Fourier. Συγκεκριμένα, είναι γνωστό ότι (α) αν η f έχει μικρή διασπορά τότε το μεγαλύτερο μέρος της μάζας Fourier της βρίσκεται πάνω από το επίπεδο $d = \Omega(\log(1/\text{Var}(f)))$, και (β) αν η f έχει μικρές επιρροές τότε το μεγαλύτερο μέρος της μάζας Fourier της βρίσκεται πάνω από το επίπεδο $d = \Omega(\log(1/\sum_{i=1}^n I_i[f]^2))$.

Ανάλυση Fourier στον διακριτό κύβο

Υπενθυμίζουμε εδώ κάποιους βασικούς ορισμούς από την ανάλυση Fourier, τους οποίους προσαρμόζουμε στον διακριτό κύβο $\{0,1\}^n$. Θεωρούμε τον χώρο των πραγματικών συναρτήσεων $f : \{0,1\}^n \rightarrow \mathbb{R}$, εφοδιασμένον με το εσωτερικό γινόμενο $\langle f, g \rangle = \mathbb{E}[f(x)g(x)]$. Είναι γνωστό ότι οι συναρτήσεις $\{\mathbb{1}_S\}_{S \subseteq [n]}$ που ορίζονται από τις $\mathbb{1}_S(x) = \prod_{i \in S} (-1)^{x_i}$ σχηματίζουν ορθοκανονική βάση, μπορούμε λοιπόν να αναπτύξουμε κάθε συνάρτηση f μονοσήμαντα ως $f(x) = \sum_{S \subseteq [n]} \widehat{f}(S) \mathbb{1}_S(x)$ όπου $\widehat{f}(S) = \langle f, \mathbb{1}_S \rangle$. Θεωρούμε επίσης L_p νόρμες για $p \geq 1$, οι οποίες ορίζονται ως $\|f\|_p = (\mathbb{E}|f(x)|^p)^{1/p}$.

Για δοθέν $z \in \{0,1\}^n$, συμβολίζουμε με $z_{-i} \in \{0,1\}^{n-1}$ το διάνυσμα που προκύπτει αν πετάξουμε την i -οστή συντεταγμένη z . Συμβολίζουμε με $(x_i = 1, x_{-i} = z_{-i})$ το διάνυσμα n -ψηφίων που έχει 1 στην i -οστή συντεταγμένη και συμπίπτει με το z σε κάθε άλλη συντεταγμένη.

Ορισμός 6.1.3. Η παράγωγος της f στη διεύθυνση $i \in [n]$ είναι η $\partial_i f : \{0,1\}^n \rightarrow \mathbb{R}$ που ορίζεται από την $\partial_i f(z) = f(x_i = 0, x_{-i} = z_{-i}) - f(x_i = 1, x_{-i} = z_{-i})$. Η κλίση της f , $\nabla f : \{0,1\}^n \rightarrow \mathbb{R}^n$, ορίζεται από την $\nabla f(z) = (\partial_1 f(z), \dots, \partial_n f(z))$.

Σημειώνουμε ότι για κάθε $x \in \{0,1\}^n$ ισχύει ότι $\partial_i f(x) \neq 0$ αν και μόνο αν η συντεταγμένη i είναι ευαίσθητη στο x , και σε αυτήν την περίπτωση η τιμή της είναι είτε 1 ή -1 . Συνεπώς, $\|\partial f(x)\|_2 = \sqrt{h_f(x)}$, και συχνά θα μας είναι βολικό να χρησιμοποιούμε την ℓ_2 -νόρμα της κλίσης της f αντί της $\sqrt{h_f(x)}$.

Λήμμα 6.1.4. Έστω $f : \{0, 1\}^n \rightarrow \mathbb{R}$ και έστω $i \in [n]$. Τότε,

$$\partial_i f(x) = 2 \sum_{S \ni i} \widehat{f}(S) \mathbb{1}_{S \setminus \{i\}}(x).$$

Ειδικότερα, έχουμε ότι $\mathbb{E} \partial_i f(x) = 2 \widehat{f}(\{i\})$.

Στη συνέχεια ορίζουμε το βάρος Fourier επιπέδου d μιας συνάρτησης f καθώς και το προσεγγιστικό βάρος επιπέδου d μιας συνάρτησης f .

Ορισμός 6.1.5. Το βάρος Fourier επιπέδου d της f ορίζεται ως εξής:

$$W_{=d}[f] = \sum_{|S|=d} \widehat{f}(S)^2.$$

Το προσεγγιστικό βάρος επιπέδου d μιας συνάρτησης f ορίζεται ως εξής:

$$W_{\approx d}[f] = \sum_{d \leq j < 2d} W_{=j}[f].$$

6.2 Τυχαίοι περιορισμοί και ισοπεριμετρικές ανισότητες στον διακριτό κύβο

Για μια συνάρτηση $f : \{0, 1\}^n \rightarrow \mathbb{R}$, ένα σύνολο συντεταγμένων $I \subseteq [n]$ και κάποιο $z \in \{0, 1\}^I$, ορίζουμε τον περιορισμό της f ως προς το ζεύγος (I, z) να είναι η συνάρτηση $f_{I \rightarrow z} : \{0, 1\}^{[n] \setminus I} \rightarrow \mathbb{R}$ που ορίζεται από την

$$f_{I \rightarrow z}(y) = f(x_I = z, x_{\bar{I}} = y).$$

Αποκαλούμε τις συντεταγμένες που δεν ανήκουν στο I «ζωντανές», και τις συντεταγμένες του I σταθερές. Θα θεωρήσουμε τυχαίους περιορισμούς μιας συνάρτησης f . Με αυτόν τον όρο εννοούμε ότι το σύνολο I επιλέγεται τυχαία έτσι ώστε κάθε $i \in [n]$ να ανήκει σε αυτό με κάποια πιθανότητα, και το $z \in \{0, 1\}^I$ επιλέγεται ομοιόμορφα. Το επόμενο, πολύ γνωστό, λήμμα εξασφαλίζει μια σχέση ανάμεσα στο προσεγγιστικό βάρος επιπέδου d της f και το βάρος επιπέδου 1 ενός τυχαίου περιορισμού της f στον οποίο κάθε $i \in [n]$ επιλέγεται να ανήκει στο I με πιθανότητα $1 - 1/d$.

Λήμμα 6.2.1. Έστω $f : \{0, 1\}^n \rightarrow \mathbb{R}$ και $d \in \mathbb{N}$, και έστω (J, z) ένας τυχαίος περιορισμός όπου κάθε $j \in [n]$ επιλέγεται να ανήκει στο J με πιθανότητα $\frac{1}{d}$, και το $z \in \{0, 1\}^J$ επιλέγεται ομοιόμορφα. Τότε,

$$\mathbb{E}_{J,z} [W_{=1}[f_{J \rightarrow z}]] \gg W_{\approx d}[f].$$

Απόδειξη. Για σταθερό J και $S \subseteq J$, έχουμε ότι

$$f_{J \rightarrow z}(S) = \sum_{T \subseteq \bar{J}} \widehat{f}(S \cup T) \mathbb{1}_T(z).$$

Συνεπώς,

$$\mathbb{E}_{J,z} [W_{=1}[f_{\bar{J} \rightarrow z}]] = \mathbb{E}_J \left[\mathbb{E}_z \left[\sum_{j \in J} \left(\sum_{S \subseteq \bar{J}} \widehat{f}(S \cup \{j\}) \mathbb{1}_S(z) \right)^2 \right] \right] = \mathbb{E}_J \left[\sum_{j \in J} \sum_{S \subseteq \bar{J}} \widehat{f}(S \cup \{j\})^2 \right],$$

όπου στο τελευταίο βήμα περνάμε τη μέση τιμή ως προς z μέσα και χρησιμοποιούμε την ταυτότητα Parseval. Έπεται ότι

$$\mathbb{E}_{J,z} [W_{=1}[f_{\bar{J} \rightarrow z}]] = \mathbb{E}_J \left[\sum_T \widehat{f}(T)^2 \mathbb{1}_{|T \cap J|=1} \right] = \sum_T \widehat{f}(T)^2 \mathbb{P}_J[|T \cap J| = 1],$$

και καθώς $\mathbb{P}_J[|T \cap J| = 1] \gg 1$ για κάθε T που έχει μέγεθος από d έως $2d$ και οι υπόλοιποι προσθετέοι είναι μη αρνητικοί, η απόδειξη είναι πλήρης. $\square$

Θα χρειαστούμε την ακόλουθη ανισότητα υπερσυσταλτότητας, γνωστή και ως «λήμμα της Bonami».

Θεώρημα 6.2.2. Έστω $f : \{0, 1\}^n \rightarrow \mathbb{R}$ μια συνάρτηση βαθμού d . Τότε,

$$\|f\|_4 \leq \sqrt{3^d} \|f\|_2.$$

Απόδειξη. Υποθέτουμε ότι $d \geq 1$, αλλιώς η f είναι σταθερή και το λήμμα ισχύει τετριμμένα. Η απόδειξη γίνεται με επαγωγή ως προς n . Γράφουμε $f(x) = g(x)(-1)^{x_n} + h(x)$, όπου οι g, h έχουν βαθμό το πολύ $d-1$ και δεν εξαρτώνται από τη μεταβλητή x_n . Για συντομία γράφουμε $f = g \cdot w + h$, όπου $w(x) = (-1)^{x_n}$. Τότε,

$$\begin{aligned} \mathbb{E}[f^4] &= \mathbb{E}[(g \cdot w + h)^4] \\ &= \mathbb{E}[g^4 w^4] + 4\mathbb{E}[g^3 w^3 h] + 6\mathbb{E}[g^2 w^2 h^2] + 4\mathbb{E}[g w h^3] + \mathbb{E}[h^4] \\ &= \mathbb{E}[g^4] \mathbb{E}[w^4] + 4\mathbb{E}[g^3 h] \mathbb{E}[w^3] + 6\mathbb{E}[g^2 h^2] \mathbb{E}[w^2] + 4\mathbb{E}[g h^3] \mathbb{E}[w] + \mathbb{E}[h^4] \\ &= \mathbb{E}[g^4] + 6\mathbb{E}[g^2 h^2] + \mathbb{E}[h^4], \end{aligned}$$

διότι $\mathbb{E}[w] = \mathbb{E}[w^3] = 0$ και $\mathbb{E}[w^2] = \mathbb{E}[w^4] = 1$. Με τον ίδιο τρόπο ελέγχουμε ότι

$$\mathbb{E}[f^2] = \mathbb{E}[g^2] + \mathbb{E}[h^2].$$

Εφαρμόζοντας την επαγωγική υπόθεση για τις g και h έχουμε ότι

$$\mathbb{E}[g^4] = \|g\|_4^4 \leq 9^{d-1} \|g\|_2^4 = 9^{d-1} (\mathbb{E}[g^2])^2$$

διότι η g είναι βαθμού το πολύ $d-1$, και

$$\mathbb{E}[h^4] = \|h\|_4^4 \leq 9^d \|h\|_2^4 = 9^d (\mathbb{E}[h^2])^2$$

διότι η h είναι βαθμού το πολύ d . Επίσης, από την ανισότητα Cauchy-Schwarz έχουμε ότι

$$\mathbb{E}[g^2 h^2] \leq \sqrt{\mathbb{E}[g^4]} \sqrt{\mathbb{E}[h^4]} \leq 3^{d-1} \mathbb{E}[g^2] 3^d \mathbb{E}[h^2].$$

Έπεται ότι

$$\begin{aligned} \mathbb{E}[f^4] &\leq 9^{d-1} \left(\mathbb{E}[g^2] \right)^2 + 6 \cdot 3^{d-1} \mathbb{E}[g^2] 3^d \mathbb{E}[h^2] + 9^d \left(\mathbb{E}[h^2] \right)^2 \\ &\leq 9^d \left(\left(\mathbb{E}[g^2] \right)^2 + 2 \mathbb{E}[g^2] \mathbb{E}[h^2] + \left(\mathbb{E}[h^2] \right)^2 \right) \\ &= 9^d \left(\mathbb{E}[g^2] + \mathbb{E}[h^2] \right)^2 = 9^2 \left(\mathbb{E}[f^2] \mathbb{E}[f^2] \right)^2, \end{aligned}$$

δηλαδή $\|f\|_4 \leq \sqrt{3^d} \|f\|_2$. □

Αποδεικνύουμε τώρα τα Θεωρήματα 6.1.1 και 6.1.2. Ξεκινάμε με ένα πολύ απλό φράγμα για το σύνορο Talagrand, και στη συνέχεια το βελτιώνουμε χρησιμοποιώντας τυχαίους περιορισμούς.

Λήμμα 6.2.3. Έστω $f : \{0, 1\}^n \rightarrow \mathbb{R}$. Τότε, $\mathbb{E} [\|\nabla f(x)\|_2] \geq 2 \sqrt{W_{=1}[f]}$.

Απόδειξη. Από την τριγωνική ανισότητα και το Λήμμα 6.1.4 έχουμε ότι

$$\mathbb{E} [\|\nabla f(x)\|_2] \geq \|\mathbb{E}[\nabla f(x)]\|_2 = \left\| 2(\widehat{f}(\{1\}), \dots, \widehat{f}(\{n\})) \right\|_2 = 2 \sqrt{\sum_{i=1}^n \widehat{f}(\{i\})^2} = 2 \sqrt{W_{=1}[f]}.$$

□

Λήμμα 6.2.4. Έστω $d \in \mathbb{N}$ και $f : \{0, 1\}^n \rightarrow \{0, 1\}$. Τότε, $\mathbb{E} [\|\nabla f(x)\|_2] \mathbb{E} [\|\nabla f(x)\|_2] \gg \sqrt{d} W_{\approx d}[f]$.

Απόδειξη. Έστω (J, z) τυχαίος περιορισμός για την f , όπου κάθε $j \in [n]$ συμπεριλαμβάνεται στο J με πιθανότητα $1/d$, και το $z \in \{0, 1\}^J$ επιλέγεται ομοιόμορφα. Σταθεροποιούμε $x \in \{0, 1\}^n$, και παρατηρούμε ότι

$$\mathbb{E}_J [\|\nabla f_{\bar{J} \rightarrow x_J}(x_J)\|_2] \leq \sqrt{\mathbb{E}_J [\|\nabla f_{\bar{J} \rightarrow x_J}(x_J)\|_2^2]} = \sqrt{\mathbb{E}_J \left[\sum_j |\partial_j f(x)|^2 \mathbb{1}_{j \in J} \right]} = \frac{\|\nabla f(x)\|_2}{\sqrt{d}}.$$

Συνεπώς,

$$(6.2.1) \quad \frac{1}{\sqrt{d}} \mathbb{E} [\|\nabla f(x)\|_2] \geq \mathbb{E}_{J,z} \left[\mathbb{E}_{x \in \{0,1\}^J} [\|\nabla f_{\bar{J} \rightarrow z}(x)\|_2] \right].$$

Από την άλλη πλευρά, για κάθε J και $z \in \{0, 1\}^J$ μπορούμε να εφαρμόσουμε το Λήμμα 6.2.3 για την $f_{\bar{J} \rightarrow z}$ και έχουμε ότι

$$\mathbb{E}_{x \in \{0,1\}^J} [\|\nabla f_{\bar{J} \rightarrow z}(x)\|_2] \gg \sqrt{W_{=1}[f_{\bar{J} \rightarrow z}]} \geq W_{=1}[f_{\bar{J} \rightarrow z}].$$

Παίρνοντας μέση τιμή ως προς J και z βλέπουμε ότι

$$\mathbb{E}_{J,z} \left[\mathbb{E}_{x \in \{0,1\}^J} \left[\|\nabla f_{J \rightarrow z}(x)\|_2 \right] \right] \gg \mathbb{E}_{J,z} \left[W_{=1}[f_{J \rightarrow z}] \right] \gg W_{\approx d}[f],$$

όπου η τελευταία ανισότητα εξασφαλίζεται από το Λήμμα 6.2.1. Συνδυάζοντας αυτό το φράγμα με την (6.2.1) ολοκληρώνουμε την απόδειξη. $\square$

Λήμμα 6.2.5. Έστω $d \in \mathbb{N}$ και $f : \{0,1\}^n \rightarrow \{0,1\}$. Τότε, $\mathbb{E}[\|\nabla f(x)\|_2] \gg \sqrt{d}W_{\geq d}[f]$.

Απόδειξη. Εφαρμόζοντας το Λήμμα 6.2.4 με τον $d2^j$ στη θέση του d για $j = 0, 1, 2, \dots$ παίρνουμε

$$2^{-j/2} \mathbb{E}_J \left[\|\nabla f_{J \rightarrow x_j}(x_J)\|_2 \right] \gg \sqrt{d}W_{\approx 2^j d}[f],$$

και αθροίζοντας πάνω από όλα τα j έχουμε το ζητούμενο. $\square$

Απόδειξη του Θεωρήματος 6.1.1. Αν $\text{Var}(f) \geq 2^{-16}$, τότε αρκεί να δείξουμε ένα κάτω φράγμα $\gg \text{Var}(f)$. Εφαρμόζοντας το Λήμμα 6.2.4 για τα d που είναι δυνάμεις του 2 μεταξύ του 1 και του n , παίρνουμε

$$\sum_{j=1}^{\log n} \frac{1}{\sqrt{2^j}} \mathbb{E}[\|\nabla f(x)\|_2] \gg \sum_{j=1}^{\log n} W_{\approx 2^j}[f] = \text{Var}(f),$$

και από το άθροισμα της γεωμετρικής σειράς το αριστερό μέλος είναι $\ll \mathbb{E}[\|\nabla f(x)\|_2]$.

Υποθέτουμε λοιπόν ότι $\text{Var}(f) \leq 2^{-16}$, και χωρίς περιορισμό της γενικότητας υποθέτουμε ότι η πλειοψηφούσα τιμή της $f(x)$ είναι 0. Τότε, $\mathbb{E}[f] \leq 2\text{Var}(f)$. Θέτοντας $d = \frac{1}{8} \log(1/\text{Var}(f))$, έχουμε ότι

$$\begin{aligned} \sum_{0 < |S| \leq d} \widehat{f}(S)^2 &= \|f^{\leq d}\|_2^2 = \langle f, f^{\leq d} \rangle \leq \|f\|_{4/3} \|f^{\leq d}\|_4 \leq \sqrt{3^d} \mathbb{E}[f]^{3/4} \|f\|_2 \\ &\leq 2^d \mathbb{E}[f]^{5/4} \leq 0.9 \text{Var}(f), \end{aligned}$$

όπου χρησιμοποιήσαμε το φράγμα $\|f^{\leq d}\|_4 \leq \sqrt{3^d} \|f^{\leq d}\|_2$ από το Θεώρημα 6.2.2. Συνεπώς, $W_{\geq d}[f] \geq 0.1 \text{Var}(f)$, και από το Λήμμα 6.2.5 παίρνουμε το αποτέλεσμα. $\square$

Για την απόδειξη του Θεωρήματος 6.1.2 θέτουμε

$$M[f] = \sum_{i=1}^n I_i[f]^2.$$

Θα χρειαστούμε το ακόλουθο αποτέλεσμα από τα [50, 5, 36] (η ακριβής εκδοχή που χρησιμοποιούμε οφείλεται στους Keller και Kindler [36]). Σημειώνουμε ότι το Θεώρημα 2.5.4 είναι μια βελτιωμένη εκδοχή του.

Θεώρημα 6.2.6. Υπάρχουν απόλυτες σταθερές $c_1, c_2 > 0$ τέτοιες ώστε

$$W_{\leq c_1 \log(1/M[f])}[f] \leq M[f]^{c_2}.$$

Μπορούμε τώρα να δώσουμε την απόδειξη του Θεωρήματος 6.1.2.

Απόδειξη του Θεωρήματος 6.1.2. Αν $M[f] \geq \text{Var}(f)^{2/c_2}$, τότε το αποτέλεσμα έπεται από το Θεώρημα 6.1.1. Αλλιώς, θεωρούμε τις σταθερές c_1, c_2 από το Θεώρημα 6.2.6 και θέτουμε $d = c_1 \log(1/M[f])$. Από το Θεώρημα 6.2.6 έχουμε ότι

$$W_{\leq d}[f] \leq M[f]^{c_2} \leq \text{Var}(f)^2 \leq \frac{1}{2} \text{Var}(f),$$

άρα $W_{> d}[f] \geq \frac{1}{2} \text{Var}(f)$, και το αποτέλεσμα έπεται από το Λήμμα 6.2.5. $\square$

Παρατηρούμε ότι συνέπεια του Θεωρήματος 6.1.2 είναι ένα αποτέλεσμα ευστάθειας για το θεώρημα KKL. Αν όλες οι επιρροές μιας Boolean συνάρτησης είναι το πολύ $O(\log n/n)$, τότε η f πρέπει να έχει σταθερού μεγέθους σύνορο κορυφών. Η ακριβής διατύπωση είναι η ακόλουθη.

Πόρισμα 6.2.7. Για κάθε $K > 0$ υπάρχουν $c, \delta > 0$ ώστε να ισχύει το ακόλουθο. Ας υποθέσουμε ότι για την $f : \{0, 1\}^n \rightarrow \{0, 1\}$ ισχύει ότι

$$\max_i I_i[f] \leq K \text{Var}(f) \frac{\log n}{n}.$$

Τότε,

$$\mathbb{P} \left[\|\nabla f(x)\|_2 \geq c \text{Var}(f) \sqrt{\log n} \right] \geq \delta \text{Var}(f).$$

Απόδειξη. Από την υπόθεση έχουμε ότι

$$M[f] \leq K^2 \frac{\log^2 n}{n} \leq \frac{K^2}{\sqrt{n}},$$

άρα το Θεώρημα 6.1.2 συνεπάγεται ότι

$$Z := \mathbb{E} [\|\nabla f(x)\|_2] \geq K_1 \text{Var}(f) \sqrt{\log n}$$

για κάποια σταθερά K_1 που εξαρτάται μόνο από την τιμή της σταθεράς K . Επίσης, έχουμε ότι

$$\mathbb{E} [\|\nabla f(x)\|_2^2] = I_1[f] + \dots + I_n[f] \leq K \text{Var}(f) \log n.$$

Από την ανισότητα Paley-Zygmund έπεται ότι

$$\mathbb{P} \left[\|\nabla f(x)\|_2 \geq \frac{1}{2} Z \right] \geq \frac{3}{4} \frac{Z^2}{\mathbb{E} [\|\nabla f(x)\|_2^2]} \geq \frac{3}{4} \frac{K_1^2}{K} \text{Var}(f),$$

και έχουμε δείξει τον ισχυρισμό του πορίσματος με $c = K_1/2$ και $\delta = \frac{3}{4} \frac{K_1^2}{K}$. $\square$

6.3 Επεκτάσεις και ισχυρότερες εκδοχές

Το Λήμμα 6.2.3 εφαρμόζεται για κάθε L_p -νόρμα και όχι μόνο στην περίπτωση της L_2 -νόρμας.

Λήμμα 6.3.1. Έστω $1 \leq p \leq 2$ και $f : \{0, 1\}^n \rightarrow \{-1, 1\}$. Τότε, $\mathbb{E} [\|\nabla f(x)\|_p] \geq W_{=1}[f]^{1/p}$.

Απόδειξη. Από την ανισότητα Jensen και το Λήμμα 6.2.3 έχουμε

$$\mathbb{E} [\|\nabla f(x)\|_p] = \mathbb{E} [\|\nabla f(x)\|_2^{2/p}] \geq \mathbb{E} [\|\nabla f(x)\|_2]^{2/p} \geq W_{=1}[f]^{1/p}.$$

□

Όπως στο Λήμμα 6.2.4, το προηγούμενο λήμμα μας οδηγεί γρήγορα στο επόμενο συμπέρασμα.

Λήμμα 6.3.2. Έστω $d \in \mathbb{N}$, $\alpha \in [1/2, 1]$ και $f : \{0, 1\}^n \rightarrow \{0, 1\}$. Τότε, $\mathbb{E} [h_f^\alpha] \geq d^\alpha W_{\geq d}[f]$.

Χρησιμοποιώντας το ίδιο επιχείρημα όπως πριν και αντικαθιστώντας τη χρήση του Λήμματος 6.2.5 με τη χρήση του Λήμματος 6.3.2, μπορούμε να αποδείξουμε παραλλαγές των Θεωρημάτων 6.1.1 και 6.1.2 στις οποίες η τετραγωνική ρίζα αντικαθίσταται από οποιαδήποτε δύναμη $p \in [1/2, 1]$.

Δείχνουμε τώρα ότι όταν έχουμε μικρό σύνολο Talagrand τότε έχουμε ευστάθεια ως προς τον θόρυβο. Για μια παράμετρο $\rho \in [0, 1]$ και ένα σημείο $x \in \{0, 1\}^n$, η κατανομή των τιμών που είναι ρ -συσχετισμένες με το x , την οποία συμβολίζουμε με $y \sim_\rho x$, ορίζεται μέσω της ακόλουθης τυχαιοποιημένης διαδικασίας. Για κάθε $i \in [n]$, θέτουμε $y_i = x_i$ με πιθανότητα ρ , αλλιώς επιλέγουμε το $y_i \in \{0, 1\}$ ομοιόμορφα. Ο τελεστής $T_\rho : L_2(\{0, 1\}^n) \rightarrow L_2(\{0, 1\}^n)$ ορίζεται ως εξής:

$$T_\rho f(x) = \mathbb{E}_{y \sim_\rho x} f(y).$$

Ορισμός 6.3.3. Για δοθέν $\varepsilon > 0$, η ευστάθεια θορύβου της f με ρυθμό θορύβου ε είναι η ποσότητα

$$\text{Stab}_{1-\varepsilon}(f) = \langle f, T_{1-\varepsilon} f \rangle.$$

Η σχέση που εξασφαλίζει το Λήμμα 6.3.2 ανάμεσα στο βάρος Fourier της f και την $\mathbb{E} [h_f^\alpha]$ για $1/2 \leq \alpha \leq 1$ συνεπάγεται μια σχέση ανάμεσα σε αυτήν την ποσότητα και την ευστάθεια θορύβου.

Πόρισμα 6.3.4. Υπάρχει απόλυτη σταθερά $C > 0$, τέτοια ώστε για κάθε $\delta \in [0, 1/2]$ και $f : \{0, 1\}^n \rightarrow \{0, 1\}$ που ικανοποιούν την $\mathbb{E} [h_f^{\frac{1}{2}+\delta}] \leq A$ ισχύει ότι

$$\text{Stab}_{1-\varepsilon}(f) \geq 1 - C A \varepsilon^{\frac{1}{2}+\delta}.$$

Απόδειξη. Γράφουμε

$$\begin{aligned} 1 - \text{Stab}_{1-\varepsilon}(f) &= \mathbb{P}_{x,y} (1-\varepsilon)\text{-συσχετισμένα } [f(x) \neq f(y)] = 2\langle 1-f, T_{1-\varepsilon}f \rangle \\ &= 2 \sum_{k=0}^n (1 - (1-\varepsilon)^k) W_{=k}[f]. \end{aligned}$$

Θέτουμε $d = 1/\varepsilon$, χωρίζουμε το άθροισμα σε δύο αθροίσματα, πάνω από τα $k < d$ και $k \geq d$, και φράσσουμε από πάνω τη συνεισφορά κάθε μέρους χωριστά. Για $k \geq d$, χρησιμοποιώντας το Λήμμα 6.3.2 έχουμε ότι

$$2 \sum_{k=d}^n (1 - (1-\varepsilon)^k) W_{=k}[f] \leq 2W_{\geq d}[f] \ll A d^{-(1/2+\delta)} = A\varepsilon^{1/2+\delta}.$$

Για $k \leq d$, γράφουμε

$$\sum_{k=0}^d (1 - (1-\varepsilon)^k) W_{=k}[f] \leq \sum_{k=1}^d k\varepsilon W_{=k}[f] \leq \varepsilon \sum_{k=1}^d W_{\geq k}[f].$$

Χρησιμοποιώντας το Λήμμα 6.3.2 παίρνουμε $W_{\geq k}[f] \ll A k^{-(1/2+\delta)}$, άρα

$$\sum_{k=0}^d (1 - (1-\varepsilon)^k) W_{=k}[f] \ll A \varepsilon \sum_{k=0}^d \frac{1}{k^{1/2+\delta}} \ll A \varepsilon d^{1/2-\delta} \ll A \varepsilon^{1/2+\delta}.$$

□

Τέλος, παρουσιάζουμε κάποιες ισχυρότερες εκδοχές των Θεωρημάτων 6.1.1 και 6.1.2. Σταθεροποιούμε έναν 2-χρωματισμό $\text{col} : E(\{0,1\}^n) \rightarrow \{\text{red}, \text{blue}\}$, και ορίζουμε $\nabla_{\text{red}} f(x)$ το Boolean διάνυσμα που έχει i -οστή συντεταγμένη 1 αν $f(x) = 1$ και $n(x, x \oplus e_i)$ είναι κόκκινη ευαίσθητη ακμή, δηλαδή $f(x) \neq f(x \oplus e_i)$. Όμοια, ορίζουμε το $\nabla_{\text{blue}} f(x)$.

Ορίζουμε επίσης την κόκκινη ευαισθησία $s_{f,\text{red}}(x) = 0$ αν $f(x) = 0$, και αλλιώς $s_{f,\text{red}}(x)$ να είναι το πλήθος των ευαίσθητων κόκκινων ακμών που είναι γειτονικές στο x . Όμοια, ορίζουμε την μπλε ευαισθησία $s_{f,\text{blue}}(x) = 0$ αν $f(x) = 1$, και αλλιώς $s_{f,\text{blue}}(x)$ να είναι το πλήθος των μπλε ευαίσθητων ακμών που είναι γειτονικές στο x . Κατόπιν, ορίζουμε

$$\text{Talagrand}_{\text{col}}(f) = \mathbb{E} \left[\sqrt{s_{f,\text{red}}(x)} \right] + \mathbb{E} \left[\sqrt{s_{f,\text{blue}}(x)} \right].$$

Έχουμε το εξής ανάλογο του Λήμματος 6.2.3.

Λήμμα 6.3.5. Έστω $f : \{0,1\}^n \rightarrow \mathbb{R}$, και $\text{col} : E(\{0,1\}^n) \rightarrow \{\text{red}, \text{blue}\}$ τυχών χρωματισμός. Τότε,

$$\mathbb{E} [\|\nabla_{\text{red}} f(x)\|_2 + \|\nabla_{\text{blue}} f(x)\|_2] \geq \frac{1}{2} \sqrt{W_{=1}[f]}.$$

Απόδειξη. Από την τριγωνική ανισότητα έχουμε

$$\mathbb{E} [\|\nabla_{\text{red}} f(x)\|_2] \geq \|\mathbb{E} [\nabla_{\text{red}} f(x)]\|_2,$$

και κοιτάζοντας το διάνυσμα $v = \mathbb{E} [\nabla_{\text{red}} f(x)]$ βλέπουμε ότι η i -οστή του συντεταγμένη είναι $v_i = \mathbb{P} [f(x) = 1, (x, x \oplus e_i)]$ είναι κόκκινη ευαίσθητη ακμή. Όμοια,

$$\mathbb{E} [\|\nabla_{\text{blue}} f(x)\|_2] \geq \|\mathbb{E} [\nabla_{\text{blue}} f(x)]\|_2$$

και κοιτάζοντας το διάνυσμα $u = \mathbb{E} [\nabla_{\text{red}} f(x)]$ βλέπουμε ότι η i -οστή του συντεταγμένη είναι $u_i = \mathbb{P} [f(x) = 0, (x, x \oplus e_i)]$ είναι μπλε ευαίσθητη ακμή. Έπεται ότι

$$\mathbb{E} [\|\nabla_{\text{red}} f(x)\|_2 + \|\nabla_{\text{blue}} f(x)\|_2] \geq \|v\|_2 + \|u\|_2 \geq \|u + v\|_2.$$

Έστω $R_i(x)$ η δείκτρια του ενδεχομένου η $(x, x \oplus e_i)$ να είναι κόκκινη ευαίσθητη ακμή και $f(x) = 1$, και $B_i(x)$ η δείκτρια του ενδεχομένου η $(x, x \oplus e_i)$ να είναι μπλε ευαίσθητη ακμή και $f(x) = 0$. Τότε,

$$\begin{aligned} u_i + v_i &= \frac{1}{2^{n-1}} \sum_x (R_i(x) + R_i(x \oplus e_i) + B_i(x) + B_i(x \oplus e_i)) \\ &\geq \frac{1}{2^{n-1}} \sum_x \mathbb{1}_{(x, x \oplus e_i) \text{ ευαίσθητη ακμή}} = \frac{1}{2} I_i[f]. \end{aligned}$$

Συνεπώς,

$$\|u + v\|_2 \geq \frac{1}{2} \sqrt{\sum_{i=1}^n I_i[f]^2} \geq \frac{1}{2} W_{=1}[f].$$

Έτσι, έχουμε αποδείξει το λήμμα. $\square$

Τώρα, για κάθε f , κάθε χρωματισμό col και κάθε x , επιλέγοντας $J \subseteq [n]$ τυχαία ώστε κάθε συντεταγμένη να συμπεριλαμβάνεται με πιθανότητα $1/d$, έχουμε ότι

$$\mathbb{E}_J \left[\sqrt{s_{fJ \rightarrow x_J, \text{red}}(x_J)} \right] \leq \sqrt{\mathbb{E}_J \left[s_{fJ \rightarrow x_J, \text{red}}(x_J) \right]} = \sqrt{\frac{s_{f, \text{red}}(x)}{d}},$$

και όμοια για τις μπλε ευαίσθητες ακμές. Έτσι παίρνουμε το ανάλογο της (6.2.1) για τις κόκκινες και μπλε ευαίσθητες ακμές, και επαναλαμβάνοντας το επιχείρημα των Λημμάτων 6.2.4 και 6.2.5 συμπεραίνουμε ότι

$$\text{Talagrand}_{\text{col}}(f) \gg \sqrt{d} W_{\geq d}[f]$$

για κάθε χρωματισμό col . Οι ισχυρές εκδοχές των Θεωρημάτων 6.1.1 και 6.1.2 έπονται τώρα, αφού έχουμε δείξει ότι

$$W_{\geq d}[f] \geq \Omega(\text{Var}(f))$$

αν επιλέξουμε είτε $d = \Omega(\log(1/\text{Var}(f)))$ ή $d = \Omega(\log(1/M(f)))$.

Ακριβείς ισοπεριμετρικές ανισότητες στον διακριτό κύβο

7.1 Εισαγωγή

Έστω $n \geq 1$ ένας φυσικός αριθμός και $\{0, 1\}^n$ ο διακριτός κύβος διάστασης n . Μπορούμε να βλέπουμε τον $\{0, 1\}^n$ ως ένα γράφημα, όπου δύο κορυφές $x, y \in \{0, 1\}^n$ συνδέονται με ακμή αν τα διανύσματα $x = (x_1, \dots, x_n)$ και $y = (y_1, \dots, y_n)$ διαφέρουν σε ακριβώς μία συντεταγμένη. Συμβολίζουμε αυτήν την ακμή με (x, y) . Σε αυτό το κεφάλαιο παρουσιάζουμε πρόσφατη εργασία των Beltran, Ivanisvili και Madrid [4] στην οποία αποδεικνύονται διάφορες ακριβείς ισοπεριμετρικές ανισότητες για υποσύνολα του $\{0, 1\}^n$.

7.1.1 Μονόπλευρο σύνορο

Έστω $A \subseteq \{0, 1\}^n$ και $A^c := \{0, 1\}^n \setminus A$. Ορίζουμε το σύνορο ακμών του A να είναι το $\nabla A := \{(x, y) : x \in A, y \in A^c\}$. Στο σύνορο ακμών ∇A του A αντιστοιχεί φυσιολογικά η συνάρτηση $h_A : \{0, 1\}^n \rightarrow \mathbb{R}$ που ορίζεται ως εξής: $h_A(x) = 0$ αν $x \in A^c$, ενώ αν $x \in A$ τότε $h_A(x)$ είναι το πλήθος των ακμών που συνδέουν το x με κορυφή που ανήκει στο A^c . Έστω μ_n το ομοιόμορφο μέτρο πιθανότητας στον $\{0, 1\}^n$ και $\mathbb{E}$ ο τελεστής της μέσης τιμής. Ένας από τους στόχους μας είναι να μελετήσουμε την ποσότητα

$$(7.1.1) \quad B(t, p, n) = \left\{ \mathbb{E}[h_A^p] : A \subseteq \{0, 1\}^n, \mu_n(A) = t \right\}$$

για κάθε $p > 0$, $t \in [0, 1]$ και $n \geq 1$. Σημειώνουμε ότι για σταθερό $n \geq 1$, οι επιτρεπτές τιμές για το $t \in [0, 1]$ είναι της μορφής $t = m/2^n$ για ακραίους $0 \leq m \leq 2^n$.

Όταν $p = 1$, η ποσότητα $\mathbb{E}[h_A]$ έχει πρόσθετη δομή, δηλαδή, $\mathbb{E}[h_A] = \mathbb{E}[h_{A^c}]$, και ισχύει ότι $\mathbb{E}[h_A] = \frac{|\nabla A|}{2^n}$, όπου με $|\cdot|$ συμβολίζουμε τον πληθάριθμο ενός συνόλου. Σε αυτήν την

περίπτωση, ο Hart [28] έχει αποδείξει ότι

$$(7.1.2) \quad B(t, 1, n) = nt - \frac{1}{2^{n-1}} \sum_{j=1}^{2^n t - 1} s(j),$$

όπου $s(j)$ είναι το πλήθος των ψηφίων 1 στη δυαδική αναπαράσταση του ακεραίου j . Το ελάχιστο στην (7.1.1) επιτυγχάνεται για το σύνολο A μέτρου $\mu_n(A) = t$ που αποτελείται από εκείνες τις κορυφές $x = (x_1, \dots, x_n) \in \{0, 1\}^n$ για τις οποίες $\sum_{j=1}^n 2^{j-1} x_j = \ell$, για κάθε $\ell = 1, 2, \dots, t2^n$. Η αντίστοιχη ισοπεριμετρική ανισότητα που προκύπτει από την (7.1.2) έχει εφαρμογές στη θεωρία παιγνίων [28], τους κατανεμμένους αλγορίθμους, την πολυπλοκότητα επικοινωνιών και την επιστήμη δικτύων (βλέπε [46] και τις αναφορές εκεί).

Παρόλο που η ταυτότητα (7.1.2) δίνει την ακριβή τιμή της $B(t, 1, n)$, συχνά είναι πιο βολικό να χρησιμοποιούμε το κάτω φράγμα

$$(7.1.3) \quad \mathbb{E}[h_A] \geq \mu_n(A)^* \log_2 \left(\frac{1}{\mu_n(A)^*} \right),$$

όπου $t^* = \min\{t, 1 - t\}$ για κάθε $t \in [0, 1]$. Αυτή η ανισότητα είναι γνωστή ως η κλασική ισοπεριμετρική ανισότητα για τον διακριτό κύβο.

Ένα πλεονέκτημα της ανισότητας $B(t, 1, n) \geq t^* \log_2(1/t^*)$ είναι ότι (α) το δεξιό μέλος είναι ανεξάρτητο από τη διάσταση n και (β) γίνεται ισότητα στα σημεία $t = 2^{-k}$ για $0 \leq k \leq n$ (η ανισότητα (7.1.3) γίνεται κι αυτή ισότητα όταν $\mu_n(A) = 1 - 2^{-k}$, λόγω της ταυτότητας $\mathbb{E}[h_A] = \mathbb{E}[h_{A^c}]$). Σε αυτήν την περίπτωση ($p = 1$, $t = 2^{-k}$), το ελάχιστο στην (7.1.1) επιτυγχάνεται στους υπο-κύβους συνδιάστασης k , για κάθε $0 \leq k \leq n$. Σε ό,τι ακολουθεί, ενδιαφερόμαστε για το $\inf_{n \geq 1} B(t, p, n)$.

Θεωρώντας Hamming μπάλες $\{x_1 + \dots + x_n \leq [n/2]\}$ στην (7.1.1) βλέπουμε ότι

$$(7.1.4) \quad \text{αν } p < 1/2 \text{ τότε } \inf_{n \geq 1} B(t, p, n) = 0$$

(βλ. Παρατήρηση 7.1.5). Από την άλλη πλευρά, θεωρώντας υπο-κύβους στην (7.1.1) βλέπουμε ότι

$$(7.1.5) \quad t(\log_2(1/t))^p \geq \inf_{n \geq 1} B(t, p, n) \quad \text{στα σημεία } t = 2^{-k}, \quad k = 0, 1, 2, \dots$$

Εφαρμόζοντας την ανισότητα Hölder στην (7.1.3) παίρνουμε

$$\mathbb{E}[h_A^p] \geq \mu_n(A) \left(\log_2 \left(\frac{1}{\mu_n(A)} \right) \right)^p \quad \text{για κάθε } p \geq 1,$$

και συνδυάζοντας αυτή την ανισότητα με την (7.1.5) συμπεραίνουμε ότι $\inf_{n \geq 1} B(t, p, n) = t(\log_2(1/t))^p$ όταν $t = 2^{-k}$ για κάθε $k \geq 0$ και κάθε $p \geq 1$. Από εδώ και πέρα, θα μας απασχολήσει η περίπτωση $1/2 \leq p < 1$.

Η περίπτωση $p = 1/2$ μελετήθηκε για πρώτη φορά από τον Talagrand στο [48]. Με

ένα επαγωγικό επιχείρημα, απέδειξε την αντίστροφη ανισότητα της (7.1.5) έως μια απόλυτη σταθερά $K > 1$, δηλαδή έδειξε ότι

$$(7.1.6) \quad \mathbb{E} \left[\sqrt{h_A} \right] \geq K \mu_n(A)^* \sqrt{\log_2(1/\mu_n(A)^*)}$$

για κάθε $A \subseteq \{0, 1\}^n$. Επιπλέον, απέδειξε ότι

$$(7.1.7) \quad \mathbb{E} \left[\sqrt{h_A} \right] \geq \sqrt{2} \mu_n(A)(1 - \mu_n(A))$$

για κάθε $A \subseteq \{0, 1\}^n$. Οι Bobkov και Götze [8] έδωσαν ένα βελτιωμένο και κομψό επαγωγικό επιχείρημα που οδηγεί στην (7.1.7) και, ειδικότερα, ένα από τα αποτελέσματά τους συνεπάγεται ότι η σταθερά $\sqrt{2}$ μπορεί να αντικατασταθεί από τη σταθερά $\sqrt{3}$. Αργότερα, οι Kahn και Park [33] έδειξαν ότι

$$(7.1.8) \quad \mathbb{E} \left[h_A^{\log_2(3/2)} \right] \geq 2 \mu_n(A)(1 - \mu_n(A)),$$

και η σταθερά 2 στο δεξιό μέλος της (7.1.8) είναι βέλτιστη: η ανισότητα γίνεται ισότητα στους υπο-κύβους συνδιάστασης 1 και 2. Μάλιστα, για κάθε δοσμένο $p \geq 1/2$, η μεγαλύτερη δυνατή σταθερά C_p για την οποία ισχύει ότι $\mathbb{E} \left[h_A^p \right] \geq C_p \mu_n(A)(1 - \mu_n(A))$ για κάθε $A \subseteq \{0, 1\}^n$ πρέπει να ικανοποιεί την $C_p \leq 2$. Αυτό προκύπτει αν δοκιμάσουμε την ανισότητα σε έναν ημι-κύβο. Επιπλέον, θεωρώντας υπο-κύβους συνδιάστασης 2 βλέπουμε ότι η σταθερά $C_p = 2$ είναι δυνατή μόνο αν $p \geq \log_2(3/2)$.

Οι Beltran, Ivanisvili και Madrid απέδειξαν ένα θεώρημα το οποίο, ειδικότερα, αντιστρέφει την ανισότητα (7.1.5) για κάθε $p \geq \log_2(3/2) = 0.5849 \dots$ και συνεπώς δίνει την ακριβή τιμή του $\inf_{n \geq 1} B(t, p, n)$ γι' αυτό το εύρος τιμών του p για όλα τα $t = 2^{-k}$, $k \geq 0$.

Θεώρημα 7.1.1. Έστω $p_0 := \log_2(3/2)$. Ισχύει ότι

$$(7.1.9) \quad \mathbb{E} \left[h_A^{p_0} \right] \geq \mu_n(A)^* (\log_2(1/\mu_n(A)^*))^{p_0}$$

για κάθε $A \subseteq \{0, 1\}^n$. Ειδικότερα, αν $\mu_n(A) \leq 1/2$ έχουμε

$$(7.1.10) \quad \mathbb{E} \left[h_A^p \right] \geq \mu_n(A) (\log_2(1/\mu_n(A)))^p \quad \text{για κάθε } p \geq \log_2(3/2).$$

Ισότητα στις (7.1.9) και (7.1.10) ισχύει για κάθε υπο-κύβο $A \subseteq \{0, 1\}^n$.

Το Θεώρημα 7.1.1 είναι μια ακριβής εκδοχή της (7.1.8) σε όλους τους υπο-κύβους. Η σύγκριση ανάμεσα στις ανισότητες (7.1.8) και (7.1.9) δείχνει ότι η συνάρτηση $2t(1-t)$ είναι μεγαλύτερη ή ίση της $t^*(\log_2(1/t^*))^{p_0}$ στο $[1/4, 3/4]$, ενώ η δεύτερη συνάρτηση είναι μεγαλύτερη ή ίση της πρώτης στο συμπληρωματικό εύρος τιμών. Γενικά, το ίδιο φαινόμενο συμβαίνει όταν συγκρίνουμε τις $2t(1-t)$ και $t^*(\log_2(1/t^*))^p$ για άλλες τιμές του $1/2 \leq p \leq 1$: η δεύτερη είναι αισθητά μεγαλύτερη σε περιοχές των $t = 0$ και $t = 1$ και, επιπλέον, ικανοποιεί την (7.1.5). Είναι ενδιαφέρον το ερώτημα να κατανοήσουμε αν η ανισότητα (7.1.9) ισχύει για κάθε $p_0 \geq 1/2$.

Έστω $\partial A = \text{supp}(h_A) \subseteq A$ το σύνολο κορυφών του A . Εφαρμόζοντας την ανισότητα Hölder

$$\mathbb{E}[h_A^p] \leq \left(\mathbb{E}[h_A^{p/q}]\right)^q \mu_n(\partial A)^{1-q}$$

που ισχύει για κάθε $q \in [0, 1]$ παίρνουμε το ακόλουθο πόρισμα, το οποίο ισχυροποιεί την κλασική ισοπεριμετρική ανισότητα (7.1.3).

Πόρισμα 7.1.2. Αν $\gamma = \frac{1}{\log_2(3/2)} - 1 = 0.709 \dots$ τότε για κάθε $A \subseteq \{0, 1\}^n$ με $\mu_n(A) \leq 1/2$ έχουμε ότι

$$(7.1.11) \quad \mathbb{E}[h_A] \geq \left(\frac{\mu_n(A)}{\mu_n(\partial A)}\right)^\gamma \mu_n(A) \log_2\left(\frac{1}{\mu_n(A)}\right).$$

Επιπλέον, ισχύει ισότητα για κάθε υπο-κύβο.

Η σταθερά $\sqrt{3}$ στην ανισότητα του Bobkov $\mathbb{E}[\sqrt{h_A}] \geq \sqrt{3}\mu_n(A)(1-\mu_n(A))$ δεν είναι βέλτιστη. Θα δούμε ότι ισχύει η ακόλουθη ισχυρότερη ανισότητα.

Θεώρημα 7.1.3. Για κάθε $p \in [1/2, \log_2(3/2)]$ ισχύει ότι

$$(7.1.12) \quad \mathbb{E}[h_A^p] \geq C_p \mu_n(A)(1-\mu_n(A))$$

για κάθε $A \subseteq \{0, 1\}^n$, όπου $C_p = 2\sqrt{2^{p+1}-2}$.

Εφαρμόζοντας το Θεώρημα 7.1.3 για την περίπτωση $p = 1/2$ παίρνουμε τη βελτιωμένη σταθερά $C_{1/2} = 1.82 \dots > \sqrt{3} = 1.73 \dots$. Δοκιμάζοντας την ανισότητα $\mathbb{E}[h_A^p] \geq C_p \mu_n(A)(1-\mu_n(A))$ σε υπο-κύβους συνδιάστασης 2, παίρνουμε το άνω φράγμα $C_p \leq 2^{p+2}/3$ για $1/2 \leq p \leq \log_2(3/2)$. Αν $p = \log_2(3/2)$ τότε η ανισότητα (7.1.12) συμπίπτει με την (7.1.8).

7.1.2 Διαμερίσεις του κύβου

Αν θεωρήσουμε σύνολα $A \subseteq \{0, 1\}^n$ με μέτρο $\mu_n(A) \geq 1/2$, θα μπορούσαμε να εικάσουμε ότι η ανισότητα $\mathbb{E}[h_A^p] \geq 2\mu_n(A)(1-\mu_n(A))$ ισχύει για ένα μεγαλύτερο εύρος εκθετών $p \geq 1/2$. Σε αυτήν την κατεύθυνση οι Beltran, Ivanisvili και Madrid απέδειξαν το εξής.

Θεώρημα 7.1.4. Για κάθε $A \subseteq \{0, 1\}^n$,

$$(7.1.13) \quad \mathbb{E}[h_A^{0.53}] \geq 8\mu_n(A)(1-\mu_n(A)) \left[\left(1 - \frac{2\sqrt{2}}{3}\right)\mu_n(A) + \frac{\sqrt{2}}{3} - \frac{1}{4} \right].$$

Με άμεση σύγκριση του δεξιού μέλους της (7.1.13) με την ποσότητα $2\mu_n(A)(1-\mu_n(A))$ παίρνουμε μια επέκταση της (7.1.8).

Πόρισμα 7.1.5. Αν $\mu_n(A) \geq 1/2$ τότε

$$\mathbb{E}[h_A^{0.53}] \geq 2\mu_n(A)(1-\mu_n(A)).$$

Το Πόρισμα 7.1.5 βρίσκει άμεση εφαρμογή σε ισοπεριμετρικές ανισότητες που αφορούν διαμερίσεις του $\{0, 1\}^n$. Έστω (A, B, W) μια διαμέριση του $\{0, 1\}^n$, όπου τυπικά σκεφτόμαστε ότι το W είναι μικρό. Σε αυτό το πλαίσιο, ορίζουμε $\nabla(A, B) = \{(x, y) : x \in A, y \in B\}$. Για $p \geq 1/2$, οι Kahn και Park [33] μελέτησαν ανισότητες της μορφής

$$(7.1.14) \quad |\nabla(A, B)| + Kn^p|W| \geq 2^{n-1}$$

για σύνολα A με μέτρο $\mu_n(A) = 1/2$, όπου $K > 0$ απόλυτη σταθερά. Σημειώνουμε ότι αν $W = \emptyset$ (δηλαδή, $B = A^c$), τότε η (7.1.14) γίνεται $|\nabla A| \geq |A|$, που είναι η κλασική ισοπεριμετρική ανισότητα (7.1.3) για σύνολα μέτρου $\mu_n(A) = 1/2$. Έκαναν την εικασία (βλ. [33, Εικασία 1.3]) ότι η (7.1.14) ισχύει για $p = 1/2$. Εδώ, είναι πάλι αναγκαίο να έχουμε $p \geq 1/2$, κάτι που βλέπουμε δοκιμάζοντας Hamming μπάλες (βλ. Παρατήρηση 7.1.5). Με ένα στοιχειώδες επιχείρημα (δείτε το τέλος της Ενότητας 7.2.3 για πληρότητα), έδειξαν ότι η ανισότητα (7.1.8) συνεπάγεται την (7.1.14) για $p = \log_2(3/2)$ και $K = 1$. Το Πόρισμα 7.1.5 δίνει περαιτέρω βελτίωση για τον εκθέτη p .

Πόρισμα 7.1.6. Έστω (A, B, W) μια διαμέριση του $\{0, 1\}^n$ τέτοια ώστε $\mu_n(A) = 1/2$. Τότε,

$$|\nabla(A, B)| + n^{0.53}|W| \geq 2^{n-1}.$$

7.1.3 Αμφίπλευρο σύνολο

Για κάθε $A \subseteq \{0, 1\}^n$ ορίζουμε την αμφίπλευρη συνάρτηση συνόρου $\zeta_A : \{0, 1\}^n \rightarrow \mathbb{R}$ θέτοντας $\zeta_A(x) = h_A(x)$ αν $x \in A$ και $\zeta_A(x) = h_{A^c}(x)$ αν $x \in A^c$. Παρατηρούμε ότι μπορούμε να πάρουμε φράγματα για την $\mathbb{E}[\zeta_A^p]$ αν έχουμε φράγματα για την $\mathbb{E}[h_A^p]$, χρησιμοποιώντας την ταυτότητα $\zeta_A^p = h_A^p + h_{A^c}^p$. Πράγματι, αν υπάρχει μια συνάρτηση $G : [0, 1] \rightarrow \mathbb{R}$ τέτοια ώστε $\mathbb{E}[h_A^p] \geq G(\mu_n(A))$ για κάθε $A \subseteq \{0, 1\}^n$, τότε

$$\mathbb{E}[\zeta_A^p] \geq G(\mu_n(A)) + G(1 - \mu_n(A)).$$

Ειδικότερα, χρησιμοποιώντας το Θεώρημα 7.1.1 παίρνουμε

$$\inf \left\{ \mathbb{E}[\zeta_A^{p_0}] : A \subseteq \{0, 1\}^n, \mu_n(A) = t \right\} \geq 2t^*(\log_2(1/t^*))^{p_0}, \quad p_0 = \log_2(3/2),$$

και συνδυάζοντας το Θεώρημα 7.1.1 με το Θεώρημα 7.1.4 έχουμε ότι

$$\inf \left\{ \mathbb{E}[\zeta_A^p] : A \subseteq \{0, 1\}^n, \mu_n(A) = t \right\} \geq t^*(\log_2(1/t^*))^p + P(t_*) \quad \text{για } p > \log_2(3/2),$$

όπου $t_* := \max(t, 1-t)$ και P είναι το κυβικό πολυώνυμο στο δεξιό μέλος της (7.1.13). Στην ειδική περίπτωση $\mu_n(A) = 1/2$ παίρνουμε ακριβή κάτω φράγματα για κάθε $p \geq 0.53$ χρησιμοποιώντας το Πόρισμα 7.1.5.

Πόρισμα 7.1.7. Για κάθε $p \geq 0.53$,

$$\inf \left\{ \mathbb{E}[\zeta_A^p] : A \subseteq \{0, 1\}^n, \mu_n(A) = 1/2 \right\} = 1.$$

Επιπλέον, για κάθε $0.5 \leq p < 0.53$,

$$(7.1.15) \quad 1 \geq \inf \left\{ \mathbb{E}[\zeta_A^p] : A \subseteq \{0, 1\}^n, \mu_n(A) = 1/2 \right\} \geq \sqrt{2^{p+1} - 2}.$$

Το άνω φράγμα $1 \geq \inf \left\{ \mathbb{E}[\zeta_A^p] : A \subseteq \{0, 1\}^n, \mu_n(A) = 1/2 \right\}$ προκύπτει αν δοκιμάσουμε τον ημικόβο και το όχι ακριβές κάτω φράγμα στην (7.1.15) για το εύρος τιμών $0.5 \leq p < 0.53$ έπεται από το Θεώρημα 7.1.3. Αυτό το αποτέλεσμα βελτιώνει εκτιμήσεις των Ivanisvili, Nazarov και Volberg από το [29, Ενότητα 3.6] που δείχνουν ότι

$$1 \geq \inf \left\{ \mathbb{E}[\zeta_A^p] : A \subseteq \{0, 1\}^n, \mu_n(A) = 1/2 \right\} \geq \max \left\{ \sqrt{\frac{2}{\pi}}, s_{(2p)'}^{2p} \right\} \quad \text{για } 1/2 \leq p \leq 1.$$

Εδώ, για κάθε $1 \leq q < \infty$ ορίζουμε $q' = \frac{q}{q-1}$, και s_q είναι η μικρότερη θετική ρίζα της υπεργεωμετρικής συνάρτησης $x \mapsto {}_1F_1\left(-\frac{q}{2}, \frac{1}{2}, \frac{x^2}{2}\right)$. Παρατηρούμε ότι $s_{p'} \geq \sqrt{\frac{2(p-1)}{p}}$ για κάθε $p \in [1, 2]$ και $s_{p'} \rightarrow 0$ ως $p \rightarrow 1$. Το Πρόγραμμα 7.1.7 για $p = 1/2$ δίνει

$$\inf \left\{ \mathbb{E} \left[\sqrt{\zeta_A} \right] : A \subseteq \{0, 1\}^n, \mu_n(A) = 1/2 \right\} \geq \sqrt{2^{3/2} - 2} = 0.91 \dots > \sqrt{\frac{2}{\pi}} = 0.79 \dots$$

Παρατήρηση 7.1.8. Έστω C_1 η μεγαλύτερη σταθερά στην L^1 ανισότητα Poincaré

$$(7.1.16) \quad C \mathbb{E} [|f - \mathbb{E}f|] \leq \mathbb{E} [|\nabla f|]$$

που ισχύει για κάθε $f : \{-1, 1\}^n \rightarrow \mathbb{R}$ και κάθε $n \geq 1$ (βλέπε Ενότητα 7.1.4 για τον ορισμό της $|\nabla f|$). Έστω $C_{1,G}$ η μεγαλύτερη σταθερά για την οποία η (7.1.16) ισχύει για κάθε $f : \{-1, 1\}^n \rightarrow \{0, 1\}$. Είναι σαφές ότι $C_1 \leq C_{1,G}$. Οι Ivanisvili, van Handel και Volberg απέδειξαν στο [30] ότι $\frac{2}{\pi} < C_1 \leq \sqrt{\frac{2}{\pi}}$, και εικάζεται ότι $C_1 = \sqrt{\frac{2}{\pi}}$. Θα μπορούσε κανείς να υποπτευθεί ότι $C_1 = C_{1,G}$. Όμως, παρατηρήστε ότι για μια δείκτρια συνάρτηση $f = 1_A$, $A \subseteq \{-1, 1\}^n$, έχουμε

$$\frac{\mathbb{E} [|\nabla f|]}{\mathbb{E} [|f - \mathbb{E}f|]} = \frac{\mathbb{E} [\sqrt{\zeta_A}]}{4\mu_n(A)(1 - \mu_n(A))}$$

(βλέπε Ενότητα 7.1.4 για τη σχέση ανάμεσα στην $|\nabla f|$ και την ζ_A). Από την άλλη πλευρά, αν συνδυάσουμε το Θεώρημα 7.1.3 με το γεγονός ότι $\sqrt{\zeta_A} = \sqrt{h_A} + \sqrt{h_{A^c}}$ βλέπουμε ότι $\sqrt{2^{3/2} - 2} \leq C_{1,G}$ το οποίο συνεπάγεται ότι $C_{1,G} - C_1 \geq \sqrt{2^{3/2} - 2} - \sqrt{2/\pi} = 0.112 \dots > 0$.

7.1.4 Οι ισοπερμετρικές ανισότητες του Talagrand για τη διακριτή κλίση

Για τα αποτελέσματα αυτής της ενότητας θεωρούμε τον διακριτό κύβο $\{-1, 1\}^n$ αντί του $\{0, 1\}^n$. Οι συναρτήσεις h_A και ζ_A ορίζονται ανάλογα για σύνολα $A \subseteq \{-1, 1\}^n$. Για κάθε $f : \{-1, 1\}^n \rightarrow \mathbb{R}$ και κάθε $x = (x_1, \dots, x_n) \in \{-1, 1\}^n$ θέτουμε

$$D_j f(x) = \frac{f(x_1, \dots, x_j, \dots, x_n) - f(x_1, \dots, -x_j, \dots, x_n)}{2}.$$

Ονομάζουμε Boolean μια συνάρτηση $f : \{-1, 1\}^n \rightarrow \mathbb{R}$ που παίρνει μόνο τις τιμές -1 ή 1 . Ορίζουμε

$$|\nabla f(x)| = \sqrt{\sum_{j=1}^n |D_j f|^2}, \quad \text{και} \quad |Mf(x)| = \sqrt{\sum_{j=1}^n (D_j f)_+^2},$$

όπου $(a)_+ = \max\{0, a\}$ για κάθε $a \in \mathbb{R}$. Για κάθε $A \subseteq \{-1, 1\}^n$ συμβολίζουμε με $\mathbb{1}_A$ τη δείκτρια συνάρτηση του A .

Παρατηρήστε ότι αν $x \notin A$ τότε $|M\mathbb{1}_A(x)|^2 = 0$, και αν $x \in A$ τότε η $|M\mathbb{1}_A(x)|^2$ ισούται με $1/4$ φορές το πλήθος των γειτόνων του x στο A^c . Συνεπώς, $|M\mathbb{1}_A(x)|^2 = h_A(x)/4$. Από την άλλη πλευρά, η «αμφίπλευρη» κλίση $|\nabla \mathbb{1}_A|^2$ παίρνει τη μορφή $|\nabla \mathbb{1}_A(x)|^2 = h_A(x)/4 + h_{A^c}(x)/4 = \zeta_A(x)/4$.

Κάθε συνάρτηση $f : \{-1, 1\}^n \rightarrow \mathbb{R}$ έχει μια αναπαράσταση Fourier-Walsh

$$f(x) = \sum_{S \subseteq \{1, \dots, n\}} \widehat{f}(S) w_S(x), \quad w_S(x) = \prod_{j \in S} x_j.$$

Για $p \geq 1$ ορίζουμε $\|f\|_p = (\mathbb{E}[|f|^p])^{1/p}$. Στο προηγούμενο κεφάλαιο συζητήσαμε το ακόλουθο θεώρημα των Eldan, Kindler, Lifshitz και Minzer από το [21].

Θεώρημα 7.1.9. Για κάθε $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$ και κάθε $p \in [1, 2]$ ισχύει ότι

$$(7.1.17) \quad \|\nabla f\|_p \gtrsim \sup_{d \geq 0} \left(\sum_{|S| \geq d} |\widehat{f}(S)|^2 \right)^{1/p} \sqrt{d}.$$

Επιλέγοντας $d \approx \log\left(\frac{e}{\text{Var}(f)}\right)$, από την ανισότητα (7.1.17) παίρνουμε την ακριβή ισοπεριμετρική ανισότητα του Talagrand

$$(7.1.18) \quad \|\nabla f\|_p \gtrsim (\text{Var}(f))^{1/p} \sqrt{\log\left(\frac{e}{\text{Var}(f)}\right)}$$

για κάθε $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$, και επιλέγοντας $d \approx \log\left(\frac{e}{W(f)}\right)$, όπου

$$W(f) = \sum_{j=1}^n \left(\mathbb{E}[|D_j f|] \right)^2,$$

από την ανισότητα (7.1.17) παίρνουμε την εικασία του Talagrand (υπενθυμίζουμε ότι η περίπτωση $p \in (1, 2]$ ήταν γνωστή στον Talagrand, ο οποίος ρωτούσε τι συμβαίνει στην περίπτωση $p = 1$)

$$(7.1.19) \quad \|\nabla f\|_p \gtrsim (\text{Var}(f))^{1/p} \sqrt{\log\left(\frac{e}{W(f)}\right)}$$

για κάθε $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$, την οποία απέδειξαν πρώτοι οι Eldan και Gross στο [20]. Παρατηρούμε ότι για μονότονες Boolean συναρτήσεις f , δηλαδή αυτές που ικανοποιούν την $f(x) \geq f(y)$ όταν $x_i \geq y_i$ για κάθε $i = 1, \dots, n$, έχουμε

$$W(f) = \sum_{j=1}^n \widehat{f}(\{j\})^2 \leq \sum_{S \neq \emptyset} \widehat{f}(S)^2 = \text{Var}(f).$$

Παρατήρηση 7.1.10. Πρέπει να σημειώσουμε ότι η (7.1.19) για $p = 1$ προκύπτει από το [40, Θεώρημα 31], το οποίο ισχυρίζεται ότι $\text{NS}_\delta(f) \lesssim \sqrt{\delta} \mathbb{E}|\nabla f|$ για κάθε $\delta \in [0, 1]$ και κάθε $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$, όπου $\text{NS}_\delta(f)$ είναι η ευστάθεια θορύβου της f . Χρησιμοποιώντας το φράγμα

$$\text{NS}_\delta(f) \gtrsim \text{Var}(f)(1 - (W(f)/\text{Var}(f))^{\frac{\delta}{2-\delta}})$$

από το [36], και επιλέγοντας

$$\delta = \frac{1}{\log(\text{Var}(f)/W(f))},$$

όπου μπορούμε να υποθέσουμε ότι $\text{Var}(f)/W(f) > 100$ λόγω της (7.1.18), παίρνουμε

$$\mathbb{E}|\nabla f| \gtrsim \text{Var}(f) \sqrt{\log\left(\frac{\text{Var}(f)}{W(f)}\right)}.$$

Επικαλούμενοι πάλι την (7.1.18) μπορούμε να υποθέσουμε ότι $\text{Var}(f) \geq \sqrt{W(f)}$ και έτσι έχουμε άλλη μία απόδειξη της (7.1.19).

Σε αυτό το κεφάλαιο θα δώσουμε νέα απόδειξη της (7.1.17), η οποία σε αντίθεση με τις μεθόδους τυχαίου περιορισμού του [21] ή τις επαγωγικές αποδείξεις του [40] επεκτείνει τα αποτελέσματα σε συναρτήσεις με τιμές σε τυχόντα χώρο Banach που έχει πεπερασμένο συν-τύπο. Υπενθυμίζουμε ότι ένας χώρος Banach $(X, \|\cdot\|)$ έχει συν-τύπο Rademacher $q \in [2, \infty]$ αν υπάρχει θετική σταθερά C τέτοια ώστε

$$\sum_{j=1}^n \|x_j\|^q \leq C^q \mathbb{E} \left\| \sum_{j=1}^n \varepsilon_j x_j \right\|^q$$

για κάθε $n \geq 1$ και κάθε $x_1, \dots, x_n \in X$, όπου $\varepsilon_1, \dots, \varepsilon_n$ είναι ανεξάρτητες ισοκατανεμημένες συμμετρικές ± 1 τυχαίες μεταβλητές Bernoulli. Η βέλτιστη σταθερά $C = C_q(X)$ ονομάζεται σταθερά συν-τύπου q του X . Έστω $\Delta = \sum_{j=1}^n D_j$, και έστω $e^{-t\Delta}$ η ημιομάδα θερμότητας.

Θεώρημα 7.1.11. Έστω $f : \{-1, 1\}^n \rightarrow X$, $\|f\|_\infty \leq 1$. Τότε, για κάθε $p \in [1, 2]$ έχουμε

$$\|Df\|_p \gtrsim \frac{1}{q^{3/2} C_q(X)} \|f\|_2^{2/p} \left(\log \left(\frac{e\|f\|_2}{\|f\|_1} \right) \right)^{1/q},$$

όπου $\|Df\|_p^p = \mathbb{E} \left[\left\| \sum_{1 \leq j \leq n} x'_j D_j f(x) \right\|^p \right]$ και x' είναι ένα ανεξάρτητο αντίγραφο του x .

Παρατήρηση 7.1.12. Αν η f είναι Boolean τότε μπορούμε να εφαρμόσουμε το θεώρημα

για την $f - \mathbb{E}[f]$ και να χρησιμοποιήσουμε την ανισότητα Khintchine για να πάρουμε την ισοπεριμετρική ανισότητα (7.1.18) του Talagrand.

Το Θεώρημα 7.1.11 θα προκύψει ως συνέπεια της ακόλουθης ανισότητας «ευστάθειας θορύβου» για διανυσματικές συναρτήσεις.

Θεώρημα 7.1.13. Για κάθε $p \in [1, 2]$ και για κάθε $f : \{-1, 1\}^n \rightarrow X$ έχουμε ότι

$$(7.1.20) \quad \|f - e^{-t\Delta} f\|_p \lesssim C_q(X) q^{3/2} (1 - e^{-2t})^{\frac{1}{q}} \|Df\|_p, \quad t \geq 0.$$

Παρατήρηση 7.1.14. Αν η f είναι Boolean, τότε το Θεώρημα 7.1.13 μαζί με την απλή εκτίμηση

$$\|f - e^{-\Delta/d} f\|_p \gtrsim \left(\sum_{|S| \geq d} \widehat{f}(S)^2 \right)^{1/p}, \quad p \in [1, 2]$$

συνεπάγεται την (7.1.17).

7.1.5 Hamming μπάλες

Για λόγους πληρότητας, αποδεικνύουμε ότι η συνθήκη $p \geq 1/2$ είναι αναγκαία για τις ισοπεριμετρικές ανισότητες (7.1.1) και (7.1.14). Το αντιπαράδειγμα, και στις δύο περιπτώσεις, δίνεται από μπάλες της μετρικής Hamming. Έστω $a \in \{0, 1\}^n$ και $r > 0$. Συμβολίζουμε με $B(a, r)$ τη μπάλα Hamming με κέντρο a και ακτίνα r , ως προς την ℓ^1 -απόσταση στον $\{0, 1\}^n$. Σημειώνουμε ότι το πλήθος των στοιχείων της $B(a, r)$ είναι ίσο με $\sum_{i=0}^{r^*} \binom{n}{i}$, όπου $r^* := \min\{r, n\}$. Αν $r \in \mathbb{N}$, $r \leq n$, τότε η σφαίρα $S(a, r)$ έχει $\binom{n}{r}$ στοιχεία.

Ξεκινάμε με τον ισχυρισμό στην (7.1.4). Έστω n ένας θετικός άρτιος ακέραιος και $a \in \{0, 1\}^n$. Θεωρούμε το σύνολο $A = B(a, n/2)$. Τότε, $\text{supp}(h_A) = S(a, n/2)$. Εφόσον $h_A(x) = n/2$ για $x \in S(0, n/2)$, και $h_A(x) = 0$ αλλιώς, έχουμε ότι

$$\mathbb{E}[h_A^p] = (n/2)^p \mu_n(\text{supp}(h_A)) = (n/2)^p \mu_n(S(0, n/2)) = (n/2)^p \frac{\binom{n}{n/2}}{2^n}.$$

Από τον τύπο του Stirling παίρνουμε $\lim_{n \rightarrow \infty} \mathbb{E}[h_A^p] = 0$ για $p < 1/2$.

Για να δείξουμε ότι η συνθήκη $p \geq 1/2$ είναι αναγκαία για την (7.1.14) θεωρούμε έναν θετικό περιττό ακέραιο n και τυχόν $a \in \{0, 1\}^n$. Θεωρούμε τη μπάλα Hamming $A = B(a, (n-1)/2)$, που έχει μέτρο $\mu_n(A) = 2^{-n} \sum_{i=0}^{(n-1)/2} \binom{n}{i} = 1/2$, γράφουμε W για το σύνολο του A , και ορίζουμε $B := \{0, 1\}^d \setminus (A \cup W)$. Παρατηρήστε ότι $|\nabla(A, B)| = 0$. Τότε, η ανισότητα (7.1.14) ανάγεται στην

$$(7.1.21) \quad K|W|n^p \geq 2^{n-1}.$$

Εφόσον $|W| = \binom{n}{(n-1)/2}$, από τον τύπο του Stirling έπεται ότι η συνθήκη $p \geq 1/2$ είναι αναγκαία για να ισχύει η ανισότητα $\liminf_{n \rightarrow \infty} K \frac{|W|n^p}{2^{n-1}} \geq 1$.

7.2 Απόδειξεις των θεωρημάτων

Για την απόδειξη των Θεωρημάτων 7.1.1, 7.1.3 και 7.1.4 θα χρησιμοποιήσουμε ένα γνωστό επαγωγικό επιχείρημα το οποίο εμφανίζεται στα [48], [8] και σε βελτιωμένη μορφή στο [33, σελ. 4217-4219].

Λήμμα 7.2.1. Έστω $p \in [1/2, 1]$. Αν n $G : [0, 1] \rightarrow [0, \infty)$ ικανοποιεί τις $G(0) = G(1) = 0$, $G(1/2) \leq 1/2$ και

$$(7.2.1) \quad \max\{((y-x)^{1/p} + G(y)^{1/p})^p, y-x + (2^p-1)G(y)\} + G(x) \geq 2G\left(\frac{y+x}{2}\right)$$

για κάθε $0 \leq x \leq y \leq 1$, τότε

$$\mathbb{E}[h_A^p] \geq G(\mu_n(A))$$

για κάθε $n \geq 1$ και κάθε $A \subseteq \{0, 1\}^n$.

Παρατήρηση 7.2.2. Για κάθε $\lambda > 0$, η συνάρτηση $T(z) = (\lambda^{1/p} + z^{1/p})^p - z$ είναι φθίνουσα στο $[0, +\infty)$. Έπεται ότι

$$(2^p - 1)\lambda = (\lambda^{1/p} + \lambda^{1/p})^p - \lambda = T(\lambda) \leq T(z) = (\lambda^{1/p} + z^{1/p})^p - z$$

αν και μόνο αν $z \leq \lambda$. Συνεπώς,

$$\max\{((y-x)^{1/p} + G(y)^{1/p})^p, y-x + (2^p-1)G(y)\} = ((y-x)^{1/p} + G(y)^{1/p})^p$$

αν και μόνο αν $G(y) \geq y-x$.

Περιγραφή της απόδειξης. Δουλεύουμε με τον διακριτό κύβο $V = \{0, 1\}^n$. Η απόδειξη θα γίνει με επαγωγή ως προς n . Εύκολα ελέγχουμε ότι το λήμμα ισχύει για $n = 1$, οπότε υποθέτουμε ότι $n \geq 2$. Για δοθέν A , σταθεροποιούμε $i \in [n]$ και ορίζουμε $V_0 = \{v \in V : v_i = 0\}$, $V_1 = \{v \in V : v_i = 1\}$, $A_0 = A \cap V_0$ και

$$A_1 = (A \cap V_1)^i := \{v^i : v \in A, v_i = 1\} \subseteq V_0$$

όπου v^i είναι η n -άδα που διαφέρει από το v μόνο στην i -οστή συντεταγμένη.

Θεωρούμε το ομοιόμορφο μέτρο μ' στο V_0 . Θέτουμε $y = \mu'(A_0)$, $x = \mu'(A_1)$ και τότε $\mu(A) = (y+x)/2$. Από την επαγωγική υπόθεση, για $i = 0, 1$ έχουμε

$$\int h_{A_0}^p d\mu' \geq G(y) \quad \text{και} \quad \int h_{A_1}^p d\mu' \geq G(x).$$

Μπορούμε να υποθέσουμε ότι $y \geq x$. Σημειώνουμε ότι

$$h_A(v) = \begin{cases} h_{A_0}(v) + 1 & \text{αν } v \in A_0 \setminus A_1 \\ h_{A_0}(v) & \text{αν } v \in A_0 \cap A_1 \\ h_{A_1}(v^i) + 1 & \text{αν } v \in A_1 \setminus A_0 \\ h_{A_1}(v^i) & \text{αν } v \in A_0 \cap A_1. \end{cases}$$

Συνεπώς,

$$\begin{aligned} \int h_A^p d\mu &= \int_{A_0} h_A^p d\mu + \int_{(A_1)^i} h_A^p d\mu = \int_{A_0} h_A^p d\mu + \int_{A_1 \setminus A_0} (h_{A_1} + 1)^p d\mu + \int_{A_0 \cap A_1} h_{A_1}^p d\mu \\ &\geq \int_{A_0} h_A^p d\mu + \int_{A_1} h_{A_1}^p d\mu \geq \int_{A_0} h_A^p d\mu + \frac{1}{2} G(x). \end{aligned}$$

Θα θέλαμε λοιπόν να δείξουμε ότι

$$(7.2.2) \quad \int_{A_0} h_A^p d\mu \geq G\left(\frac{x+y}{2}\right) - \frac{1}{2} G(x).$$

Θέτουμε $Z = \text{supp}(h_{A_0}) \setminus A_1$ και $X = \text{supp}(h_{A_0}) \cap A_1$. Τότε,

$$(7.2.3) \quad 2 \int_{A_0} h_A^p d\mu = \int_Z (h_{A_0} + 1)^p d\mu' + \int_X h_{A_0}^p d\mu' + \int_{A_0 \setminus (A_1 \cup Z)} 1 d\mu'.$$

Ισχυρισμός 7.2.3. Μπορούμε να υποθέσουμε ότι $A_1 \subseteq A_0$.

Απόδειξη. Έστω ότι υπάρχει $v \in A_1 \setminus A_0$. Τότε, μπορούμε να βρούμε $w \in A_0 \setminus A_1$ διότι $\mu'(A_0) \geq \mu'(A_1)$. Ορίζουμε $B_1 = (A_1 \setminus \{v\}) \cup \{w\}$, $B = A_0 \cup (B_1)^i$ και $B_0 = B \cap V_0 = A_0$. Παρατηρήστε ότι $|A| = |B|$, $|A_0| = |B_0|$ και $|A_1| = |B_1|$. Επίσης,

$$\int_{B_0} h_B^p d\mu < \int_{A_0} h^p d\mu,$$

διότι, αν $Z_B = \text{supp}(h_{B_0}) \setminus B_1$ και $X_B = \text{supp}(h_{B_0}) \cap B_1$ τότε η θέση του w μεταβλαλλεται είτε από το Z στο X_B ή από το $A_0 \setminus (A_1 \cup Z)$ στο $(B_0 \cap B_1) \setminus X_B$, και σε κάθε περίπτωση η συνεισφορά του στο δεξιό μέλος της (7.2.3) μειώνεται. Συνεπώς, αν το A_1 δεν είναι υποσύνολο του A_0 τότε μπορούμε να το μετατρέψουμε σε ένα «χειρότερο» σύνολο. $\square$

Θέτουμε $\sigma = \int_Z h_{A_0}^p d\mu'$, $\gamma = \int_X h_{A_0}^p d\mu'$, $\lambda = \sigma + \gamma = \int h_{A_0}^p d\mu'$ και $z = \mu'(Z)$. Αφού $A_1 \subseteq A_0$ το δεξιό μέλος της (7.2.3) είναι ίσο με

$$\begin{aligned} \int_Z (h_{A_0} + 1)^p d\mu' + \gamma + \mu'(A_0 \setminus (A_1 \cup Z)) &\geq (\sigma^{1/p} + z^{1/p})^p + \gamma + (y - x - z) \\ &= ((\lambda - \gamma)^{1/p} + z^{1/p})^p + \gamma + (y - x - z) \\ &\geq (\lambda^{1/p} + z^{1/p})^p + (y - x - z), \end{aligned}$$

όπου η πρώτη ανισότητα ισχύει διότι, γενικά, αν $X \subseteq V$ και $\frac{1}{\mu(X)} \int_X f^p d\mu = T^p$ για κάποια $f \geq 0$ τότε

$$\frac{1}{\mu(X)} \int_X (f + 1)^p d\mu \geq (T + 1)^p,$$

ενώ η δεύτερη ανισότητα ισχύει διότι η συνάρτηση $\gamma \mapsto ((\lambda - \gamma)^{1/p} + z^{1/p})^p + \gamma$ είναι αύξουσα.

Για να ολοκληρώσουμε την απόδειξη αρκεί να ελέγξουμε ότι

$$(\lambda^{1/p} + z^{1/p})^p + (y - x - z) \geq 2G\left(\frac{x+y}{2}\right) - G(x),$$

όπου μπορούμε να υποθέσουμε ότι

$$\lambda = \int h_{A_0}^p d\mu' \geq G(y)$$

και

$$z \leq \min\{\lambda, y - x\}$$

(το δεύτερο φράγμα ισχύει διότι $Z \subseteq A_0 \setminus A_1$). Διακρίνουμε δύο περιπτώσεις:

Περίπτωση 1: $G(y) \leq y - x$.

Από την Παρατήρηση 7.2.2 έχουμε ότι $2G\left(\frac{y+x}{2}\right) \leq y-x+(2^p-1)G(y)+G(x)$. Χρησιμοποιώντας και την $G(y) \leq \lambda$ βλέπουμε ότι αρκεί να ελέγξουμε την

$$(2^p - 1)\lambda \leq (\lambda^{1/p} + z^{1/p})^p - z,$$

η οποία είναι ακριβώς η $T(\lambda) \leq T(z)$ και ισχύει διότι $z \leq \lambda$.

Περίπτωση 2: $G(y) \geq y - x$.

Από την Παρατήρηση 7.2.2 έχουμε ότι $2G\left(\frac{y+x}{2}\right) \leq ((y-x)^{1/p} + G(y)^{1/p})^p + G(x)$. Αφού $G(y) \leq \lambda$, αρκεί λοιπόν να ελέγξουμε ότι

$$((y-x)^{1/p} + \lambda^{1/p})^p \leq (\lambda^{1/p} + z^{1/p})^p + (y-x-z)$$

δηλαδή ότι $T(y-x) \leq T(z)$, η οποία ισχύει λόγω της $z \leq y-x$. □

7.2.1 Απόδειξη του Θεωρήματος 7.1.1: λογαριθμική συνάρτηση

Η «ανισότητα δύο σημείων» του επόμενου λήμματος θα παίξει βασικό ρόλο στις αποδείξεις μας.

Λήμμα 7.2.4. Για κάθε $p \in [\frac{1}{2}, 1]$ και κάθε $0 \leq x \leq y \leq \frac{1}{2}$ έχουμε ότι

$$(7.2.4) \quad ((y-x)^{1/p} + G(y)^{1/p})^p + G(x) \geq 2G\left(\frac{y+x}{2}\right),$$

όπου $G(x) = x(\log_2(1/x))^p$.

Απόδειξη. Αρχικά ξαναγράφουμε την ανισότητα (7.2.4) στη μορφή

$$(7.2.5) \quad (y^{1/p} \log_2(1/y) + (y-x)^{1/p})^p + x(\log_2(1/x))^p \geq (x+y) \left(\log_2 \left(\frac{2}{x+y} \right) \right)^p.$$

Θα χωρίσουμε την απόδειξη σε τρία βήματα.

Βήμα 1. Αν η ανισότητα ισχύει για $p = 1/2$, τότε ισχύει για κάθε $p \geq 1/2$. Πράγματι, ξαναγράφουμε την ανισότητα (7.2.5) στη μορφή

$$(7.2.6) \quad \left[\frac{y}{x+y} \left(\log_2(1/y) + \left(\frac{y-x}{y} \right)^{1/p} \right)^p + \frac{x}{x+y} (\log_2(1/x))^p \right]^{1/p} \geq \log_2 \left(\frac{2}{x+y} \right)$$

και στη συνέχεια εφαρμόζουμε την ανισότητα $\left(\frac{y-x}{y} \right)^{1/p} \geq \left(\frac{y-x}{y} \right)^2$ και το γεγονός ότι η συνάρτηση $p \mapsto (\mathbb{E}[|X|^p])^{1/p}$ είναι αύξουσα στο $[0, \infty)$ για κάθε τυχαία μεταβλητή X .

Σε ό,τι ακολουθεί υποθέτουμε ότι $p = \frac{1}{2}$ στην (7.2.5).

Βήμα 2. Αν η ανισότητα ισχύει για $y = \frac{1}{2}$ τότε ισχύει για κάθε $y \in [x, \frac{1}{2}]$. Πράγματι, θέτουμε $x = yt$ όπου $0 \leq t \leq 1$, και $a = \log_2(1/y) \in [1, \infty)$. Τότε, μπορούμε να ξαναγράψουμε την (7.2.6) ως εξής:

$$\left[\frac{1}{t+1} \left(a + (1-t)^{1/p} \right)^p + \frac{t}{t+1} (a + \log_2(1/t))^p \right]^{1/p} - a \geq \log_2 \left(\frac{2}{t+1} \right).$$

Παρατηρούμε ότι η $\varphi(a) = (\mathbb{E}[(a + |X|)^p])^{1/p} - a$ είναι αύξουσα στο $(0, \infty)$ για κάθε τυχαία μεταβλητή X . Πράγματι,

$$\varphi'(a) = (\mathbb{E}[(a + |X|)^p])^{\frac{1}{p}-1} \mathbb{E}[(a + |X|)^{p-1}] - 1 \geq 0$$

από την ανισότητα Hölder.

Αρκεί λοιπόν να επαληθεύσουμε την ανισότητα για $y = 1/2$.

Βήμα 3. Δείχνουμε ότι ισχύει η ανισότητα

$$(7.2.7) \quad \sqrt{1 + (1-t)^2} + t \sqrt{1 + \log_2(1/t)} \geq (t+1) \sqrt{1 + \log_2 \left(\frac{2}{t+1} \right)}, \quad t \in [0, 1].$$

Αν θέσουμε $1-t = s$, τότε η ανισότητα (7.2.7) παίρνει τη μορφή

$$(7.2.8) \quad \sqrt{1 + s^2} + (1-s) \sqrt{1 - \log_2(1-s)} \geq (2-s) \sqrt{1 + \log_2 \left(\frac{2}{2-s} \right)}.$$

Χρησιμοποιούμε δύο απλές ανισότητες:

$$(7.2.9) \quad -\log_2(1-s) \geq \frac{s}{\log(2)} + \frac{s^2}{2\log(2)}$$

και

$$(7.2.10) \quad \log_2 \left(\frac{2}{2-s} \right) \leq \frac{s}{2\log(2)} + \left(1 - \frac{1}{2\log(2)} \right) s^2.$$

Η πρώτη ανισότητα προκύπτει από το θεώρημα Taylor. Για τη δεύτερη ανισότητα θεωρούμε τη συνάρτηση $g(s) = \frac{s}{2\log(2)} + \left(1 - \frac{1}{2\log(2)}\right)s^2 - \log_2\left(\frac{2}{2-s}\right)$. Παρατηρούμε ότι $g(0) = g(1) = 0$. Επιπλέον, η

$$\frac{g'(s)}{s} = -\frac{(4\log(2) - 2)s + 5 - 8\log(2)}{2(2-s)\log(2)}$$

αλλάζει πρόσημο μόνο μία φορά από $+$ σε $-$ στο $(0, 1)$ (παρατηρήστε ότι $5 - 8\log(2) < 0$ και $3 - 4\log(2) > 0$). Συνεπώς, $g(s) \geq 0$ στο $[0, 1]$.

Αντικαθιστούμε τώρα τους λογαρίθμους στην (7.2.8) με τα αντίστοιχα κάτω και άνω φράγματα που δώσαμε στις (7.2.9) και (7.2.10). Υψώνοντας στο τετράγωνο βλέπουμε ότι αρκεί να δείξουμε την

$$\begin{aligned} & 1 + s^2 + 2(1-s)\sqrt{(1+s^2)\left(1 + \frac{s}{\log(2)} + \frac{s^2}{2\log(2)}\right)} + (1-s)^2\left(1 + \frac{s}{\log(2)} + \frac{s^2}{2\log(2)}\right) \\ & \geq (2-s)^2\left(1 + \frac{s}{2\log(2)} + \left(1 - \frac{1}{2\log(2)}\right)s^2\right). \end{aligned}$$

Απομονώνουμε την τετραγωνική ρίζα στο αριστερό μέλος, υψώνουμε πάλι τα δύο μέλη στο τετράγωνο, και έχουμε να δείξουμε ότι

$$\begin{aligned} 4(1-s)^2(1+s^2)\left(1 + \frac{s}{\log(2)} + \frac{s^2}{2\log(2)}\right) & \geq \left[-(1-s)^2\left(1 + \frac{s}{\log(2)} + \frac{s^2}{2\log(2)}\right) \right. \\ & \quad \left. + (2-s)^2\left(1 + \frac{s}{2\log(2)} + \left(1 - \frac{1}{2\log(2)}\right)s^2\right) - 1 - s^2\right]^2. \end{aligned}$$

Ανοίγοντας τις παρενθέσεις βλέπουμε ότι η διαφορά του αριστερού μέλους από το δεξιό μέλος απλοποιείται σε $-\frac{s^2(1-s)^2}{4\log^2(2)} \times \nu(s)$, όπου

$$\begin{aligned} \nu(s) &= 4(\log(2) - 1)^2 s^4 + 12(1 - \log(2))(2\log(2) - 1)s^3 - (17 - 18\log(2))(2\log(2) - 1)s^2 \\ &\quad - (12 - 24\log(2) + 16\log^2(2))s + 32\log^2(2) - 32\log(2) + 4. \end{aligned}$$

Ο παράγοντας $s^2(1-s)^2$ εμφανίζεται φυσιολογικά αφού για $s = 0$ και $s = 1$ έχουμε τις περιπτώσεις ισότητας στην (7.2.8) και στη διάρκεια των αναγωγών που κάναμε κρατήσαμε τις τιμές στο $s = 0$ και στο $s = 1$ αμετάβλητες.

Τα πρόσημα των συντελεστών στην $\nu(s)$, από τη μεγαλύτερη προς τη μικρότερη δύναμη, είναι $+, +, -, -, -$. Επομένως, από τον κανόνα προσήμων του Descartes, το πλήθος των θετικών ριζών του πολυωνύμου ν είναι το πολύ 1. Εφόσον οι τιμές $\nu(s)$ είναι θετικές στο άπειρο, και $\nu(1) = 32\log^2(2) - 32\log(2) + 1 < 0$, έπεται ότι $\nu(s) < 0$ στο $[0, 1]$, άρα $g(s) > 0$ στο $[0, 1]$. Αυτό αποδεικνύει την (7.2.7) και ολοκληρώνει την απόδειξη του Λήμματος 7.2.4. $\square$

Το επόμενο λήμμα θα χρησιμοποιηθεί για σύνολα με μέτρο μεγαλύτερο από $1/2$.

Λήμμα 7.2.5. Για κάθε $p \in [\frac{1}{2}, \log(2)]$ και για κάθε $0 \leq x \leq y \leq \frac{1}{2}$ έχουμε ότι

$$(7.2.11) \quad ((y-x)^{1/p} + G(x)^{1/p})^p + G(y) \geq 2G\left(\frac{y+x}{2}\right),$$

όπου $G(x) = x(\log_2(1/x))^p$.

Απόδειξη. Παρατηρούμε ότι η G είναι αύξουσα στο $[0, 1/2]$ διότι

$$G'(x) = (\log_2(1/x))^{p-1} \left(\log_2(1/x) - \frac{p}{\log(2)} \right) \geq 0$$

για $p \in [1/2, \log(2)]$. Για κάθε $0 \leq x \leq y \leq 1/2$ έχουμε

$$(7.2.12) \quad ((y-x)^{1/p} + G(x)^{1/p})^p - ((y-x)^{1/p} + G(y)^{1/p})^p + G(y) - G(x) \geq 0.$$

Πράγματι, για κάθε $p \in (0, 1]$ η απεικόνιση $\psi(t) = ((y-x)^{1/p} + t^{1/p})^p - t$ είναι φθίνουσα στο $[0, \infty)$. Χρησιμοποιώντας το γεγονός ότι $G(x) \leq G(y)$ βλέπουμε ότι $\psi(G(x)) \geq \psi(G(y))$, δηλαδή ισχύει η (7.2.12). Τέλος, το αριστερό μέλος της (7.2.11) μείον το αριστερό μέλος της (7.2.12) ισούται με το αριστερό μέλος της (7.2.4) και έχουμε το λήμμα. $\square$

Το επόμενο λήμμα είναι το βασικό τεχνικό βήμα στο [33]. Το αποτέλεσμα αυτό περιέχεται επίσης στα Λήμματα 7.2.7 και 7.2.8 για $p = \log_2(3/2)$.

Λήμμα 7.2.6. Η συνάρτηση $G(x) = 2x(1-x)$ ικανοποιεί την (7.2.1) για $p = \log_2(3/2)$.

Απόδειξη του Θεωρήματος 7.1.1. Θέτουμε $p = p_0 = \log_2(3/2)$ στο Λήμμα 7.2.1, και θεωρούμε την

$$(7.2.13) \quad G(t) := \begin{cases} t(\log_2(1/t))^{p_0} & t \in [0, 1/4], \\ 2t(1-t) & t \in [1/4, 3/4], \\ (1-t)(\log_2(1/(1-t)))^{p_0} & t \in (3/4, 1]. \end{cases}$$

Παρατηρούμε ότι

$$G(t) = \max\{t^*(\log_2(1/t^*))^{p_0}, 2t(1-t)\} =: \max\{G_1(t), G_2(t)\}, \quad t \in [0, 1],$$

όπου $t^* = \min\{t, 1-t\}$. Μπορούμε να γράψουμε το αριστερό μέλος της ανισότητας (7.2.1) ως $\Phi(G(x), G(y), (y-x))$, όπου

$$\Phi(u, v, w) := \max\{(w^{1/p_0} + v^{1/p_0})^{p_0}, w + (2^{p_0} - 1)v\} + u.$$

Η απεικόνιση $(u, v, w) \in [0, \infty)^3 \mapsto \Phi(u, v, w)$ είναι αύξουσα ως προς u και v . Στόχος μας είναι να δείξουμε ότι η

$$(7.2.14) \quad \Phi(G(x), G(y), (y-x)) \geq 2G\left(\frac{x+y}{2}\right)$$

ισχύει για κάθε $0 \leq x \leq y \leq 1$ για τη συνάρτηση G που ορίσαμε στην (7.2.13). Θα χρειαστεί να διακρίνουμε έξι περιπτώσεις.

- (i) Αν $x, y \in [0, 1/4]$ τότε $(x + y)/2 \in [0, 1/4]$, άρα η G στην (7.2.14) συμπίπτει με την G_1 και τότε η (7.2.14) έπεται από το Λήμμα 7.2.4.
- (ii) Αν $x, y \in [3/4, 1]$ τότε $(x + y)/2 \in [3/4, 1]$, άρα η G στην (7.2.14) συμπίπτει με την G_1 . Θέτουμε $x = 1 - b$, $y = 1 - a$. Τότε $0 \leq a \leq b \leq 1/4$, και έχουμε

$$\begin{aligned}\Phi(G(x), G(y), (y - x)) &= \Phi(G_1(x), G_1(y), (y - x)) = \Phi(G_1(b), G_1(a), (b - a)) \\ &\geq 2G_1((a + b)/2) = 2G((x + y)/2),\end{aligned}$$

όπου η ανισότητα προκύπτει από το Λήμμα 7.2.5.

- (iii) Αν $x, y \in [1/4, 3/4]$ τότε $(x + y)/2 \in [1/4, 3/4]$, άρα η G στην (7.2.14) συμπίπτει με την G_2 και τότε η (7.2.14) έπεται από το Λήμμα 7.2.6.
- (iv) Αν $x \in [0, 1/4]$, $y \in [3/4, 1]$ τότε $(x + y)/2 \in [1/4, 3/4]$. Συνεπώς,

$$\begin{aligned}\Phi(G(x), G(y), (y - x)) &\geq \Phi(G_2(x), G_2(y), (y - x)) \\ &\geq 2G_2((x + y)/2) = 2G((x + y)/2),\end{aligned}$$

όπου η δεύτερη ανισότητα προκύπτει από το Λήμμα 7.2.6.

- (v) Αν $x \in [0, 1/4]$, $y \in [1/4, 3/4]$ τότε $(x + y)/2 \in [1/8, 1/2]$. Διακρίνουμε δύο περιπτώσεις:

- (v1) Αν $(x + y)/2 \in [0, 1/4]$, τότε $y \leq 1/2$ και έχουμε

$$\begin{aligned}\Phi(G(x), G(y), (y - x)) &\geq \Phi(G_1(x), G_1(y), (y - x)) \\ &\geq 2G_1((x + y)/2) = 2G((x + y)/2),\end{aligned}$$

όπου στη δεύτερη ανισότητα χρησιμοποιήσαμε το Λήμμα 7.2.4.

- (v2) Αν $(x + y)/2 \in [1/4, 3/4]$ τότε

$$\begin{aligned}\Phi(G(x), G(y), (y - x)) &\geq \Phi(G_2(x), G_2(y), (y - x)) \\ &\geq 2G_2((x + y)/2) = 2G((x + y)/2),\end{aligned}$$

όπου στη δεύτερη ανισότητα χρησιμοποιήσαμε το Λήμμα 7.2.6.

- (vi) Αν $x \in [1/4, 3/4]$, $y \in [3/4, 1]$ τότε $(x + y)/2 \in [1/2, 7/8]$. Διακρίνουμε δύο περιπτώσεις:

- (vi1) Αν $(x + y)/2 \in [1/4, 3/4]$ τότε χρησιμοποιούμε το φράγμα $G(y) \geq G_2(y)$ και προχωράμε με το Λήμμα 7.2.6, όμοια με την περίπτωση (iv), ή την περίπτωση (v2).

(vi2) Αν $(x + y)/2 \in [3/4, 1]$, τότε $x \geq 1/2$. Όπως στην περίπτωση (ii), θέτουμε $x = 1 - b$, $y = 1 - a$. Σημειώνουμε ότι $0 \leq a \leq 1/4 \leq b \leq 1/2$ και έχουμε

$$\begin{aligned}\Phi(G(x), G(y), (y - x)) &\geq \Phi(G_1(x), G_1(y), (y - x)) = \Phi(G_1(b), G_1(a), (b - a)) \\ &\geq 2G_1((a + b)/2) = 2G((x + y)/2),\end{aligned}$$

όπου στη δεύτερη ανισότητα χρησιμοποιήσαμε το Λήμμα 7.2.5.

Έτσι, έχουμε ελέγξει ότι η (7.2.14) ισχύει για τη συνάρτηση G που ορίστηκε στην (7.2.13). Συνδυάζοντας αυτή την ανισότητα με το Λήμμα 7.2.1 παίρνουμε

$$\mathbb{E} \left[h_A^{p_0} \right] \geq G(\mu_n(A)) \geq \mu_n(A)^* (\log_2(1/\mu_n(A)^*))^{p_0}.$$

Αυτό ολοκληρώνει την απόδειξη της κύριας ανισότητας (7.1.9).

Για να ελέγξουμε την (7.1.10), εφαρμόζουμε την ανισότητα Hölder στην (7.1.9) για κάθε $q \in [0, 1]$ και κάθε $A \subseteq \{0, 1\}^n$ με $\mu_n(A) \leq 1/2$. Αυτό μας δίνει

$$\left(\mathbb{E} \left[h_A^{p_0/q} \right] \right)^q \mu_n(A)^{1-q} \geq \left(\mathbb{E} \left[h_A^{p_0/q} \right] \right)^q \mu_n(\partial A)^{1-q} \geq \mathbb{E} \left[h_A^{p_0} \right] \geq \mu_n(A) (\log_2(1/\mu_n(A)))^{p_0},$$

όπου η πρώτη ανισότητα χρησιμοποιεί το τετριμμένο φράγμα $\mu_n(\partial A) \leq \mu_n(A)$. Έτσι βλέπουμε ότι

$$\left(\mathbb{E} \left[h_A^{p_0/q} \right] \right)^q \mu_n(A)^{1-q} \geq \mu_n(A) (\log_2(1/\mu_n(A)))^{p_0}.$$

Διαιρώντας τα δύο μέλη της ανισότητας με $\mu_n(A)^{1-q}$ και υψώνοντας στη δύναμη $1/q$ παίρνουμε την (7.1.10). Αυτό ολοκληρώνει την απόδειξη του Θεωρήματος 7.1.1. $\square$

7.2.2 Απόδειξη του Θεωρήματος 7.1.3: τετραγωνικό πολυώνυμο

Η απόδειξη του Θεωρήματος 7.1.3 προκύπτει αν συνδυάσουμε τα επόμενα Λήμματα 7.2.7 και 7.2.8 με το Λήμμα 7.2.1.

Λήμμα 7.2.7. Έστω $p \in [1/2, \log_2(3/2)]$. Αν $C_p = 2\sqrt{2^{p+1}-2}$ και $G(x) = C_p x(1-x)$, τότε η ανισότητα

$$(7.2.15) \quad y - x + (2^p - 1)G(y) + G(x) \geq 2G\left(\frac{y+x}{2}\right)$$

ισχύει για κάθε $0 \leq y \leq 1$ και $0 \leq x \leq y - G(y)$.

Απόδειξη. Έστω $f(x, y)$ το αριστερό μέλος της (7.2.15) μείον το δεξιό μέλος της (7.2.15), δηλαδή,

$$(7.2.16) \quad f(x, y) := C_p(2^p - 2)y(1-y) - \frac{1}{2}(x-y)(C_p(x-y) + 2).$$

Η απεικόνιση $x \mapsto f(x, y)$ είναι κοίλη, άρα για να δείξουμε ότι $f(x, y) \geq 0$ για $y \in [0, 1]$ και $x \in [0, y - G(y)]$, αρκεί να ελέγξουμε τις εξής δύο ανισότητες: (i) $f(y - G(y), y) \geq 0$ και (ii)

$f(0, y) \geq 0$. Για να επαληθεύσουμε την (i) βλέπουμε πρώτα, με πράξεις, ότι

$$f(y - G(y), y) = \frac{C_p}{2} y(1 - y)(2^{p+1} - 2 - C_p^2 y(1 - y))$$

και παρατηρούμε ότι $2^{p+1} - 2 - C_p^2 y(1 - y) \geq 2^{p+1} - 2 - \frac{C_p^2}{4} = 0$ για κάθε $y \in [0, 1]$. Για να επαληθεύσουμε την (ii) βλέπουμε πρώτα ότι

$$2f(0, y)/y = yC_p(3 - 2^{p+1}) + 2^{p+1}C_p - 4C_p + 2.$$

Παρατηρούμε ότι $3 - 2^{p+1} \geq 0$ για κάθε $p \in [1/2, \log_2(3/2)]$. Επιπλέον, από την $y - G(y) \geq 0$ έπεται ότι $\frac{C_p^{-1}}{C_p} \leq y$. Η τιμή της $2f(0, y)/y$ στο σημείο $y = \frac{C_p^{-1}}{C_p}$ ισούται με $(\sqrt{2^{p+1} - 2} - 1)^2 \geq 0$. $\square$

Λήμμα 7.2.8. Έστω $p \in [1/2, \log_2(3/2)]$. Αν $C_p = 2\sqrt{2^{p+1} - 2}$ και $G(x) = C_p x(1 - x)$, τότε η ανισότητα

$$(7.2.17) \quad ((y - x)^{1/p} + G(y)^{1/p})^p + G(x) \geq 2G\left(\frac{y + x}{2}\right)$$

ισχύει για κάθε $0 \leq x \leq y \leq 1$ με $x \in [y - G(y), y]$.

Απόδειξη. Έστω $g(x, y)$ το αριστερό μέλος της (7.2.17) μείον το δεξιό μέλος της (7.2.17), δηλαδή,

$$g(x, y) := [(C_p y(1 - y))^{1/p} + (y - x)^{1/p}]^p - C_p y(1 - y) - \frac{C_p}{2}(y - x)^2.$$

Παρατηρούμε ότι $g(y, y) = 0$ και, από την απόδειξη του Λήμματος 7.2.7 έχουμε $g(y - G(y), y) = f(y - G(y), y) \geq 0$ στο $[0, 1]$, όπου η f είναι όπως στην (7.2.16). Τότε, για σταθερό $y \in [0, 1]$, για να αποδείξουμε ότι $g(x, y) \geq 0$ για κάθε $x \in [y - G(y), y]$, αρκεί να δείξουμε τους παρακάτω δύο ισχυρισμούς:

(i) η απεικόνιση $x \mapsto g(x, y)$ έχει το πολύ ένα κρίσιμο σημείο στο $(y - G(y), y)$.

(ii) υπάρχει $\varepsilon > 0$ τέτοιο ώστε $g(x, y) > 0$ για κάθε $x \in (y - \varepsilon, y)$.

Αποδεικνύουμε πρώτα το (i). Έχουμε ότι

$$(7.2.18) \quad \partial_x g(x, y) = -(y - x)^{\frac{1}{p}-1} \left[(C_p y(1 - y))^{\frac{1}{p}} + (y - x)^{\frac{1}{p}} \right]^{p-1} + C_p(y - x).$$

Θέτουμε $A := y - x$ και $G := C_p y(1 - y)$. Παρατηρούμε ότι $0 \leq A \leq G$, και το G δεν εξαρτάται από το x . Έχουμε

$$\partial_x g(x, y) = -A^{\frac{1}{p}-1} [G^{\frac{1}{p}} + A^{\frac{1}{p}}]^{p-1} + C_p A,$$

άρα η $\partial_x g(x, y) = 0$ είναι ισοδύναμη με την

$$(7.2.19) \quad G^{\frac{1}{p}} + A^{\frac{1}{p}} = C_p^{\frac{1}{p-1}} A^{\frac{2p-1}{p(p-1)}}.$$

Για $p \geq 1/2$, το αριστερό μέλος αυξάνει με το A , ενώ το δεξιό μέλος φθίνει με το A . Συνεπώς, η (7.2.19) μπορεί να ισχύει σε ένα το πολύ σημείο, το οποίο αποδεικνύει το (i).

Ο ισχυρισμός (ii) προκύπτει αν παρατηρήσουμε ότι $\lim_{x \rightarrow y} \frac{\partial_x g(x,y)}{(y-x)^{\frac{1}{p}-1}} < 0$ και $g(y,y) = 0$. $\square$

7.2.3 Απόδειξη του Θεωρήματος 7.1.4: κυβικό πολυώνυμο

Σε αυτήν τη ενότητα θα ασχοληθούμε με το πολυώνυμο

$$(7.2.20) \quad G(x) = 8x(1-x) \left[\left(1 - \frac{2^{q+1}}{3}\right)x + \left(\frac{2^q}{3} - \frac{1}{4}\right) \right].$$

Παρατηρούμε ότι η G έχει επιλεγεί έτσι ώστε να ικανοποιεί τις $G(0) = G(1) = 0$, $G(1/2) = 1/2$ και $G(1/4) = 2^q/4$. Είναι βολικό να αναπτύξουμε την G ως

$$(7.2.21) \quad G(x) = 8\left(\frac{2^q}{3} - \frac{1}{4}\right)x - 8\left(2^q - \frac{5}{4}\right)x^2 - 8\left(1 - \frac{2^{q+1}}{3}\right)x^3.$$

Το Θεώρημα 7.1.4 προκύπτει αν συνδυάσουμε το Λήμμα 7.2.1 με τα επόμενα δύο Λήμματα 7.2.9 και 7.2.10.

Λήμμα 7.2.9. Έστω $q = 1/2$ και $p = 0.53$. Η ανισότητα

$$y - x + (2^p - 1)G(y) + G(x) \geq 2G\left(\frac{y+x}{2}\right)$$

ισχύει για κάθε $0 \leq x \leq y - G(y)$, $0 \leq y \leq 1$.

Απόδειξη. Αρχικά παρατηρούμε ότι αν $G(y) \leq y$ τότε $y \geq 1/2$. Πράγματι, ένας υπολογισμός δείχνει ότι

$$(7.2.22) \quad y - G(y) = \frac{8(3 - 2^{q+1})}{3}y \left(y - \frac{1}{2}\right) \left(y - \frac{2^q - 9/8}{2^q - 3/2}\right).$$

Τώρα, ο ισχυρισμός προκύπτει αν παρατηρήσουμε ότι $\frac{2^q - 9/8}{2^q - 3/2} < 0$ για $q = 1/2 \leq \log_2(3/2)$.

Θέτουμε

$$(7.2.23) \quad f(x,y) := (2^p - 1)G(y) + (y - x) - 2G\left(\frac{x+y}{2}\right) + G(x).$$

Για να επαληθεύσουμε το Λήμμα 7.2.9 αρκεί να δείξουμε την ανισότητα $f(x,y) \geq 0$ για κάθε $1/2 \leq y \leq 1$ και κάθε $0 \leq x \leq y - G(y)$. Η ανισότητα έπεται από τους παρακάτω τρεις ισχυρισμούς:

(i) $f(0,y) \geq 0$ για κάθε $y \in [1/2, 1]$.

(ii) Η $x \mapsto f(x,y)$ είναι κοίλη στο $[0, y - G(y)]$ για κάθε σταθερό $y \in [1/2, 1]$.

(iii) $f(y - G(y), y) \geq 0$ για κάθε $y \in [1/2, 1]$.

Για να επαληθεύσουμε τον ισχυρισμό (i), παρατηρούμε ότι η $\varphi(y) := f(0, y)/y$ είναι κοίλη παραβολή, διότι

$$3\varphi''(y) = (20 - 2^{p+4})(3 - 2^{q+1}) < 0.$$

Από την άλλη πλευρά, $\varphi(1/2) = 2^p - 2^q > 0$ και $\varphi(1) = 0$. Συνεπώς, $\varphi(y) \geq 0$ στο $[1/2, 1]$, και έπεται το (i).

Για να επαληθεύσουμε τον ισχυρισμό (ii) παρατηρούμε ότι

$$\partial_{xx}^2 f(x, y) = -2(3 - 2^{q+1})(6x - 2y + 2^{q+1} + 1) \leq 0 \quad \text{για κάθε } x, y \in [0, 1].$$

Τέλος επαληθεύουμε τον ισχυρισμό (iii). Παρατηρούμε ότι η $f(y - G(y), y) = G(y)H(y)$, όπου

$$\begin{aligned} H(y) = & \left(11 - \frac{70}{9} \cdot 2^q\right) 2^7 y^6 + (2^{q+7} - 181) \frac{2^6}{3} y^5 + \left(1453 - \frac{3082}{3} \cdot 2^q\right) \frac{2^3}{3} y^4 \\ & - \left(\frac{331}{3} - 78 \cdot 2^q\right) 2^4 y^3 + \left(563 - \frac{1192}{3} \cdot 2^q\right) \frac{2}{3} y^2 - (47 - 2^{q+5}) \frac{2}{3} y \\ & + 2^p - 1 \end{aligned}$$

είναι πολυώνυμο βαθμού 6. Έχουμε $H'''(y) \geq H'''(1/2) = (5\sqrt{2} - 7)2^7 \geq 0$ στο $[1/2, 1]$ διότι τα πρόσημα όλων των συντελεστών του πολυωνύμου $H'''(y)$ είναι θετικά εκτός από αυτό του ελεύθερου συντελεστή. Επίσης $H''(1/2) = (147 - 100\sqrt{2}) \cdot 4/9 > 0$. Συνεπώς, η $H(y)$ είναι κυρτή στο $[1/2, 1]$ και έπεται ότι $H(y) \geq L(y) = H(y_0) + H'(y_0)(y - y_0)$, όπου $y_0 = 0.631$. Εφόσον $H'(y_0) = 0.00036 \dots > 0$ και $L(1/2) = 0.000067 \dots > 0$ συμπεραίνουμε ότι $H(y) \geq 0$ στο $[1/2, 1]$. $\square$

Λήμμα 7.2.10. Έστω $q = 0.5$ και $p = 0.53$. Ισχύει ότι

$$((y - x)^{1/p} + G(y)^{1/p})^p + G(x) \geq 2G\left(\frac{y + x}{2}\right)$$

για κάθε $0 \leq x, y \leq 1$ που ικανοποιούν την $x \in [\max\{y - G(y), 0\}, y]$.

Απόδειξη. Ορίζουμε

$$g(x, y) := \left(G(y)^{\frac{1}{p}} + (y - x)^{\frac{1}{p}}\right)^p - 2G\left(\frac{x + y}{2}\right) + G(x).$$

Σκοπός μας είναι να δείξουμε ότι $g(x, y) \geq 0$ για κάθε $0 \leq y \leq 1$ και κάθε $\max\{y - G(y), 0\} \leq x \leq y$. Είναι φανερό ότι $g(y, y) = 0$. Στην απόδειξη του Λήμματος 7.2.9 είδαμε ότι $\max\{y - G(y), 0\} = y - G(y)$ αν και μόνο αν $y \in [1/2, 1]$. Συνεπώς, παρατηρούμε ότι για $y \in [1/2, 1]$ έχουμε $g(y - G(y), y) = f(y - G(y), y) \geq 0$ επειδή το πολυώνυμο $H(y)$ που ορίστηκε στο Λήμμα 7.2.9 είναι θετικό στο $[1/2, 1]$. Εδώ, η f είναι όπως στην (7.2.23). Αν $y \in [0, 1/2]$, τότε

$$g(0, y) = (G(y)^{1/p} + y^{1/p})^p - 2G(y/2) \geq (G(y)^2 + y^2)^{1/2} - 2G(y/2).$$

Στη συνέχεια, έχουμε

$$G(y)^2 + y^2 - (2G(y/2))^2 = \frac{17 - 12\sqrt{2}}{3} \cdot y^2(1 - 2y)(1 - y)(10y^2 + y(29 + 28\sqrt{2}) + 51 + 36\sqrt{2}) \geq 0$$

για κάθε $y \in [0, 1/2]$. Έτσι συμπεραίνουμε ότι $g(\max\{y - G(y), 0\}, y) \geq 0$ για κάθε $y \in [0, 1]$.

Έτσι, για την απόδειξη του Λήμματος 7.2.10 αρκεί να εξηγήσουμε τους παρακάτω δύο ισχυρισμούς για κάθε σταθερό $y \in [0, 1]$:

(i) Η απεικόνιση $x \mapsto g(x, y)$ έχει το πολύ ένα κρίσιμο σημείο στο $(0, y)$.

(ii) Υπάρχει $\varepsilon > 0$ τέτοιο ώστε $\partial_x g(x, y) < 0$ για κάθε $x \in (y - \varepsilon, y)$.

Για τον ισχυρισμό (ii), άμεσος υπολογισμός δείχνει ότι

$$\partial_x g(x, y) = -\left(G(y)^{\frac{1}{p}} + (y - x)^{\frac{1}{p}}\right)^{p-1} (y - x)^{\frac{1}{p}-1} - G'\left(\frac{x+y}{2}\right) + G'(x).$$

Συνεπώς, αφού $\frac{1}{p} - 1 < 1$ έπεται ότι

$$\lim_{x \rightarrow y} \frac{\partial_x g(x, y)}{(y - x)^{\frac{1}{p}-1}} > 0.$$

Για να επαληθεύσουμε τον ισχυρισμό (i), παρατηρούμε ότι

$$(y - x)^{1-\frac{1}{p}} \partial_x g(x, y) = -\left(G(y)^{\frac{1}{p}} + (y - x)^{\frac{1}{p}}\right)^{p-1} + 2(3 - 2^{3/2})(y - x)^{2-\frac{1}{p}}(3x + y + 2^{3/2} + 1).$$

Αρκεί να δείξουμε ότι για κάθε $y \in [0, 1]$ η απεικόνιση

$$\psi(x) := -(G(y)^{\frac{1}{p}} + (y - x)^{\frac{1}{p}}) + [2(3 - 2^{3/2})(y - x)^{2-\frac{1}{p}}(3x + y + 2^{3/2} + 1)]^{\frac{1}{p-1}}$$

ικανοποιεί την $\psi' > 0$ στο $(0, y)$. Έχουμε

$$(7.2.24) \quad \frac{\psi'(x)(y - x)^{-\frac{2p-1}{p(p-1)}+1}}{(2(3 - 2\sqrt{2})(3x + y + 2^{3/2} + 1))^{\frac{1}{p-1}}} = \frac{((y - x)(3x + y + 2^{3/2} + 1))^{\frac{1}{1-p}}}{p(2(3 - 2\sqrt{2}))^{\frac{1}{p-1}}} + \frac{2p - 1}{p(1 - p)} - \frac{3(y - x)}{(1 - p)(3x + y + 2^{3/2} + 1)}.$$

Θέτουμε $x = y - t$ στην (7.2.24) όπου $y \in [t, 1]$ και $t \in (0, 1)$. Τότε το δεξιό μέλος της (7.2.24) είναι αύξουσα συνάρτηση του y για κάθε σταθερό t , και για $y = t$ παίρνει τη μορφή

$$\frac{(t(t + 2^{3/2} + 1))^{\frac{1}{1-p}}}{p(2(3 - 2\sqrt{2}))^{\frac{1}{p-1}}} + \frac{2p - 1}{p(1 - p)} - \frac{3t}{(1 - p)(t + 2^{3/2} + 1)} =: \varphi(t).$$

Η $\varphi(t)$ είναι κυρτή $[0, 1]$ ως άθροισμα δύο κυρτών συναρτήσεων. Συνεπώς, $\varphi(t) \geq L(t) := \varphi(0.22) + \varphi'(0.22)(t - 0.22)$. Αφού $\varphi'(0.22) = 0.053 \dots > 0$ και $L(0) = 0.033 \dots > 0$ έχουμε τον ισχυρισμό του λήμματος. $\square$

Ολοκληρώνουμε αυτήν την ενότητα εξηγώντας γιατί το Θεώρημα 7.1.4 (ή το Πόρισμα 7.1.5) συνεπάγεται το Πόρισμα 7.1.6. Η απόδειξη δίνεται ήδη στο [33, Πόρισμα 3.2], αλλά την συμπεριλαμβάνουμε για λόγους πληρότητας.

Απόδειξη του Πορίσματος 7.1.6. Έστω (A, B, W) μια διαμέριση του $\{0, 1\}^n$ με $\mu_n(A) = 1/2$. Υποθέτουμε ότι $\mathbb{E}[h_A^p] \geq P(\mu_n(A))$ για κάθε τέτοιο A . Τότε, έχουμε

$$P(\mu_n(A)) \leq \mathbb{E}[h_{B \cup W}^p] = \mathbb{E}[(h_{B \cup W}^p \mathbb{1}_B)] + \mathbb{E}[(h_{B \cup W}^p \mathbb{1}_W)] \leq \frac{|\nabla(A, B)|}{2^n} + n^p \mu_n(W),$$

όπου στην τελευταία ανισότητα χρησιμοποιήσαμε το φράγμα $h_{B \cup W} \leq n$. Έτσι προκύπτει το ζητούμενο αν $P(1/2) = 1/2$, το οποίο ισχύει για τη συνάρτηση στο δεξιό μέλος της (7.1.13). $\square$

7.2.4 Απόδειξη των Θεωρημάτων 7.1.11 και 7.1.13

Για κάθε $t \geq 0$ θεωρούμε ένα τυχαίο διάνυσμα $\xi(t) = (x_1 \xi_1(t), \dots, x_n \xi_n(t)) \in \{-1, 1\}^n$, όπου οι $\xi_i(t)$ είναι ανεξάρτητες ισοκατανεμημένες τυχαίες μεταβλητές με

$$\mathbb{P}(\xi_j(t) = \pm 1) = \frac{1 \pm e^{-t}}{2}$$

και ορίζουμε

$$\delta_j(t) = \frac{\xi_j(t) - \mathbb{E}\xi_j(t)}{\sqrt{\text{Var}(\xi_j(t))}} = \frac{\xi_j(t) - e^{-t}}{\sqrt{1 - e^{-2t}}}.$$

Υπενθυμίζουμε ότι

$$D_j f(x) = \frac{f(x_1, \dots, x_j, \dots, x_n) - f(x_1, \dots, -x_j, \dots, x_n)}{2}$$

για κάθε $f : \{-1, 1\}^n \rightarrow \mathbb{R}$. Θεωρούμε τον τελεστή $\Delta = \sum_{j=1}^n D_j$, και την ημιομάδα της θερμοότητας $e^{-t\Delta}$, $t > 0$. Θα χρειαστούμε το ακόλουθο λήμμα από το [31].

Λήμμα 7.2.11. Έστω $f : \{-1, 1\}^n \rightarrow \mathbb{R}$. Για κάθε $t > 0$ ισχύει ότι

$$(7.2.25) \quad -\frac{d}{dt} e^{-t\Delta} f(x) = \frac{1}{\sqrt{e^{2t} - 1}} \mathbb{E}_\xi \left[\sum_{j=1}^n \delta_j(t) D_j f(\xi(t)x) \right].$$

Η ταυτότητα αυτή επαληθεύεται εύκολα για τις συναρτήσεις Walsh και λόγω γραμμικότητας βλέπουμε ότι ισχύει για κάθε συνάρτηση f . Από την (7.2.25) και από το θεμελιώδες

θεώρημα του απειροστικού λογισμού παίρνουμε

$$\begin{aligned} \|f - e^{-t\Delta} f\|_p &\lesssim \int_0^t \left(\mathbb{E} \left[\left\| \sum_{j=1}^n \delta_j(s) D_j f(x) \right\|^p \right] \right)^{1/p} \frac{ds}{\sqrt{e^{2s} - 1}} \\ &\lesssim C_q(X) \sqrt{q} \|Df\|_p \int_0^t (1 - e^{-2s})^{\frac{1}{q}-1} ds \\ &\lesssim C_q(X) q^{3/2} (1 - e^{-2t})^{\frac{1}{q}} \|Df\|_p, \end{aligned}$$

όπου χρησιμοποιήσαμε επίσης το [31, Θεώρημα 4.1] στη δεύτερη ανισότητα. Αυτό ολοκληρώνει την απόδειξη του Θεωρήματος 7.1.13.

Για το Θεώρημα 7.1.11 παρατηρούμε ότι λόγω της $\|e^{-t\Delta} f\|_\infty \leq \|f\|_\infty \leq 1$, η παραπάνω ανισότητα και η ανισότητα Minkowski συνεπάγονται ότι

$$(7.2.26) \quad \|f\|_2^{2/p} - \|e^{-t\Delta} f\|_2^{2/p} \lesssim C_q(X) q^{3/2} (1 - e^{-2t})^{\frac{1}{q}} \|Df\|_p.$$

Θέτουμε $\varepsilon = 1 - e^{-2t}$. Από την ανισότητα υπερσυσταλτότητας και την ανισότητα Hölder έχουμε ότι

$$\|e^{-t\Delta} f\|_2 \leq \|f\|_2 \left(\frac{\|f\|_1}{\|f\|_2} \right)^{\frac{\varepsilon}{2-\varepsilon}}.$$

Συνεπώς,

$$(7.2.27) \quad \frac{1 - \left(\frac{\|f\|_1}{\|f\|_2} \right)^{\frac{2\varepsilon}{2-\varepsilon}}}{\varepsilon^{1/q}} \lesssim C_q(X) q^{3/2} \frac{\|Df\|_p}{\|f\|_2^{2/p}}.$$

Αν $\frac{\|f\|_2}{\|f\|_1} < 10$, τότε το θεώρημα προκύπτει αν πάρουμε το όριο καθώς $t \rightarrow \infty$ στην (7.2.26). Διαφορετικά, επιλέγοντας $\varepsilon = \frac{1}{\log \frac{e\|f\|_2}{\|f\|_1}}$ στην (7.2.27) παίρνουμε το ζητούμενο.

CHAPTER 8

Summary

8.1 Boolean functions

Consider the discrete cube $\{-1, 1\}^n$ equipped with the uniform probability measure

$$\left(\frac{1}{2}\delta_1 + \frac{1}{2}\delta_{-1}\right)^{\otimes n}.$$

Boolean function is any function $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$, i.e. any function that maps vectors with n digits to a digit. Boolean functions are the main mathematical object that we shall study in this thesis.

Although their definition is simple, Boolean functions can represent many different things from various branches of mathematics and it should not come as a surprise that they appear very often. We list a number of examples that demonstrate their applicability.

Social Choice: Suppose that we have n citizens of England and that the digit $x_i \in \{-1, 1\}$ represents the opinion of the i -th citizen on some political matter, for example, if he agrees on Brexit. The vector $x = (x_1, \dots, x_n) \in \{-1, 1\}^n$ represents the opinion of this crowd of n citizens, and every Boolean function $f(x)$ represents some way of making a decision about the way the society should act taking into account the opinion of these n citizens. Obviously, there are many different ways of making a decision, and this can be seen from the variety of different voting systems in various countries in the world. Probably the most natural way is through the *majority function*, which outputs 1 if $\sum_{i=1}^n x_i > 0$ and -1 if $\sum_{i=1}^n x_i < 0$. In other words, through this function, the decision is made in accordance with the opinion of the majority of the n citizens. A different way, less democratic, is through the *dictatorship function* which takes into account the opinion of just one of the citizens, say the k -th citizen, and outputs the value $f(x) = x_k$ that he has chosen.

Algorithms and Graph Theory: A large part of Computer Science and Algorithms deals with concrete problems where, given a discrete structure, we want to check if this structure has a specific property. For example, given a graph with N vertices, how can we decide if it contains

a Hamiltonian path? If it is a tree? If it is 4-colorable? If it is isomorphic to some other graph? If it is planar? A graph with N vertices has at most $n = \binom{N}{2}$ edges and is completely described if we know which of these edges appear in the graph and which don't. Therefore, we can identify the graph with a vector $x \in \{-1, 1\}^n$, where each digit x_i corresponds to some possible edge: we have $x_i = 1$ if the edge is in the graph and $x_i = -1$ if it is not. In this way, for any "graph property" that we may think we obtain a Boolean function.

Combinatorics: Consider the set $S = \{1, 2, \dots, n\}$ of the first n natural numbers. Then, we may ask several questions about the way that various subsets of this set interact. For example, which is the largest family of subsets of S with the property that no set in the family contains another set of the family? We may think any family $\mathcal{F}$ of subsets of S as a Boolean function, in the following way. Each $x \in \{-1, 1\}^n$ represents a subset of S , where $x_i = 1$ if i belongs to the subset and $x_i = -1$ otherwise. Similarly, every family $\mathcal{F}$ of subsets of S defines a Boolean function $f_{\mathcal{F}}$: for any $x \in \{-1, 1\}^n$ we have that $f_{\mathcal{F}}(x) = 1$ if the subset representing x belongs to $\mathcal{F}$ and $f_{\mathcal{F}}(x) = -1$ otherwise.

8.2 Isoperimetric inequalities on the discrete cube

In this thesis we do not really care about some concrete Boolean functions and their computation. We are mainly interested in the general properties of this class of functions and on the way these properties are related.

We shall study isoperimetric inequalities on the discrete cube. Classical results of this type are due to Margulis [42], Talagrand [48, 51, 49] and Kahn, Kalai and Linial [34]. All these inequalities concern different notions of boundaries and their relationship for Boolean functions.

Given a function $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$ and a point $x \in \{-1, 1\}^n$, we define the sensitivity of f at x , and denote it by $h_f(x)$, to be the number of coordinates $i \in [n]$ for which $f(\sigma_i(x)) \neq f(x)$, where $\sigma_i(x)$ is the vertex of the cube which differs from x only at the i -th coordinate. The vertex boundary V_f of a function f is defined to be the set of all x for which $h_f(x) > 0$, and the edge boundary E_f of f is defined to be the set of all edges $(x, \sigma_i(x))$ of the discrete cube on the vertices of which f assigns different values. The most basic isoperimetric inequality for Boolean functions is the Poincaré inequality, which states that

$$(8.2.1) \quad \frac{|E_f|}{2^n} \geq \text{Var}(f),$$

i.e. that the edge boundary of a Boolean function f cannot be too small. Margulis [42] improved the Poincaré inequality, showing that for any function f we have that

$$(8.2.2) \quad \frac{|V_f|}{2^n} \frac{|E_f|}{2^n} \geq \Omega(\text{Var}(f)^2).$$

Talagrand considered a different notion of boundary for a given function f , the quantity $\mathbb{E}[\sqrt{h_f}]$, which we shall call the Talagrand boundary of f . He obtained a number of results

about this quantity [48, 51], the most basic of them stating that

$$(8.2.3) \quad \mathbb{E} \left[\sqrt{h_f} \right] \geq \Omega(\text{Var}(f)).$$

This inequality strengthens the theorem of Margulis, as one can check with a simple application of the Cauchy-Schwarz inequality. Later, Talagrand strengthened this result in the case of functions with small variance, proving the next theorem.

Theorem 8.2.1 (Talagrand). *There exists an absolute constant $c > 0$ such that, for any function $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$ we have that*

$$(8.2.4) \quad \mathbb{E} \left[\sqrt{h_f} \right] \geq c \text{Var}(f) \sqrt{\log \left(\frac{1}{\text{Var}(f)} \right)}.$$

Although this inequality is sharp up to absolute constants (as one can see considering sub-cubes), it leaves some space for improvements because it is not comparable to some other important isoperimetric result, the (KKL) theorem of Kahn, Kalai and Linial [34].

The (KKL)-theorem concerns a different isoperimetric quantity associated with a given function f , the so-called influence. The influence of a coordinate $i \in [n]$, denoted by $I_i(f)$, is the number of edges in E_f in the i -direction i divided by 2^n , namely

$$(8.2.5) \quad I_i(f) = \mathbb{P}[f(x) \neq f(\sigma_i(x))].$$

The (KKL)-theorem [34] asserts that any function must have at least one variable with non-negligible influence. More precisely, Kahn, Kalai and Linial proved that

$$(8.2.6) \quad \max_i I_i(f) \geq \Omega \left(\frac{\log n}{n} \text{Var}(f) \right).$$

The two isoperimetric inequalities above, Talagrand's inequality and the inequality of Kahn, Kalai and Linial, look very different (the same is true for the techniques that are used in their proofs). Talagrand asked if there exists a single isoperimetric inequality which is strong enough to imply both results at the same time. In [51], Talagrand made a conjecture about the possible form of such an inequality, and made some progress towards its proof. He showed that there exists $\alpha \in (0, 1/2]$ such that every function $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$ satisfies

$$(8.2.7) \quad \mathbb{E} \left[\sqrt{h_f} \right] \geq c \text{Var}(f) \left(\log \left(\frac{1}{\text{Var}(f)} \right) \right)^{1/2-\alpha} \left(\log \left(1 + 1 / \left(\sum_{i=1}^n I_i(f)^2 \right) \right) \right)^\alpha,$$

for some absolute constant $c > 0$. This inequality implies a weaker inequality of (KKL)-type: There exists $\beta > 0$ such that if f is balanced then

$$(8.2.8) \quad \max_i I_i(f) \geq \Omega((\log n)^\beta / n).$$

Note that (8.2.7) does not imply the (KKL)-theorem in its full strength. Talagrand conjectured that (8.2.7) also holds for $\alpha = 1/2$, a statement which is strong enough to give us immediately the (KKL)-theorem.

Eldan and Gross [19] proved that Talagrand's conjecture is true. In other words, we may have $\alpha = 1/2$ in (8.2.7).

Theorem 8.2.2. *There exists an absolute constant $c > 0$ such that, for any function $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$, we have that*

$$(8.2.9) \quad \mathbb{E} \left[\sqrt{h_f} \right] \geq c \operatorname{Var}(f) \sqrt{\log \left(1 + 1 / \left(\sum_{i=1}^n I_i(f)^2 \right) \right)}.$$

We present the proof of Theorem 8.2.2 in Chapters 3, 4 and 5. The method which is introduced by Eldan and Gross is based on stochastic calculus and avoids the use of hypercontractivity that played an essential role in all the previous results. A concise description of the method is given in Chapter 3.

We also present other results from [19], which are proved with the same technique and show that almost-extremal functions for a number of classical inequalities involving the influences of a Boolean function must have large vertex boundary. For example, we give a lower bound for the size of the vertex boundary of a function that makes the inequality

$$\operatorname{Var}(f) \leq C \sum_{i=1}^n \frac{I_i(f)}{1 + \log(1/I_i(f))}$$

almost equality. A corollary of this result is the fact that if, for some set, the Kahn-Kalai-Linial inequality holds almost as an equality then a large portion of the set is contained in its inner vertex boundary. We also describe a strengthened version of an inequality of Keller and Kindler that gives a quantitative relation of influences and noise sensitivity.

In Chapter 6 we present the subsequent paper [21] of Eldan, Kindler, Lifshitz and Minzer, which among other results provides an alternative proof of Theorem 8.2.2. For every $f : \{-1, 1\}^n \rightarrow \mathbb{R}$ and any $x = (x_1, \dots, x_n) \in \{-1, 1\}^n$ we set

$$D_j f(x) = \frac{f(x_1, \dots, x_j, \dots, x_n) - f(x_1, \dots, -x_j, \dots, x_n)}{2}.$$

The main technical result is the next theorem, which is proved with the “method of random restrictions”.

Theorem 8.2.3. *For any $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$ and any $p \in [1, 2]$ we have that*

$$(8.2.10) \quad \|\nabla f\|_p \gtrsim \sup_{d \geq 0} \left(\sum_{|S| \geq d} |\widehat{f}(S)|^2 \right)^{1/p} \sqrt{d}.$$

Choosing $d \approx \log\left(\frac{e}{\text{Var}(f)}\right)$, from (8.2.10) we obtain Talagrand's sharp isoperimetric inequality

$$(8.2.11) \quad \|\nabla f\|_p \gtrsim (\text{Var}(f))^{1/p} \sqrt{\log\left(\frac{e}{\text{Var}(f)}\right)}$$

for every $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$, and choosing $d \approx \log\left(\frac{e}{W(f)}\right)$, where

$$W(f) = \sum_{j=1}^n \left(\mathbb{E} [|D_j f|] \right)^2,$$

from the inequality (8.2.10) we get Talagrand's conjecture

$$(8.2.12) \quad \|\nabla f\|_p \gtrsim (\text{Var}(f))^{1/p} \sqrt{\log\left(\frac{e}{W(f)}\right)}$$

for every $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$, which is equivalent to Theorem 8.2.2.

In Chapter 7 we present the recent work [4] of Beltran, Ivanisvili and Madrid, which establishes various sharp isoperimetric inequalities on the discrete cube. An example is the inequality

$$\mathbb{E} \left[h_A^{\log_2(3/2)} \right] \geq \mu_n(A)^* (\log_2(1/\mu_n(A)^*))^{\log_2(3/2)}$$

for any $A \subseteq \{-1, 1\}^n$, where $\mu_n(A)^* = \min\{\mu_n(A), 1 - \mu_n(A)\}$ and the function h_A is supported on A and maps x to the number of neighbors of x that belong to the complement of A . We also provide lower bounds for $\mathbb{E} [h_A^\beta]$ in terms of $\mu_n(A)$ for every $\beta \in [1/2, 1]$, which improve previous known results and are optimal in some cases. An example is the inequality $\mathbb{E} [h_A^{0.53}] \geq 2\mu_n(A)(1 - \mu_n(A))$ which holds true for every set A with $\mu_n(A) \geq 1/2$, and strengthens an inequality of Kahn and Park.

The same work contains a new proof of (8.2.10) which, compared with the method of random restrictions from [21], has the advantage that it extends the results to functions with values in an arbitrary Banach spaces with finite co-type. Let $\Delta = \sum_{j=1}^n D_j$ and let $e^{-t\Delta}$ be the heat semigroup.

Theorem 8.2.4. *For any $p \in [1, 2]$ and any $f : \{-1, 1\}^n \rightarrow X$ we have that*

$$(8.2.13) \quad \|f - e^{-t\Delta} f\|_p \lesssim C_q(X) q^{3/2} (1 - e^{-2t})^{\frac{1}{q}} \|Df\|_p, \quad t \geq 0$$

where $\|Df\|_p^p = \mathbb{E} \left[\left\| \sum_{1 \leq j \leq n} x'_j D_j f(x) \right\|^p \right]$ and x' is an independent copy of x .

If f is a Boolean function then combining Theorem 8.2.4 with the simple estimate

$$\|f - e^{-\Delta/d} f\|_p \gtrsim \left(\sum_{|S| \geq d} \widehat{f}(S)^2 \right)^{1/p}, \quad p \in [1, 2]$$

we obtain (8.2.10), and Theorem 8.2.2 follows.

Βιβλιογραφία

- [1] F. Barthe and B. Maurey, *Some remarks on isoperimetry of Gaussian type*, Ann. Inst. H. Poincaré Probab. Statist. 36(4), 419–434 (2000).
- [2] W. Beckner, *Inequalities in Fourier analysis*, Ann. Math., 102:159–182, 1975.
- [3] V. Beffara and H. Duminil-Copin, *The self-dual point of the two-dimensional random-cluster model is critical for $q \geq 1$* , Probab. Theory Relat. Fields 153(3-4), 511–542 (2012).
- [4] D. Beltran, P. Ivanisvili and J. Madrid, *On sharp isoperimetric inequalities on the hypercube*, Preprint, arXiv:2303.06738.
- [5] I. Benjamini, G. Kalai and O. Schramm, *Noise sensitivity of Boolean functions and applications to percolation*, Inst. Hautes Etudes Sci. Publ. Math. 90, 5–43 (1999).
- [6] I. Benjamini, G. Kalai and O. Schramm, *First passage percolation has sublinear distance variance*, Ann. Probab. 31(4), 1970–1978 (2003).
- [7] S. G. Bobkov, *An isoperimetric inequality on the discrete cube, and an elementary proof of the isoperimetric inequality in Gauss space*, Ann. Probab. 25(1), 206–214 (1997).
- [8] S. G. Bobkov, F. Götze, *Discrete isoperimetric and Poincaré-type inequalities*, Probab. Theory Relat Fields 114 (1999), no. 2, 245–277.
- [9] A. Bonami, *Étude des coefficients de Fourier des fonctions de $L^p(G)$* , Ann. Inst. Fourier (Grenoble) 20 (1971), no. 2:335–402.
- [10] J. Bourgain, *On the distribution of the fourier spectrum of boolean functions*, Israel J. Math. 131(1):269–276, 2002.
- [11] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, Cambridge (2004).
- [12] M. Capitaine, E. P. Hsu and M. Ledoux, *Martingale representation and a simple proof of logarithmic Sobolev inequalities on path spaces*, Electron. Commun. Probab. 2, 71–81 (1997).
- [13] D. Cordero-Erausquin and M. Ledoux, *Hypercontractive Measures, Talagrand's Inequality, and Influences*, pp. 169–189. Springer, Berlin (2012).
- [14] I. Dinur and S. Safra, *On the hardness of approximating minimum vertex cover*, Ann. Math. (2) 162(1), 439–485 (2005).
- [15] R. Durrett, *Probability-theory and examples*, In: Cambridge Series in Statistical and Probabilistic Mathematics, vol. 49. Cambridge University Press, Cambridge (2019).
- [16] R. Eldan, *A two-sided estimate for the Gaussian noise stability deficit*, Invent. Math. 201(2), 561–624 (2015).
- [17] R. Eldan, *Second-order bounds on correlations between increasing families*, Lecture Notes.
- [18] R. Eldan, *Analysis of high-dimensional distributions using pathwise methods*, In: 2022 ICM Proceedings (2021).
- [19] R. Eldan and R. Gross, *Concentration on the Boolean hypercube via pathwise stochastic analysis*, STOC '20- Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing, 208–221, ACM, New York (2020).

- [20] R. Eldan and R. Gross, *Concentration on the Boolean hypercube via pathwise stochastic analysis*, Invent. Math. 230 (2022), no. 3, 935–994.
- [21] R. Eldan, G. Kindler, N. Lifshitz, D. Minzer, *Isoperimetric Inequalities Made Simpler*, arXiv:2204.06686.
- [22] D. Ellis, *Almost isoperimetric subsets of the discrete cube*, Combin. Probab. Comput. 20(3), 363–380 (2011).
- [23] E. Friedgut, *Boolean functions with low average sensitivity depend on few coordinates*, Combinatorica 18(1), 27–35 (1998).
- [24] E. Friedgut, *Sharp thresholds of graph properties, and the k -sat problem*, J. Am. Math. Soc. 12(4), 1017–1054 (1999).
- [25] C. Garban and J. E. Steif, *Noise sensitivity of Boolean functions and percolation*, In: Institute of Mathematical Statistics Textbooks, vol. 5. Cambridge University Press, New York (2015).
- [26] D. Gavinsky, J. Kempe, I. Kerenidis, R. Raz, and R. de Wolf, *Exponential separation for one-way quantum communication complexity, with applications to cryptography*, SIAM J. Comput. 38(5), 1695–1708 (2008/2009).
- [27] L. Gross, *Logarithmic Sobolev inequalities*, Amer. J. Math. 97(4):1061–1083, 1975.
- [28] S. Hart, *A note on the edges of the n -cube*, Discrete Math. 14(2), 157–163 (1976).
- [29] P. Iwanisvili, A. Nazarov and A. Volberg, *Square functions and the Hamming cube: Duality*, Discrete Anal. 1 (2018), pp. 1–18.
- [30] P. Iwanisvili, R. van Handel and A. Volberg, *Improving constant in end-point Poincaré inequality on Hamming cube*, arXiv: 1811.05584.
- [31] P. Iwanisvili, R. van Handel and A. Volberg, *Rademacher type and Enflo type coincide*, Ann. of Math. (2) 192 (2020), no. 2, 665–678.
- [32] J. Kahn and G. Kalai, *Thresholds and expectation thresholds*, Combin. Probab. Comput. 16(3), 495–502 (2007).
- [33] J. Kahn, J. Park, *An isoperimetric inequality for the Hamming cube and some consequences*, Proc. Amer. Math. Soc. 148 (2020), 4213–4224.
- [34] J. Kahn, G. Kalai and N. Linial, *The influence of variables on Boolean functions*, 29th Annual Symposium on Foundations of Computer Science, 68–80, IEEE Comput. Soc. Press, Washington, DC, 1988.
- [35] G. Kalai and S. Safra, *Threshold phenomena and influence: perspectives from mathematics, computer science, and economics*, In: Computational Complexity and Statistical Physics, St. Fe Inst. Stud. Sci. Complex., Oxford University Press, New York, pp. 25–60 (2006).
- [36] N. Keller and G. Kindler, *Quantitative relation between noise sensitivity and influences*, Combinatorica 33(1), 45–71 (2013).
- [37] S. Khot, D. Minzer and M. Safra, *On monotonicity testing and Boolean isoperimetric-type theorems*, SIAM J. Comput., 47(6):2238–2276, 2018.
- [38] J. F. C. Kingman, *Poisson processes*, In: Oxford Studies in Probability, vol. 3, The Clarendon Press, Oxford University Press, New York, Oxford Science Publications (1993).
- [39] S. Klein, A. Levi, M. Safra, C. Shikhelman and Y. Spinka, *On the converse of Talagrand’s influence inequality*, arXiv:1506.06325 (2015).
- [40] A. Klivans, R. O’Donnell, R. Servedio, *Learning Geometric Concepts via Gaussian Surface Area*, <https://www.cs.cmu.edu/~odonnell/papers/perimeter.pdf>
- [41] R. Krauthgamer and Y. Rabani, *Improved lower bounds for embeddings into L_1* , SIAM J. Comput. 38(6), 2487–2498 (2009).
- [42] G. A. Margulis, *Probabilistic characteristics of graphs with large connectivity*, Problemy peredachi informatsii, 10(2):101–108, 1974.
- [43] E. Mossel, J. Neeman and O. Tamuz, *Majority dynamics and aggregation of information in social networks*, Auton. Agent. Multi-Agent Syst. 28(3), 408–429 (2014).

-
- [44] R. O'Donnell, *Analysis of Boolean Functions*, Cambridge University Press, New York (2014).
 - [45] R. O'Donnell and R. A. Servedio, *Learning monotone decision trees in polynomial time*, SIAM J. Comput. 37(3), 827–844 (2007).
 - [46] C. Rashtchian and W. Raynaud, *Edge Isoperimetric Inequalities for Powers of the Hypercube*, Electron. J. Combin. 29 (2022), no. 1, Paper No. 1.47, 26 pp.
 - [47] R. Raz, *Fourier analysis for probabilistic communication complexity*, Comput. Complex. 5(3–4), 205–221 (1995).
 - [48] M. Talagrand, *Isoperimetry, logarithmic Sobolev inequalities on the discrete cube, and Margulis' graph connectivity theorem*, Geom. Funct. Anal. 3(3), 295–314 (1993).
 - [49] M. Talagrand, *On Russo's approximate zero-one law*, Ann. Probab. 22(3), 1576–1587 (1994).
 - [50] M. Talagrand, *How much are increasing sets positively correlated?*, Combinatorica 16(2), 243–258 (1996).
 - [51] M. Talagrand, *On boundaries and influences*, Combinatorica 17(2), 275–285 (1997).
 - [52] R. Vershynin, *High-dimensional probability. An introduction with applications in data science*, Cambridge Series in Statistical and Probabilistic Mathematics, 47. Cambridge University Press, Cambridge (2018).

Abstract

In this thesis we study isoperimetric inequalities on the discrete cube. Classical results of this type are due to Margulis, Talagrand and Kahn, Kalai and Linial. All these inequalities concern different notions of boundaries and their relationship for Boolean functions. Given a function $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$ and a point $x \in \{-1, 1\}^n$, we define the sensitivity of f at x , and denote it by $h_f(x)$, to be the number of coordinates $i \in [n]$ for which $f(\sigma_i(x)) \neq f(x)$, where $\sigma_i(x)$ is the vertex of the cube which differs from x only at the i -th coordinate. Talagrand considered the quantity $\mathbb{E}[\sqrt{h_f}]$ and made a conjecture about the possible form of a stronger inequality. He showed that there exists $\alpha \in (0, 1/2]$ such that every function $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$ satisfies

$$\mathbb{E}[\sqrt{h_f}] \geq c \operatorname{Var}(f) \left(\log \left(\frac{1}{\operatorname{Var}(f)} \right) \right)^{1/2-\alpha} \left(\log \left(1 + 1 / \left(\sum_{i=1}^n I_i(f)^2 \right) \right) \right)^\alpha,$$

for some absolute constant $c > 0$ and conjectured that this inequality also holds for $\alpha = 1/2$; this would imply the isoperimetric inequalities of Margulis, Talagrand and Kahn, Kalai and Linial. Eldan and Gross proved that Talagrand's conjecture is true: There exists an absolute constant $c > 0$ such that, for any function $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$, we have that

$$\mathbb{E}[\sqrt{h_f}] \geq c \operatorname{Var}(f) \sqrt{\log \left(1 + 1 / \left(\sum_{i=1}^n I_i(f)^2 \right) \right)}.$$

We present the proof of this theorem, which is based on stochastic calculus and avoids the use of hypercontractivity that played an essential role in all the previous results. In the last two chapters of the thesis we discuss alternative subsequent approaches to the Eldan-Gross theorem.