
A-optimale Blockpläne für Behandlungs-Kontroll-Vergleiche bei AR(1)-korrelierten Fehlern

Dipl.-Stat. Oliver Melsheimer
Dortmund im Juli 2011

eingereicht an der Technischen Universität Dortmund
zur Erlangung des Doktorgrades der Naturwissenschaften

Inhaltsverzeichnis

1	Einleitung	1
2	Grundlagen	2
2.1	Hintergrund und Bezeichnungen	2
2.2	Die Suche nach optimalen Plänen	4
2.3	Allgemeine Aussagen über die Einträge der Matrix W	8
3	Reduzierung der Anzahl zu betrachtender Sequenzen	13
3.1	Die Ansätze von Kunert, Martin und Eccleston	13
3.2	Weitere Ansätze zur Reduzierung	15
4	Reduzierung der Anzahl zu betrachtender Schnittpunkte	24
5	Das Vorgehen bei der Suche nach optimalen Sequenzen	31
6	Allgemeine Aussagen über die optimalen Sequenzen	36
7	Zusammenfassung	56

1 Einleitung

Diese Arbeit soll eine Untersuchung fortführen, die Kunert, Martin und Eccleston in ihrem Artikel „*A-optimal block designs for the comparison with a control for correlated errors and analysis with the weighted least squares estimate*“ begonnen haben. In ihrem Artikel versuchen die Autoren, optimale approximative Versuchspläne für den Vergleich mehrerer Behandlungen mit einer Kontrolle in einem einfachen Blockmodell bei AR(1)-korrelierten Störgrößen zu finden oder zumindest für gegebene Rahmenbedingungen Untergrenzen für den Wert des betrachteten Gütekriteriums, der durchschnittlichen Varianz der Kontrastschätzer zwischen Behandlungseffekten und dem Effekt der Kontrollbehandlung, zu ermitteln. Dabei beschränken sich die Autoren aber hauptsächlich auf zwei Spezialfälle, nämlich Blockpläne mit Blocklängen von 3 und 4. Die Frage, wie Pläne bzw. Untergrenzen für beliebige Blocklängen k aussehen, wird dort also noch nicht beantwortet und soll daher in dieser Arbeit untersucht werden.

Die Grundlagen zum Verständnis dieser Arbeit sollen dazu in Kapitel 2 vermittelt werden, ab Kapitel 3 wird dann systematisch auf das Auffinden der optimalen Sequenzen hingearbeitet. Zu dieser Suche seien an dieser Stelle noch einige Anmerkungen gegeben: Zunächst einmal sei betont, dass sie natürlich unter genau denselben Annahmen stattfindet, die auch [Kunert et al., 2010] ihrer Arbeit zugrunde gelegt haben. Damit ist insbesondere eine Beschränkung auf Fälle gemeint, in denen die Blocklänge k die Zahl der mit der Kontrolle zu vergleichenden Behandlungen v nicht übersteigt, und in denen der Korrelationsparameter λ des AR(1)-Prozesses im Intervall $[0, 1)$ liegt. [Kunert et al., 2010] untersuchten den Fall $k = 3$ zwar noch für beliebiges $\lambda \in (-1, 1)$, beschränkten sich aber auch schon im Fall $k = 4$ auf nicht-negative λ -Werte, da sie sonst eines ihrer grundlegenden Resultate nicht hätten anwenden können. Da das Auftreten einer negativen Korrelation innerhalb der Blöcke praktisch eh ohne Relevanz sein dürfte, stellt diese Beschränkung auch keinen wirklichen Nachteil dar.

Als Hilfsmittel bei allen Berechnungen wurden die beiden Programmpakete [Maxima, 2010] und [R, 2009] eingesetzt. Letzteres wurde dabei insbesondere auch zum Erstellen der Grafiken verwendet. Da bei vielen der im Rahmen dieser Arbeit vorgenommenen Berechnungen es nur darum ging zu zeigen, dass ein Ausdruck ein bestimmtes Vorzeichen hat, und viele Umformungen folgenlos für dieses bleiben, sei folgende Kurzschreibweise vereinbart: Durch das Proportionalitätszeichen \propto seien Umformungsschritte, durch die sich das Vorzeichen eines Ausdrucks nicht verändert, gekennzeichnet. Wenn A und B zwei mathematische Ausdrücke sind, dann soll die Schreibweise $A \propto B$ also einfach nur stellvertretend zu verstehen sein für $\text{sig}(A) = \text{sig}(B)$.

2 Grundlagen

In diesem Kapitel sollen die Hintergründe der Arbeit dargelegt und so die Grundlage für das Verständnis der kommenden Kapitel geschaffen werden. Dabei werden im ersten Unterkapitel zunächst einige sehr grundlegende Begriffe und Bezeichnungen eingeführt und erklärt, sowie die zu betrachtende Ausgangssituation kurz dargestellt. Wie man in einer Situation wie der in diesem Abschnitt dargelegten nach guten Versuchsplänen suchen kann, wird unter Verweis auf ein Resultat von [Kunert et al., 2010] im darauf folgenden Abschnitt erläutert. Das dritte Unterkapitel beschäftigt sich intensiv mit einer Matrix, die für diese Suche von allergrößter Bedeutung ist: Eigentlich alle der in dieser Arbeit folgenden Beweise basieren nämlich auf einigen elementaren Aussagen über die Einträge einer Matrix W . All diese Aussagen werden in diesem letzten Abschnitt des Kapitels hergeleitet.

2.1 Hintergrund und Bezeichnungen

Die Ausgangsbasis aller Überlegungen in dieser Arbeit stellt das einfache Blockmodell dar:

$$\mathbf{y} = (T_d, B) \begin{pmatrix} \boldsymbol{\tau} \\ \boldsymbol{\kappa} \end{pmatrix} + \mathbf{e}$$

In diesem Modell bezeichnet \mathbf{y} den *Vektor der Beobachtungen* und \mathbf{e} den *Vektor der Störgrößen*. $\boldsymbol{\tau}$ steht für den *Vektor der Behandlungseffekte*, der $v + 1$ Einträge umfasst: Neben dem mit τ_0 bezeichneten Effekt einer Kontrollbehandlung sind dies die Effekte von v verschiedenen Behandlungen, die mit der Kontrollbehandlung verglichen werden sollen. Bei dem Vektor $\boldsymbol{\kappa}$, der für jeden der b gleichlangen Blöcke des Designs einen Eintrag enthält, handelt es sich um den *Vektor der Blockeffekte*.

Die drei Vektoren haben also das folgende Aussehen:

$$\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_{bk} \end{pmatrix} \sim (bk, 1) \quad \boldsymbol{\tau} = \begin{pmatrix} \tau_0 \\ \tau_1 \\ \vdots \\ \tau_v \end{pmatrix} \sim (v + 1, 1) \quad \boldsymbol{\kappa} = \begin{pmatrix} \kappa_1 \\ \kappa_2 \\ \vdots \\ \kappa_b \end{pmatrix} \sim (b, 1)$$

k steht hierbei für die Länge eines jeden Blocks.

Bei den Matrizen T_d und B handelt es sich um die *Behandlungs-Design-Matrix* und die *Block-Design-Matrix*. Diese beiden Matrizen sind so aufgebaut, dass sie in jeder Zeile genau eine Eins enthalten und sonst nur Nullen. Die Positionen der Einsen in den jeweils i -ten Zeilen dieser Matrizen geben Aufschluss darüber, welche Behandlung bei Versuchsobjekt i eingesetzt wurde und in welchem Block von Beobachtungen es sich befunden hat.

Die Matrizen sind entsprechend von den folgenden Dimensionen:

$$T_d \sim (bk, v + 1) \quad B = I_b \otimes 1_k \sim (bk, b)$$

Welche Behandlungen die Versuchsobjekte bekommen, wird auch in Form eines *Blockplans* angegeben: Ein solcher Blockplan d ist eine Matrix mit b Zeilen und k Spalten, deren Einträge jeweils der Nummer der eingesetzten Behandlung entsprechen. Die Zeilen der Matrix stehen also für die verschiedenen Blöcke, die Spalten für die Positionen im Block. Die Darstellung als Blockplan enthält nicht mehr und nicht weniger Informationen, als sie auch schon die Matrizen T_d und B bieten, stellt diese aber in einer kompakteren Form dar.

Als *Sequenz* wird eine einzelne Zeile dieses Blockplans bezeichnet. Eine Sequenz s gibt also für einen Block an, an welcher Stelle welche Behandlung verwendet wird. Die Schreibweise einer Sequenz ist üblicherweise so definiert, dass die Behandlungsnummern in der Reihenfolge ihres Auftretens im Block in eckige Klammern geschrieben werden. Die Sequenz $[0 \ 1 \ 2 \ 3 \ 0 \ 4]$ bedeutet also beispielsweise für einen Block der Länge 6, dass bei dem ersten und fünften Versuchsobjekt des Blocks die Kontrollbehandlung 0 eingesetzt wurde und die anderen vier Versuchsobjekte (in dieser Reihenfolge) mit den Behandlungen 1 bis 4 behandelt wurden.

In dieser Arbeit wird nur der Fall untersucht, dass e , der *Vektor der Störgrößen*, eine ganz spezielle Gestalt hat. Es wird nämlich davon ausgegangen, dass nur Beobachtungen aus verschiedenen Blöcken voneinander unabhängig, die Beobachtungen innerhalb eines Blocks aber gemäß eines AR(1)-Prozesses miteinander korreliert sind. Der Vektor y hat damit also die folgende Verteilung:

$$y \sim \mathbb{N}(T_d \tau + B \kappa, \Sigma)$$

mit

$$\Sigma = \sigma^2 (I_b \otimes S_\lambda)$$

und

$$S_\lambda = \begin{pmatrix} 1 & \lambda & \lambda^2 & \lambda^3 & \dots & \lambda^{k-3} & \lambda^{k-2} & \lambda^{k-1} \\ \lambda & 1 & \lambda & \lambda^2 & \dots & \lambda^{k-4} & \lambda^{k-3} & \lambda^{k-2} \\ \lambda^2 & \lambda & 1 & \lambda & \dots & \lambda^{k-5} & \lambda^{k-4} & \lambda^{k-3} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \lambda^{k-2} & \lambda^{k-3} & \lambda^{k-4} & \lambda^{k-5} & \dots & \lambda & 1 & \lambda \\ \lambda^{k-1} & \lambda^{k-2} & \lambda^{k-3} & \lambda^{k-4} & \dots & \lambda^2 & \lambda & 1 \end{pmatrix}$$

Die Größe λ steht hierbei für den Korrelationsparameter des AR(1)-Prozesses, σ^2 für die ihm zugrundeliegende Fehlervarianz.

In der Versuchsplanung beschäftigt man sich mit der Fragestellung, welche Behandlung welches Versuchsobjekt erhalten sollte, damit die Beantwortung der zugrundeliegenden statistischen Fragestellung am besten erfolgen kann. Es gilt also, durch geschickte Wahl des Plans d eine Optimalität hinsichtlich eines bestimmten Gütekriteriums zu erzielen. Dabei gibt es natürlich eine Vielzahl möglicher Gütekriterien, die alle herangezogen werden könnten. Das Gütekriterium,

das jedoch in dieser Arbeit ausschließlich betrachtet werden soll, ist das *A-Kriterium für den Vergleich mit einer Kontrolle*. Dieses Kriterium ist definiert als die durchschnittliche Varianz der Kontrastschätzer zwischen den Behandlungseffekten und dem Effekt der Kontrollbehandlung:

$$A_{tc}(d) = \frac{1}{v} \sum_{i=1}^v \text{var}(\widehat{\tau}_i - \widehat{\tau}_0)$$

Dabei bezeichne d den jeweiligen Blockplan aus der Menge aller möglichen Blockpläne \mathcal{D} mit b Blöcken, Blocklänge k sowie einer Kontroll- und v Vergleichsbehandlungen.

2.2 Die Suche nach optimalen Plänen

Als optimal wird ein Plan bezeichnet, wenn er unter allen Plänen aus \mathcal{D} den kleinsten Wert des Kriteriums erzielt. Ein Plan d^* wird demnach *A-optimal für den Vergleich mit einer Kontrolle* genannt, wenn für ihn gilt:

$$A_{tc}(d^*) = \min_{d \in \mathcal{D}} A_{tc}(d).$$

Es gelten nun die folgenden Bezeichnungen:

- Mit $s_{[1]}, \dots, s_{[b]}$ seien die nicht notwendigerweise verschiedenen Sequenzen bezeichnet, die (in dieser Reihenfolge) in den b Blöcken des Versuchsplans Verwendung finden.
- W sei die Matrix $S_\lambda^{-1} - \frac{1}{\mathbf{1}_k^T S_\lambda^{-1} \mathbf{1}_k} S_\lambda^{-1} \mathbf{1}_k \mathbf{1}_k^T S_\lambda^{-1}$ ($\sim (k, k)$).
- Mit w_{ij} seien die einzelnen Einträge dieser Matrix W bezeichnet, wobei i die Zeile und j die Spalte des jeweiligen Elements der Matrix angibt (mit $i, j = 1, \dots, k$).
- $\mathbf{t}^{(u)}(s)$ bezeichne einen Vektor der Länge k , der an all den Stellen eine 1 enthält, an denen in Sequenz s Behandlung u eingesetzt wurde, und sonst nur Nullen. Für den q -ten Eintrag dieses Vektors, $t_q^{(u)}(s)$, gilt also:

$$t_q^{(u)}(s) = \begin{cases} 1 & : \text{Behandlung } u \text{ steht an Stelle } q \text{ von Sequenz } s \\ 0 & : \text{sonst} \end{cases}$$

Die Behandlungs-Design-Matrix des Versuchsplans d lässt sich so also auch darstellen als

$$T_d = \begin{pmatrix} \mathbf{t}^{(0)}(s_{[1]}) & \mathbf{t}^{(1)}(s_{[1]}) & \dots & \mathbf{t}^{(v)}(s_{[1]}) \\ \mathbf{t}^{(0)}(s_{[2]}) & \mathbf{t}^{(1)}(s_{[2]}) & \dots & \mathbf{t}^{(v)}(s_{[2]}) \\ \vdots & \vdots & & \vdots \\ \mathbf{t}^{(0)}(s_{[b]}) & \mathbf{t}^{(1)}(s_{[b]}) & \dots & \mathbf{t}^{(v)}(s_{[b]}) \end{pmatrix}.$$

- Die Matrix M_d ist definiert als $\widetilde{T}_d^T (I_b \otimes W) \widetilde{T}_d$, wobei mit \widetilde{T}_d der Teil der Behandlungs-Design-Matrix T_d ohne die zur Kontrollbehandlung gehörende erste Spalte gemeint ist.
- $c_{jj}(s)$ bezeichne den Ausdruck $\mathbf{t}^{(j)}(s)^T \cdot W \cdot \mathbf{t}^{(j)}(s)$, der sich in alternativer Schreibweise auch als folgende Doppelsumme darstellen lässt:

$$c_{jj}(s) = \sum_{r_1=1}^k \sum_{r_2=1}^k t_{r_1}^{(j)}(s) \cdot w_{r_1 r_2} \cdot t_{r_2}^{(j)}(s)$$

- Aus diesen c_{jj} -Werten berechnen sich nun die m -Werte einer Sequenz s :

$$m_2(s) = \frac{1}{v} c_{00}(s)$$

$$m_1(s) = \frac{1}{v-1} \left(\sum_{j=1}^v c_{jj}(s) - m_2(s) \right).$$

- Die folgende auf diesen m -Werten basierende Funktion wird als H -Funktion einer Sequenz s bezeichnet und ist für diese Arbeit von allergrößter Bedeutung:

$$H_s(x) = (1+x)^2 m_1(s) + x^2 (v-1) m_2(s).$$

Wie [Kunert et al., 2010] mit Hilfe von Resultaten aus [Majumdar und Notz, 1983], [Kushner, 1997] und [Kunert und Martin, 2000] gezeigt haben, stellt dann die Erfüllung der folgenden drei Forderungen eine hinreichende Bedingung für die A_{tc} -Optimalität eines Plans d dar:

- (I) Bezeichne x^* die Stelle, für die

$$\max_s H_s(x^*) = \min_x \max_s H_s(x)$$

gilt. Für alle im Versuchsplan d eingesetzten Sequenzen $s_{[1]}, \dots, s_{[b]}$ muss dann gelten:

$$H_{s_{[q]}}(x^*) = \max_s H_s(x^*) \quad (q = 1, \dots, b)$$

- (II) Die im Versuchsplan eingesetzten Sequenzen sind so gewählt, dass für die Summe der Ableitungen ihrer H -Funktionen nach x gilt:

$$\sum_{q=1}^b H'_{s_{[q]}}(x^*) = 0.$$

- (III) Die Matrix M_d ist vollständig symmetrisch und lässt sich also darstellen als

$$M_d = c_1 \cdot I_v + c_2 \mathbf{1}_v \mathbf{1}_v^T \quad (c_1, c_2 \in \mathbb{R}).$$

Alle ihre Diagonalelemente nehmen also ebenso wie auch all ihre Nebendiagonalelemente jeweils denselben Wert an.

Nicht für jede Kombination der Parameter b , k und v lässt sich aber ein Plan konstruieren, der alle drei dieser Forderungen erfüllt: Während die erste Forderung immer erfüllbar ist, stellen die anderen beiden nämlich für konkrete Parameterwerte unter Umständen eine unüberwindbare Hürde dar.

Ein konstruktives Vorgehen, wie man in dieser Situation dann den optimalen Plan ermittelt, ist nicht bekannt. Da nach [Kunert et al., 2010] aber auch gilt, dass

$$A_{tc}(d) \geq \frac{v-1}{v \cdot b \cdot \max_s H_s(x^*)} \cdot \sigma^2$$

ist, lässt sich die Güte bzw. Effizienz eines konkreten Plans durch einen Vergleich mit dieser theoretischen Schranke ermitteln. Ist für einen Plan d der Quotient

$$E(d) = \frac{v-1}{v \cdot b \cdot \max_s H_s(x^*)} \lambda_{tc}(d)$$

nahe an 1, so hat man einen zwar nicht unbedingt optimalen, aber dennoch für die konkrete Situation sehr guten Plan gefunden.

Ob man sich jetzt mit einem „guten“ Plan zufrieden gibt oder auf die Suche nach dem optimalen (approximativen) Plan macht, an der Ermittlung der Stelle

$$x^* = \operatorname{argmin}_x \max_s H_s(x)$$

und dem höchsten dort angenommenen Wert der H-Funktionen

$$\max_s H_s(x^*) = \min_x \max_s H_s(x)$$

führt so oder so kein Weg vorbei. Gegenstand dieser Arbeit soll es daher sein, Aussagen über die Lage dieser Stelle x^* und die dort höchsten H-Funktionen zu treffen, sowie zu untersuchen, welche Auswirkungen Änderungen der Parameter darauf haben. Die Menge der Sequenzen s , für die an der Stelle x^* die höchsten H-Funktionswerte erreicht werden, werden in der weiteren Arbeit dabei der Einfachheit halber als *optimale Sequenzen* bezeichnet, auch wenn ein aus ihnen zusammengestellter Versuchsplan wie bereits dargelegt natürlich nicht notwendigerweise optimal sein muss.

Mit den folgenden Behauptungen lässt sich die Suche nach der Stelle x^* bereits etwas eingrenzen:

Behauptung 2.1: *Für die m -Werte einer Sequenz s gilt: $m_1(s)$ und $m_2(s)$ sind beide größer oder gleich Null.*

Beweis: Der Beweis folgt direkt aus [Majumdar und Notz, 1983]. Die Autoren zeigen dort in *Lemma 2.1*, dass es sich bei den hier als $m_1(s)$ und $m_2(s)$ bezeichneten Werten um die Eigenwerte einer Matrix $\bar{M}(d)$ für einen nur aus einem einzigen Block bestehenden Versuchsplan d handelt. Diese eben erwähnte Matrix $\bar{M}(d)$ berechnet sich dabei als Durchschnitt über sämtliche M -Matrizen, die sich bei einer Permutation der Behandlungen in d ergeben.

Da nach [Majumdar und Notz, 1983] jede einzelne dieser M -Matrizen positiv semi-definit ist, muss dies natürlich auch für $\bar{M}(d)$ gelten. Da eine positiv semi-definite Matrix aber keine negativen Eigenwerte haben kann, ist klar, dass beide m -Werte größer oder gleich Null sein müssen. \square

Behauptung 2.2: *Sämtliche H-Funktionen sind an keiner Stelle negativ. Sofern nicht $m_1(s)$ und $m_2(s)$ beide gleich Null sind, handelt es sich bei der zu einer Sequenz s gehörenden H-Funktion zudem um eine nach oben geöffnete Parabel, die als solche natürlich streng konvex ist.*

Beweis: Dass die H-Funktionen überall nicht-negativ sind, ergibt sich direkt aus ihrer Definition und der Tatsache, dass auch die m -Werte größer oder gleich Null sind (siehe Behauptung 2.1). Mit derselben Begründung ergibt sich aber weiter, dass es sich bei den H-Funktionen um nach oben geöffnete Parabeln handelt, denn schließlich ist $H_s(x)$ nichts anderes als eine (so-

fern wenigstens einer der m -Werte von Null verschieden ist) positive Linearkombination der beiden nach oben geöffneten Parabeln $(1+x)^2$ und x^2 :

$$H_s(x) = (1+x)^2 m_1(s) + x^2(v-1)m_2(s).$$

Die Konvexität folgt nun zwar schon unmittelbar aus dieser Tatsache, sie ließe sich aber natürlich auch explizit durch Betrachtung der zweiten Ableitung von $H_s(x)$ nach x nachweisen:

$$\frac{\partial^2 H_s(x)}{\partial^2 x} = 2m_1(s) + 2(v-1)m_2(s).$$

Da eben $m_1(s)$ als auch $m_2(s)$ nicht-negativ und nicht beide gleich Null sind, ist diese zweite Ableitung für alle Werte von x positiv und die Funktion somit in jedem Fall streng konvex. \square

Behauptung 2.3: Die gesuchte Stelle $x^* = \operatorname{argmin}_x \max_s H_s(x)$ kann sich nur an einem der im Intervall $[-1, 0]$ liegenden Schnitt- und Scheitelpunkte der H -Funktionen befinden.

Beweis: Sollten für eine Sequenz s beide m -Werte gleich Null sein, so ist auch ihre H -Funktion überall gleich Null. Bei dieser Funktion kann es sich dann aber nicht um $\max_s H_s(x)$ handeln, da jede Sequenz, für die auch nur einer der m -Werte größer als Null ist, überall echt positive H -Werte erzielt und damit der erstgenannten Sequenz überall überlegen wäre. Eine weitere Betrachtung dieser Sequenz würde also keinen Sinn machen, daher werden im Folgenden nur noch Sequenzen untersucht, für die mindestens einer der m -Werte größer Null ist.

Nach Behauptung 2.2 ist für solche Sequenzen $H_s(x)$ aber eine nach oben geöffnete Parabel. Sei nun \tilde{x} eine Stelle, an der kein Scheitel- oder Schnittpunkt zweier oder mehrerer Funktionen liegt, und sei s_m die Sequenz, die an dieser Stelle die höchsten H -Werte erreicht.

Da es bei \tilde{x} keinen Schnittpunkt gibt, lässt sich aufgrund der Stetigkeit der H -Funktionen in jedem Fall ein $\Upsilon > 0$ finden, so dass für alle möglichen Sequenzen s gilt:

$$H_{s_m}(x) \geq H_s(x) \quad \text{für } x \in [\tilde{x} - \Upsilon, \tilde{x} + \Upsilon].$$

Nicht nur an der Stelle \tilde{x} , sondern auch in dieser Υ -Umgebung um \tilde{x} erreicht die Sequenz s_m also höhere oder mindestens genau so hohe H -Werte wie alle anderen Sequenzen. Da es sich bei $H_{s_m}(x)$ um eine Parabel handelt und an der Stelle \tilde{x} kein Scheitelpunkt liegt, gilt für die Ableitung von $H_{s_m}(x)$ nach x aber auch

$$H'_{s_m}(\tilde{x}) \neq 0,$$

so dass es entweder im Intervall $[\tilde{x} - \Upsilon, \tilde{x})$ oder im Intervall $(\tilde{x}, \tilde{x} + \Upsilon]$ mindestens eine Stelle \tilde{x} geben muss, für die

$$H_{s_m}(\tilde{x}) > H_{s_m}(\tilde{x})$$

und damit auch

$$\max_s H_s(\tilde{x}) > \max_s H_s(\tilde{x})$$

gilt. Der Punkt \tilde{x} kommt somit nicht als gesuchte Stelle $\operatorname{argmin}_x \max_s H_s(x)$ in Frage.

Die gesuchte Stelle x^* muss daher an einem der Scheitel- oder Schnittpunkte der H-Funktionen liegen, und zwar einem innerhalb des Intervalls $[-1, 0]$: Betrachtet man nämlich die erste Ableitung von $H_s(x)$ nach x und beschränkt sich dabei wie zu Beginn des Beweises erwähnt auf Sequenzen, für die einer der beiden m -Werte $m_1(s)$ und $m_2(s)$ positiv und der andere zumindest nicht-negativ ist, so stellt man fest, dass

$$H'_s(x) = \frac{\partial H_s(x)}{\partial x} = \underbrace{2(1+x)}_{>0 \text{ für } x > -1, <0 \text{ für } x < -1} \cdot m_1(s) + \underbrace{2(v-1)x}_{>0 \text{ für } x > 0, <0 \text{ für } x < 0} \cdot m_2(s) \begin{array}{l} > 0 \text{ für } x > 0 \\ < 0 \text{ für } x < -1 \end{array}$$

ist. Außerhalb des Intervalls $[-1, 0]$ haben die Ableitungen aller H-Funktionen jeweils dasselbe Vorzeichen: Sie sind allesamt streng monoton fallend für $x < -1$ und streng monoton wachsend für $x > 0$. Dies hat aber zur Folge, dass für beliebige Sequenz s

$$H_s(x) > H_s(-1) \quad \text{für } x < -1$$

$$H_s(x) > H_s(0) \quad \text{für } x > 0$$

gilt und somit natürlich insbesondere auch

$$\max_s H_s(x) > \max_s H_s(-1) \quad \text{für } x < -1$$

$$\max_s H_s(x) > \max_s H_s(0) \quad \text{für } x > 0$$

ist. Damit ist gezeigt, dass es sich bei Stellen außerhalb des Intervalls nicht um das gesuchte Minimum handeln kann. \square

Um die Suche nach der Optimalstelle x^* nun noch weiter einzuzugrenzen, sind grundsätzliche Erkenntnisse über das Aussehen und Verhalten der zu den jeweiligen Sequenzen gehörenden H-Funktionen nötig. Um zu diesen zu gelangen, ist es hilfreich, sich zunächst einmal mit den Einträgen der Matrix W zu beschäftigen. Dies geschieht im nächsten Unterkapitel.

2.3 Allgemeine Aussagen über die Einträge der Matrix W

Die im vorigen Unterkapitel eingeführte Matrix W , die die Grundlage für alle nun folgenden Überlegungen bildet, weist - wie [Kunert et al., 2010] zeigen - beim AR(1)-Prozess höchstens sieben verschiedene Einträge auf: Die Diagonalelemente w_{ii} sind entweder

$$\alpha := 1 - \frac{1-\lambda}{k(1-\lambda)+2\lambda}, \text{ wenn } i = 1 \text{ oder } i = k, \text{ oder}$$

$$\beta := 1 + \lambda^2 - \frac{(1-\lambda)^3}{k(1-\lambda)+2\lambda}, \text{ wenn } 2 \leq i \leq k-1,$$

und die anderen Elemente der Matrix w_{ij} sind für $i < j$

$$\gamma := -\frac{1-\lambda}{k(1-\lambda)+2\lambda}, \text{ wenn } i = 1 \text{ oder } j = k,$$

$$\delta := -\frac{(1-\lambda)^2}{k(1-\lambda)+2\lambda}, \text{ wenn } i = 1 \text{ und } 3 \leq j \leq k-1 \text{ oder wenn } 2 \leq i \leq k-2 \text{ und } j = k,$$

$$\varepsilon := -\frac{(1-\lambda)^3}{k(1-\lambda)+2\lambda}, \text{ wenn } i \geq 2 \text{ oder } i+2 \leq j \leq k-1,$$

$$\zeta := -\lambda - \frac{(1-\lambda)^2}{k(1-\lambda)+2\lambda}, \text{ wenn } i = 1 \text{ und } j = 2 \text{ oder wenn } i = k-1 \text{ und } j = k,$$

$$\eta := -\lambda - \frac{(1-\lambda)^3}{k(1-\lambda)+2\lambda}, \text{ wenn } 2 \leq i \leq k-2 \text{ und } j = i+1,$$

wobei all diese Einträge noch mit dem Vorfaktor $(\frac{1}{1-\lambda^2})$ versehen werden.

Da W symmetrisch ist, ergeben sich die Einträge für $j > i$ auch aus den zuvor aufgeführten. Die Gestalt der Matrix sei an folgendem Beispiel für $k = 12$ einmal demonstriert.

$$W = \left(\frac{1}{1-\lambda^2} \right) \cdot \begin{pmatrix} \alpha & \zeta & \delta & \gamma \\ \zeta & \beta & \eta & \varepsilon & \delta \\ \delta & \eta & \beta & \eta & \varepsilon & \delta \\ \delta & \varepsilon & \eta & \beta & \eta & \varepsilon & \delta \\ \delta & \varepsilon & \varepsilon & \eta & \beta & \eta & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \delta \\ \delta & \varepsilon & \varepsilon & \varepsilon & \eta & \beta & \eta & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \delta \\ \delta & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \eta & \beta & \eta & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \delta \\ \delta & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \eta & \beta & \eta & \varepsilon & \varepsilon & \varepsilon & \delta \\ \delta & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \eta & \beta & \eta & \varepsilon & \varepsilon & \delta \\ \delta & \varepsilon & \eta & \beta & \eta & \varepsilon & \delta \\ \delta & \varepsilon & \eta & \beta & \zeta & \delta \\ \gamma & \delta & \zeta & \alpha \end{pmatrix}$$

Der Vorfaktor $(\frac{1}{1-\lambda^2})$ der Matrix W ist für die Suche nach der Optimalstelle x^* nicht von Bedeutung, da er nicht nur alle Einträge dieser Matrix, sondern auch alle H-Funktionen in gleichem Maße betrifft (siehe dazu die Definition der H-Funktionen sowie der c_{jj} - und der m -Werte auf Seite 4). Aus diesem Grund wird er bei allen weiteren Überlegungen in dieser Arbeit einfach ignoriert.

Die von diesem Vorfaktor befreiten Einträge der Matrix W werden fortan vor allem mit den neu vergebenen Bezeichnungen $\alpha, \beta, \gamma, \delta, \varepsilon, \zeta$ und η angesprochen und nur noch bei Verweisen auf [Kunert et al., 2010] mit der dort verwendeten Schreibweise w_{ij} . Der Vorteil dieser alternativen Bezeichnungen liegt einfach darin, dass gleiche Zeichen auch immer die gleiche Bedeutung haben, während bei den w_{ij} -Bezeichnungen w_{44} für $k = 4$ beispielsweise einen ganz anderen Eintrag meint als für $k = 5$.

Die Größen $\alpha, \beta, \delta, \varepsilon, \zeta$ und η der Matrix W lassen sich alle auch als Funktion von γ darstellen. So ergibt sich, dass

$$\begin{aligned} \alpha &= 1 + \gamma & \beta &= 1 + \lambda^2 + (1 - \lambda)^2 \gamma & \delta &= (1 - \lambda) \gamma \\ \varepsilon &= (1 - \lambda)^2 \gamma & \zeta &= -\lambda + (1 - \lambda) \gamma & \eta &= -\lambda + (1 - \lambda)^2 \gamma. \end{aligned}$$

Wie man direkt sieht, gilt auch die folgende Gleichung:

$$\delta + \eta = \varepsilon + \zeta.$$

Diese Gleichung wird im Folgenden noch von Bedeutung sein. Ebenfalls im Folgenden noch von Bedeutung sein werden einige Aussagen über die Monotonie der Größen selbst und der aus ihnen zusammengesetzten Ausdrücke, sowie Aussagen über das Spektrum der von ihnen angenommenen Werte.

Behauptung 2.4: $\alpha, \beta, \gamma, \delta$ und ε sind allesamt streng monoton wachsend in λ . Es gilt weiter: γ, δ und $\varepsilon \in [-\frac{1}{k}, 0)$, $\alpha \in [1 - \frac{1}{k}, 1)$, $\beta \in [1 - \frac{1}{k}, 2)$ für $\lambda \in [0, 1)$.

Beweis: Die Ableitungen von $\alpha, \beta, \gamma, \delta$ und ε bezüglich λ lauten:

$$\begin{aligned}\alpha' &:= \frac{\partial \alpha}{\partial \lambda} = \frac{2}{(k(1-\lambda)+2\lambda)^2} \\ \beta' &:= \frac{\partial \beta}{\partial \lambda} = \frac{2\lambda(k(1-\lambda)+2\lambda)^2 + 6\lambda(1-\lambda)^2 + 2k(1-\lambda)^3 + 2(1-\lambda)^3}{(k(1-\lambda)+2\lambda)^2} \\ \gamma' &:= \frac{\partial \gamma}{\partial \lambda} = \frac{2}{(k(1-\lambda)+2\lambda)^2} \\ \delta' &:= \frac{\partial \delta}{\partial \lambda} = \frac{4\lambda(1-\lambda) + k(1-\lambda)^2 + 2(1-\lambda)^2}{(k(1-\lambda)+2\lambda)^2} \\ \varepsilon' &:= \frac{\partial \varepsilon}{\partial \lambda} = \frac{6\lambda(1-\lambda)^2 + 2k(1-\lambda)^3 + 2(1-\lambda)^3}{(k(1-\lambda)+2\lambda)^2}\end{aligned}$$

Da bei allen Brüchen Zähler und Nenner für $\lambda < 1$ ganz offensichtlich positiv sind, ist die strenge Monotonie der Größen gezeigt. Die Minima und Suprema der Größen ergeben sich durch Einsetzen der Grenzen des Wertespektrums $\lambda = 0$ und $\lambda = 1$ in die Formeln. \square

Behauptung 2.5: Der Quotient $\frac{\beta}{\alpha}$ ist für $\lambda \in [0, 1)$ streng monoton wachsend in λ und liegt immer im Intervall $[1, 2)$. Es gilt also $2\alpha > \beta \geq \alpha$.

Beweis:

$$\begin{aligned}\frac{\beta}{\alpha} &= \frac{\frac{k-\lambda k+2\lambda+\lambda^2 k-\lambda^3 k+2\lambda^3-1+3\lambda-3\lambda^2+\lambda^3}{k(1-\lambda)+2\lambda}}{\frac{k-\lambda k+2\lambda-1+\lambda}{k(1-\lambda)+2\lambda}} = \frac{3\lambda^3 - \lambda^3 k + \lambda^2 k - 3\lambda^2 + 5\lambda - \lambda k + k - 1}{k - \lambda k + 3\lambda - 1} \\ &= \frac{\overbrace{(3-k)\lambda^3 + (k-3)\lambda^2 + (5-k)\lambda + (k-1)}^{=Z}}{\underbrace{(k-1) + (3-k)\lambda}_{=N}}\end{aligned}$$

Die Ableitung dieses Ausdrucks bezüglich λ sieht wie folgt aus:

$$\left(\frac{\beta}{\alpha}\right)' = \frac{Z'N - ZN'}{N^2} = \frac{2(k^2-6k+9)\lambda^3 + (-4k^2+18k-18)\lambda^2 + (2k^2-8k+6)\lambda + 2(k-1)}{N^2}$$

Wenn $\frac{\beta}{\alpha}$ streng monoton wachsend in λ ist, muss obige Ableitung für $\lambda \in (0, 1)$ positiv sein. Da es für das Vorzeichen der Ableitung unerheblich ist, wie der Nenner der Ableitung genau aussieht, da er aufgrund des Quadrats ohnehin nie negativ sein kann, kann dieser für die weitere Betrachtung also auch weggelassen werden. Die Ungleichung vereinfacht sich also zu:

$$\left(\frac{\beta}{\alpha}\right)' > 0 \Leftrightarrow \underbrace{2(k-3)^2\lambda^3 - 4\left(k^2 - \frac{9}{2}k + \frac{9}{2}\right)\lambda^2 + 2(k^2 - 4k + 3)\lambda + 2(k-1)}_{(*)} > 0$$

Der in der letzten Ungleichung stehende Ausdruck (*) kann nach k abgeleitet werden. Damit erhält man:

$$(*)' = 2 \underbrace{(\lambda-1)}_{<0} \underbrace{\left(2\lambda^2(k-3) - 2\lambda\left(k - \frac{3}{2}\right) - 1\right)}_{<0, \text{ da } \lambda^2 < \lambda \text{ und } (k-3) \leq (k-\frac{3}{2})} > 0$$

Der Ausdruck (*) ist daher monoton in k wachsend. Da (*) für $k = 3$ einen Wert von 4 annimmt, muss es für $k \geq 4$ dementsprechend noch höhere Werte annehmen. Im von uns be-

trachteten Fall $k \geq 4$ ist (*) also insbesondere immer positiv und folglich die Monotonie des Quotienten $(\frac{\beta}{\alpha})$ in λ gezeigt.

Die Unter- beziehungsweise Obergrenze des Quotienten erhält man, indem man ihn für $\lambda = 0$ und $\lambda \rightarrow 1$ berechnet. Wie man leicht sieht, beträgt der Minimalwert dann 1 und das Supremum 2. Damit ist auch klar, dass $2\alpha > \beta \geq \alpha$ gelten muss. \square

Behauptung 2.6: *Es gelten die folgenden Monotonieaussagen:*

- (a) β^2 ist streng monoton wachsend in λ . (b) ε^2 ist streng monoton fallend in λ .
 (c) $\beta\varepsilon$ ist streng monoton wachsend in λ . (d) $\beta^2\varepsilon^2$ ist streng monoton fallend in λ .

Beweis:

(a) Da - wie in Behauptung 2.4 gezeigt - sowohl β als auch β' immer positiv sind, ist auch $(\beta^2)' = 2\beta\beta' > 0$.

(b) Da nach Behauptung 2.4 aber auch ε für $\lambda \in (0, 1)$ kleiner Null und ε' größer Null ist, gilt für die Ableitung von ε^2 nach λ : $(\varepsilon^2)' = 2\varepsilon\varepsilon' < 0$.

(c) Die Ableitung von $\beta\varepsilon$ nach λ sieht wie folgt aus:

$$\frac{\partial \beta \varepsilon}{\partial \lambda} = \frac{2(1-\lambda)^2}{\underbrace{(k(1-\lambda) + 2\lambda)}_{>0}} \left(k^2(2\lambda^4 - 5\lambda^3 + 5\lambda^2 - 3\lambda + 1) + k(-10\lambda^4 + 19\lambda^3 - 19\lambda^2 + 11\lambda - 1) + 12\lambda^4 - 12\lambda^3 + 10\lambda^2 + 4\lambda - 2 \right). \quad (*)$$

Die zweite Ableitung von (*) nach k sieht für $\lambda \in (0, 1)$ wie folgt aus:

$$\frac{\partial^2 (*)}{\partial^2 k} = 2 \underbrace{(1-\lambda)^2}_{>0} \left(\underbrace{2\lambda^2}_{>0} + \underbrace{1-\lambda}_{>0} \right) > 0$$

Die erste Ableitung von (*) nach k ist dementsprechend monoton wachsend in k . Da wir ohnehin nur Fälle betrachten, für die $k \geq 3$ ist, lässt sich diese Ableitung also durch Einsetzen von $k = 3$ nach unten abschätzen:

$$\frac{\partial (*)}{\partial k} \geq 2\lambda^4 - 11\lambda^3 + 11\lambda^2 - 7\lambda + 5 = 2(1-\lambda)^4 + 3(1-\lambda^3) + \lambda(1-\lambda) > 0.$$

Der Ausdruck (*) ist selbst also monoton wachsend in k . Es gilt daher, dass

$$(*) \underset{k \geq 3}{\geq} \underbrace{-2(\lambda(\lambda-5)-2)}_{<0} > 0.$$

Also ist die Ableitung von $\beta\varepsilon$ nach λ für $k \geq 3$ auf jeden Fall positiv und $\beta\varepsilon$ damit streng monoton wachsend in λ .

(d) Da für $\lambda \in (0, 1)$ $\varepsilon < 0$ und $\beta > 0$ ist, ist das Produkt $\beta\varepsilon$ folglich negativ. Für die Ableitung von $\beta^2\varepsilon^2$ nach λ gilt aber:

$$(\beta^2\varepsilon^2)' = ((\beta\varepsilon)(\beta\varepsilon))' = 2(\beta\varepsilon)'(\beta\varepsilon).$$

$(\beta\varepsilon)'$ ist wegen (b) positiv, $(\beta\varepsilon)$ aber negativ. Ihr Produkt ist dementsprechend wiederum negativ und $(\beta\varepsilon)^2$ deshalb streng monoton fallend in λ . \square

Behauptung 2.7: Wenn der Quotient $\frac{c_2}{c_1}$ kleiner ist als $(k-1)$, so ist ein Ausdruck der Form $c_1\beta + c_2\varepsilon$ für $c_1 > 0$ auf jeden Fall positiv.

Beweis: Betrachte zunächst den Fall, dass auch c_2 größer Null ist. Da β und ε beide monoton wachsend in λ sind (siehe Behauptung 2.4), ist auch der gesamte Ausdruck $c_1\beta + c_2\varepsilon$ monoton in λ wachsend. Sein Minimum nimmt dieser also für $\lambda = 0$ an. Es gilt also:

$$c_1\beta + c_2\varepsilon \geq c_1 \frac{k-1}{k} - c_2 \frac{1}{k} = \frac{1}{k}((k-1)c_1 - c_2) = \underbrace{\frac{c_1}{k}}_{>0} \underbrace{((k-1) - \frac{c_2}{c_1})}_{>0} > 0.$$

Im Fall, dass c_2 kleiner oder gleich Null ist, lässt sich aufgrund der Tatsache, dass auch ε nie positiv ist, sogar noch einfacher zeigen, dass besagter Ausdruck positiv sein muss:

$$\underbrace{\overbrace{c_1}^{>0} \cdot \overbrace{\beta}^{>0}}_{>0} + \underbrace{\overbrace{c_2}^{\leq 0} \cdot \overbrace{\varepsilon}^{\leq 0}}_{\geq 0} > 0. \quad \square$$

3 Reduzierung der Anzahl zu betrachtender Sequenzen

Wie im Abschnitt 2.2 gezeigt wurde, gilt es bei der Suche nach guten oder optimalen approximativen Plänen, die Stelle x^* zu finden, die folgende Anforderung erfüllt:

$$\max_s H_s(x^*) = \min_x \max_s H_s(x).$$

Da es sich wie bereits erwähnt bei jedem Schnitt- und jedem Scheitelpunkt in $[-1, 0]$ um die gesuchte Stelle handeln könnte, ist es klar, dass diese Aufgabe umso schwieriger ist, je mehr Sequenzen betrachtet werden müssen. Aus diesem Grund ist es von großer Bedeutung, die Anzahl der zu betrachtenden Sequenzen schon vorab zu reduzieren. Dazu versucht man, Sequenzen entweder direkt aus dem Kreis der Kandidaten auszuschließen oder sich gleichverhaltende Sequenzen zu Äquivalenzklassen zusammenzufassen, um so nur noch einen Repräsentanten jeder Klasse betrachten zu müssen.

3.1 Die Ansätze von Kunert, Martin und Eccleston

[Kunert et al., 2010] geben in ihrem Artikel drei Regeln an, mit denen sich die Zahl der zu betrachtenden Sequenzen senken lässt. Alle diese drei Regeln basieren auf der grundlegenden Idee, dass eine Sequenz, die überall einer anderen unterlegen ist, bei der Suche nach der Minimalstelle der Maximalwerte keine Rolle spielen kann und deshalb nicht weiter betrachtet werden muss. Gilt für zwei Sequenzen s_1 und s_2 nämlich, dass $m_1(s_1) \geq m_1(s_2)$ und $m_2(s_1) \geq m_2(s_2)$ ist, so folgt daraus direkt, dass auch $H_{s_1}(x) \geq H_{s_2}(x)$ ist und zwar für alle Werte von x . Die drei Regeln sind nicht speziell für die Situation von AR(1)-Fehlern formuliert und lauten in ihrer allgemeingültigen Formulierung:

- Lässt sich eine Sequenz s in eine andere Sequenz t überführen, indem man einfach die Reihenfolge dieser Sequenz umdreht und den Behandlungen neue Bezeichnungen gibt, so haben beide Sequenzen dieselben m_1 - und m_2 -Werte und eine von ihnen muss folglich nicht weiter betrachtet werden.
- Wenn alle w_{ij} für $i \neq j$ kleiner oder gleich Null sind, dann müssen nur die Sequenzen betrachtet werden, die keine Behandlung mehr als einmal einsetzen.
- Wenn zwei Sequenzen s und t sich in der Art unterscheiden, dass an allen Stellen j_1, \dots, j_z , an denen in Sequenz s eine Kontrolle steht, in Sequenz t eine Behandlung eingesetzt wird und dass an allen Stellen, an denen in Sequenz s eine Behandlung steht, in Sequenz t eine Kontrolle eingesetzt wird, und wenn weiter gilt, dass $\sum_{r=1}^z w_{j_r, j_r} \leq \frac{1}{2} \text{tr}(W)$ ist und keine Behandlung mehr als einmal eingesetzt wird, dann muss Sequenz t nicht weiter beachtet werden.

[Kunert et al., 2010] betonen selbst, dass die Bedingung für die zweite der drei Regeln bei AR(1)-Fehlern für $0 \leq \lambda < 1$ immer erfüllt ist. Sequenzen, die eine Behandlung mehr als einmal ver-

wenden, spielen für die Untersuchungen in dieser Arbeit also keine Rolle und können außen vor gelassen werden. Wenn $0 \leq \lambda < 1$ vorausgesetzt werden kann, vereinfacht sich aber auch die dritte der drei Regeln noch:

Behauptung 3.1: *Wenn $0 \leq \lambda < 1$ ist, so müssen keine Sequenzen betrachtet werden, die...*

- ... die Kontrolle auf mehr als der Hälfte der Positionen im Block einsetzen.
- ... die Kontrolle in genau der Hälfte der Positionen im Block einsetzen, nicht jedoch auf einer der Randpositionen.

Beweis: Die Diagonalelemente w_{ii} von W nehmen nur zwei unterschiedliche Werte an: w_{11} und w_{kk} sind immer gleich α , die übrigen w_{jj} für $j = 2, \dots, k-1$ gleich β . Die Spur von W ist also $\sum_{i=1}^k w_{ii} = 2\alpha + (k-2)\beta$. Wegen den Behauptungen 2.4 und 2.5 gilt aber: $\beta \geq \alpha > 0$ und $\alpha > \frac{1}{2}\beta$ für $0 \leq \lambda < 1$.

Damit eine Sequenz überhaupt von Bedeutung sein kann, muss nach *Proposition 3.2* aus [Kunert et al., 2010] gelten, dass $\sum_{r=1}^z w_{j_r, j_r} \leq \frac{1}{2} \text{tr}(W)$ ist, wobei $j_i, i = 1, \dots, z$ die Positionen bezeichnet, an denen in ihr Kontrollen eingesetzt werden. Da $w_{ii} \in \alpha, \beta$ und α und β beide positiv sind, ist offensichtlich, dass diese Bedingung umso leichter erfüllt werden kann, je weniger Kontrollen eingesetzt werden.

Sei jetzt k ungerade und s eine Sequenz, für die z , die Anzahl eingesetzter Kontrollen, größer als $\frac{k}{2}$ ist. Wegen $\beta \geq \alpha$ ist dann für s die Summe $\sum_{r=1}^z w_{j_r, j_r}$ auf jeden Fall $\geq 2\alpha + \frac{k-3}{2}\beta$. Die Sequenz t , die überall da, wo in s Kontrollen stehen, Behandlungen einsetzt und ansonsten überall Kontrollen, erreicht in der Summe aber nur einen Wert von $\frac{k-1}{2}\beta$. Da wegen $\alpha > \frac{1}{2}\beta$ aber immer gilt, dass $\frac{k-1}{2}\beta < \frac{1}{2} \text{tr}(W) = \alpha + \frac{k-2}{2}\beta$, genügt es folglich nach *Proposition 3.2* die Sequenz t zu betrachten und die Sequenz s kann ignoriert werden. Analog folgt auch für gerades k , dass nicht mehr als die Hälfte der Positionen mit Kontrollen belegt werden dürfen.

Soll eine Sequenz s bei geradem k genau die Hälfte der Positionen mit Kontrollen belegt haben, aber keine Kontrolle auf einer der Randpositionen stehen, so ergibt sich eine Summe von $\sum_{r=1}^z w_{j_r, j_r} = \frac{k}{2}\beta$. Wegen $\beta \geq \alpha$ ist dies aber nie kleiner als $\frac{1}{2} \text{tr}(W) = \alpha + \frac{k-2}{2}\beta$, und die Sequenz ist ihrer „Gegenspielerin“ t , die Kontrollen da einsetzt, wo s Behandlungen einsetzt und umgekehrt, für $b > a$ (entspricht $\lambda > 0$) sogar unterlegen und für $b = a$ (entspricht $\lambda = 0$) auch nicht überlegen. Es macht demnach keinen Sinn, diese Sequenz weiter zu betrachten. \square

Diese Regeln sind dazu in der Lage, die Zahl der in Frage kommenden Sequenzen beträchtlich zu reduzieren. Wie Tabelle 1 zeigt, bleiben aber auch nach ihrer Anwendung für größere Werte von k immer noch viel zu viele Sequenzen zur Betrachtung übrig, als dass die Fragestellung nun einfach angegangen werden könnte. Die Zahl der zu betrachtenden Sequenzen ist bislang immer noch größer als 2^{k-2} , weitere Methoden zur Reduzierung ihrer Anzahl scheinen somit zwingend erforderlich zu sein.

Tabelle 1: Anzahl verbleibender Sequenzen nach Anwendung der Regeln aus [Kunert et al., 2010]

k	Anzahl Sequenzen	k	Anzahl Sequenzen
3	3	11	528
4	6	12	1166
5	10	13	2080
6	21	14	4590
7	36	15	8256
8	78	16	18164
9	136	17	32896
10	299	18	72099

3.2 Weitere Ansätze zur Reduzierung

[Kunert et al., 2010] haben drei allgemeine Regeln angegeben, mit denen Sequenzen direkt ausgeschlossen werden können, ohne dabei von der speziellen Korrelationsstruktur der Fehlerterme Gebrauch zu machen. Die weiteren Reduzierungsmaßnahmen, die in diesem Abschnitt vorgestellt werden, setzen im Gegensatz dazu voll auf das Wissen um die konkrete Korrelationsstruktur: In einem ersten Schritt werden dabei Sequenzen zu Äquivalenzklassen zusammengeführt, indem man sich das spezielle Aussehen der Matrix W für den AR(1)-Prozess zu Nutze macht.

Behauptung 3.2: *Im konkreten Fall von AR(1)-Fehlertermen mit nicht-negativem Korrelationsparameter λ lassen sich die $m_1(s)$ - und $m_2(s)$ -Werte der verbleibenden Sequenzen wie folgt darstellen:*

$$m_2(s) = \frac{1}{v} (p + q - 2n\lambda + q\lambda^2 + \gamma(q(1 - \lambda) + p)^2)$$

$$m_1(s) = \frac{1}{v-1} ((2 - p)\alpha + (k - 2 - q)\beta - m_2(s))$$

Dabei bezeichne p die Anzahl der Kontrollen, die auf den Randpositionen (d.h. direkt an allererster oder an allerletzter Stelle) eingesetzt werden, q die Anzahl der Kontrollen auf den übrigen Innenpositionen der Sequenz, und n gibt an, wie oft in der Sequenz Kontrollen direkt aufeinander folgen. Aus dieser Definition folgt unmittelbar, dass $p \in \{0, 1, 2\}$, $q \in \{0, \dots, k - 2\}$ und $n \in \{0, \dots, k - 1\}$ sein muss.

Beweis: Die m_1 - und m_2 -Werte waren wie folgt definiert:

$$m_2(s) = \frac{c_{00}}{v} \quad m_1(s) = \frac{1}{v-1} \left(\sum_{j=1}^v c_{jj}(s) - m_2(s) \right).$$

Dabei ist $c_{jj}(s) = \sum_{r=1}^k \sum_{u=1}^k t_r^{(j)}(s) w_{ru} t_u^{(j)}(s)$ mit

$$t_u^{(j)}(s) = \begin{cases} 1 & : \text{Behandlung } j \text{ steht an Stelle } u \text{ von Sequenz } s \\ 0 & : \text{sonst} \end{cases}$$

c_{00} , der Zähler von $m_2(s)$, lässt sich auf Grund der Symmetrie von W umschreiben zu:

$$\begin{aligned} c_{00}(s) &= \sum_{r=1}^k \sum_{u=1}^k t_r^{(0)}(s) w_{ru} t_u^{(0)}(s) \\ &= \sum_{i=1}^k t_i^{(0)}(s) w_{ii} t_i^{(0)}(s) + \sum_{r=1}^k \sum_{u \neq r, u=1}^k t_r^{(0)}(s) w_{ru} t_u^{(0)}(s) \\ &= \sum_{i=1}^k t_i^{(0)}(s)^2 w_{ii} + \sum_{r=1}^k \sum_{u \neq r, u=1}^k t_r^{(0)}(s) w_{ru} t_u^{(0)}(s). \end{aligned}$$

Der erste Summand in der obigen Darstellung, $\sum_{i=1}^k t_i^{(0)}(s)rw_{ii}$, entspricht unter Verwendung der zuvor vorgestellten Bezeichnungen $p\alpha + q\beta$.

Den zweiten Summanden, $\sum_{r=1}^k \sum_{u \neq r, u=1}^k t_r^{(0)}(s)w_{ru}t_u^{(0)}(s)$, umzuschreiben, ist etwas aufwändiger: Man halte sich zunächst einmal vor Augen, dass wenn insgesamt $p + q$ Kontrollen in einer Sequenz eingesetzt werden, es in der Doppelsumme $(p + q)^2 - p - q$ Summanden geben muss, für die nicht entweder schon $t_r^{(0)}(s)$ oder $t_u^{(0)}(s)$ gleich Null ist. Dies ergibt sich aus einfachen kombinatorischen Überlegungen.

Betrachtet man nun den Aufbau von W und überlegt sich, wann welche der Größen $\alpha, \beta, \gamma, \delta, \varepsilon, \zeta$ oder η zur Summe gehören, so stellt man folgendes fest:

- Die Größe δ kann überhaupt nur dann unter den Summanden sein, wenn p von Null verschieden ist. Ist $p = 1$, so trägt jede Kontrolle in der Mitte 2δ zur Gesamtsumme bei, ist $p = 2$, so sind es sogar 4δ . Allgemeiner ausgedrückt: Unter den $(p + q)^2 - p - q$ Summanden sind immer $(2pq)$ gleich δ .
- Die Größe γ ist genau dann unter den Summanden, wenn $p = 2$ ist, und dann sind es direkt auch zwei Summanden, die diesen Wert annehmen. Mathematischer formuliert: Unter den Summanden sind immer $(p^2 - p)$ gleich γ .
- Die Größen ζ und η können sich nur unter den Summanden befinden, wenn in der Sequenz zwei Kontrollen direkt aufeinander folgen. Wenn die 1. und 2. Stelle oder die $(k - 1)$. und k . Stelle einer Sequenz mit einer Kontrolle besetzt ist, fließen jeweils 2ζ in die Summe ein, sind zwei andere direkt aufeinanderfolgende Stellen mit Kontrollen besetzt, werden 2η dazu addiert. Auf welchen Stellen aber nun genau benachbarte Kontrollen stehen, ob nun am Rand der Sequenz oder in ihrem Mittelteil, ist letztlich aber irrelevant. Um dies zu sehen, stelle man sich eine Sequenz vor, bei der am Rand zwei Kontrollen direkt aufeinander folgen (o.B.d.A. seien dies die Positionen 1 und 2) und bei der noch mindestens eine weitere Nichtrandposition j mit einer Kontrolle besetzt ist. Nun entferne man die Kontrolle auf Position 2 und füge sie dafür auf einer noch nicht von einer Kontrolle besetzten Position direkt neben j hinzu (o.B.d.A. sei dies Position $j + 1$). Dieses Vorgehen hat aber keine Auswirkungen auf c_{00} . Für $r = 1$ und $u = 2$ trägt obige Doppelsumme statt vormals 2ζ jetzt zwar genau Null zum Gesamtergebnis bei, für $r = j$ und $u = j + 1$ wird dafür aber aus einem Beitrag von Null ein Beitrag von 2δ . Da zudem die 2ε verschwinden, die vor der Umsetzung für $r = 2$ und $u = j$ dazu gekommen wären, und sich dafür für $r = j$ und $u = j + 1$ 2η dazuaddieren, ergibt sich als Gesamtveränderung

$$\Delta = -2\zeta + 2\delta - 2\varepsilon + 2\eta.$$

Wegen der aber schon zuvor in Abschnitt 2.3 auf Seite 9 präsentierten Gleichung $\delta + \eta = \varepsilon + \zeta$ ist dies nichts anderes als Null, d.h. wir können immer so tun, dass, wenn es aufeinanderfolgende Kontrollen gibt, diese sich im Mittelteil befinden, und für jedes der n Paare direkt aufeinanderfolgender Kontrollen 2η zur Gesamtsumme hinzuzählen.

- Da $(2pq)$ der Summanden gleich δ sind, $(p^2 - p)$ gleich γ und $2n$ gleich η , es insgesamt aber nur $(p + q)^2 - p - q$ Summanden gibt und ε zudem die letzte zu betrachtende Größe darstellt, ergibt sich zwingend, dass es sich bei allen restlichen $(p + q)^2 - p - q - (2pq) - (p^2 - p) - 2n = q^2 - q - 2n$ Summanden um ε handeln muss.

Man stellt also fest: c_{00} lässt sich darstellen als

$$\begin{aligned}
 c_{00} &= \overbrace{p^\alpha + q^\beta}^{\text{1. Term}} + \overbrace{(2pq)\delta + 2n\eta + (p^2 - p)\gamma + (q^2 - q - 2n)\varepsilon}^{\text{2. Term}} \\
 &= \dots = p + q - 2n\lambda + q\lambda^2 + \gamma(q(1 - \lambda) + p)^2.
 \end{aligned}$$

Da $m_2(s)$ definiert war als $\frac{c_{00}}{v}$, folgt somit der erste Teil der Behauptung.

Um $m_1(s)$ zu bestimmen, ist es nötig, die einzelnen c_{jj} für $j > 0$ zu bestimmen. Die Bestimmung dieser c_{jj} ist aber wesentlich einfacher als die Bestimmung von c_{00} : Alle Nebendiagonalelemente der Matrix W sind in der hier ausschließlich untersuchten Situation $0 \leq \lambda < 1$ auf jeden Fall nicht-positiv (siehe dazu 2.4 und die Definition der Größen), somit müssen wie schon erwähnt nach [Kunert et al., 2010] auch nur Sequenzen betrachtet werden, die jede beliebige Behandlung höchstens einmal einsetzen. Bei solchen Sequenzen können die einzelnen c_{jj} aber nur entweder einen Wert von α oder einen Wert von β annehmen - ersteren, wenn die jeweilige Behandlung am Rand der Sequenz eingesetzt wird, letzteren, wenn sie auf einer der $(k - 2)$ Innenpositionen steht.

Mit p und q wurde die Anzahl der Kontrollen auf den Rand- bzw. Innenpositionen bezeichnet. Da es in jeder Sequenz aber genau 2 Randpositionen und $(k - 2)$ Innenpositionen gibt, hat $\sum_{j=1}^v c_{jj}(s)$ die folgende Gestalt:

$$\sum_{j=1}^v c_{jj}(s) = (2 - p)\alpha + (k - 2 - q)\beta = (2 - p)(1 + c) + (k - 2 - q)(1 + \lambda^2 + (1 - \lambda)^2 c)$$

Aus diesem Resultat und der Darstellung von $m_2(s)$ folgt nun der zweite Teil der Behauptung. \square

Anhand dieser neuen Darstellungsform der m_1 - und m_2 -Werte ist nun direkt ersichtlich, dass beispielsweise die Sequenzen $s_1 = [0\ 0\ 1\ 2\ 0\ 3]$ und $s_2 = [1\ 0\ 0\ 2\ 3\ 0]$ die gleiche Funktion $H_s(x)$ liefern, da sie beide eine Kontrolle auf einem Randfeld und zwei Kontrollen im Mittelteil der Sequenz einsetzen und dabei jeweils genau an einer Stelle Kontrollen direkt nebeneinander stehen. Die beiden Sequenzen gehören also zur selben Klasse von Sequenzen. Für weitere Überlegungen würde es daher genügen, nur noch einen Repräsentanten dieser Klasse zu betrachten.

Definition 3.3: Mit $[p = p_*, q = q_*, n = n_*]$ wird fortan die Klasse von Sequenzen bezeichnet, die p_* Kontrollen auf den Randfeldern und q_* Kontrollen im Mittelteil einsetzen und dabei insgesamt n_* -mal Kontrollen auf direkt benachbarten Positionen verwendet.

Da sich also alle Sequenzen, die das gleiche Parametertupel (p, q, n) aufweisen, jeweils zu einer Klasse zusammenfassen lassen und es wegen der verhältnismäßig einfachen Form der Parametrisierung sogar nur vergleichsweise wenige Klassen gibt, trägt die in Behauptung 3.2 gefundene alternative Darstellung bereits enorm zur Reduzierung der Anzahl zu betrachtender Sequenzen bei.

Wie im Folgenden gezeigt wird, lassen sich anhand der neuen Darstellung aber auch ganze Sequenzklassen recht einfach aussortieren:

Behauptung 3.4: *Wenn die Zahl der in einer Sequenz insgesamt eingesetzten Kontrollen kleiner ist als $\frac{k}{2}$, so müssen nur Sequenzen beachtet werden, die keine Kontrollen auf direkt benachbarten Positionen einsetzen.*

Beweis: Betrachte zwei Sequenzen s_1 und s_2 , die beide jeweils p -mal am Rand und q -mal im Innenteil einsetzen. Bei der Sequenz s_1 wird zwischen zwei Kontrollen jedoch immer mindestens eine Behandlung eingesetzt, bei der Sequenz s_2 gebe es genau m Paare direkt aufeinanderfolgender Kontrollen ($m > 0$). Beachte, dass die Konstruktion solcher Sequenzen immer möglich ist, da in der Behauptung die Zahl der insgesamt verwendeten Kontrollen als kleiner $\frac{k}{2}$ festgelegt wurde.

m_2 und m_1 haben für diese Sequenzen mit p Kontrollen auf den Außenfeldern, q Kontrollen im Mittelteil und 0 bzw. m Paaren direkt benachbarter Kontrollen das folgende Aussehen:

$$\begin{aligned} m_2(s_1) &= \frac{1}{v}(p + q + q\lambda^2 + \gamma(q(1 - \lambda) + p)^2) \\ m_1(s_1) &= \frac{1}{v-1}((2 - p)\alpha + (k - 2 - q)\beta - m_2(s_1)) \\ m_2(s_2) &= \frac{1}{v}(p + q - 2m\lambda + q\lambda^2 + \gamma(q(1 - \lambda) + p)^2) \\ m_1(s_2) &= \frac{1}{v-1}((2 - p)\alpha + (k - 2 - q)\beta - m_2(s_2)) \end{aligned}$$

Definiere nun $\Delta m_1 := m_1(s_1) - m_1(s_2)$ und $\Delta m_2 := m_2(s_1) - m_2(s_2)$ und berechne

$$\Delta H(x) = H_1(x) - H_2(x) = (1 + x)^2 \Delta m_1 + (v - 1)x^2 \Delta m_2$$

Δm_2 ist in dieser Situation $\frac{2m\lambda}{v}$ und Δm_1 ist $-\frac{2m\lambda}{v(v-1)}$.

Somit gilt für $\Delta H(x)$:

$$\Delta H(x) = (1 + x)^2 \left(-\frac{2m\lambda}{v(v-1)}\right) + (v - 1)x^2 \frac{2m\lambda}{v} = \frac{2m\lambda}{v(v-1)} \underbrace{\left(-(1 + x)^2 + (v - 1)^2 x^2\right)}_{(*)}$$

Dieser Ausdruck hat bezüglich x die folgenden beiden Nullstellen:

$$x_{N_1} = -\frac{1}{v} \text{ und } x_{N_2} = \frac{1}{v-2}.$$

An der Stelle $x = 0$, die zwischen diesen beiden Nullstellen liegt, nimmt $\Delta H(x)$ den Wert -1 an. Da es nur zwei Nullstellen geben kann und die Funktion stetig ist, bedeutet dies, dass nur in dem Intervall $[-\frac{1}{v}, \frac{1}{v-2}]$ die Sequenz s_2 , die Kontrollen in direkter Folge einsetzt, höhere H-Werte als die Sequenz s_1 erreicht, die dies nicht tut.

Betrachte man zum Vergleich jetzt einmal eine Sequenz s_0 , die komplett auf den Einsatz von Kontrollen verzichtet.

Diese Sequenz hat die folgenden m_2 - und m_1 -Werte:

$$\begin{aligned} m_2(s_0) &= 0 \\ m_1(s_0) &= \frac{1}{v-1}(2\alpha + (k-2)\beta) \end{aligned}$$

Bezeichne $\widetilde{\Delta m_1} := m_1(s_0) - m_1(s_2)$ und $\widetilde{\Delta m_2} := m_2(s_0) - m_2(s_2)$ nun die Differenzen der m_1 - und m_2 -Werte zwischen s_0 und s_2 . Berechnet man damit weiter

$$\widetilde{\Delta H(x)} = (1+x)^2 \widetilde{\Delta m_1} + (v-1)x^2 \widetilde{\Delta m_2},$$

so gelangt man zu folgendem Ergebnis:

$$\begin{aligned} \widetilde{\Delta H(x)} &= (1+x)^2 \frac{1}{v-1}(p\alpha + q\beta + m_2(s_2)) - x^2(v-1)m_2(s_2) \\ &= (1+x)^2 \frac{1}{v-1}(p\alpha + q\beta) + \frac{m_2(s_2)}{v-1}((1+x)^2 - (v-1)^2x^2). \end{aligned}$$

Der erste Summand ist größer Null, da sowohl α als auch β nicht kleiner werden können als $(1 - \frac{1}{k})$ (siehe Behauptung 2.4) und auch alle weiteren Größen echt positiv sind. Da für alle Sequenzen der m_2 -Wert aber größer oder gleich Null ist (siehe Behauptung 2.1), ist auch der Vorfaktor des zweiten Summanden nicht-negativ. Der Term in Klammern beim zweiten Summanden unterscheidet sich nur bezüglich des Vorzeichens vom Ausdruck (*), dementsprechend hat er dieselben Nullstellen $x_{N_1} = -\frac{1}{v}$ und $x_{N_2} = \frac{1}{v-2}$, nimmt aber zwischen diesen beiden Nullstellen aufgrund des veränderten Vorzeichens positive statt negative Werte an. Wenn $x \in [-\frac{1}{v}, \frac{1}{v-2}]$, so ist der gesamte Ausdruck dementsprechend echt positiv, und die Sequenz s_0 in diesem Bereich der Sequenz s_2 also bezüglich ihres H-Werts überlegen.

Da nun aber im Bereich $x \in [-\infty, -\frac{1}{v}]$ und im Bereich $x \in [\frac{1}{v-2}, \infty]$ die Sequenz s_1 höhere H-Werte liefert als die Sequenz s_2 , in dem dazwischen liegenden Bereich $x \in [-\frac{1}{v}, \frac{1}{v-2}]$ aber wiederum die Sequenz s_0 der Sequenz s_2 überlegen ist, ist klar, dass s_2 nie eine Rolle spielen kann und für $\lambda \geq 0$ die Verwendung von Kontrollen in direkter Reihe also zu vermeiden ist. \square

Behauptung 3.5: *Wenn die Zahl der in einer Sequenz insgesamt eingesetzten Kontrollen kleiner ist als $\frac{k}{2}$, so müssen nur Sequenzen beachtet werden, die keine Kontrollen auf Randfeldern einsetzen.*

Beweis: Betrachte die folgenden vier Sequenzen, die alle so gewählt seien, dass die Kontrollen in ihnen an keiner Stelle in direkter Folge eingesetzt werden (das Verwenden mehrerer Kontrollen nebeneinander ist ja nach Behauptung 3.4 schon zu vermeiden). Beachte wiederum, dass dadurch, dass in der Behauptung die Zahl der insgesamt verwendeten Kontrollen als kleiner $\frac{k}{2}$ festgelegt wurde, es immer möglich ist, derartige Sequenzen zu konstruieren, ohne dass Kontrollen auf direkten Nachbarfeldern stehen müssen. Beachte weiter, dass die Sequenzen s_{2*} und s_{3*} natürlich überhaupt nur dann existieren, wenn $q \geq 1$ bzw. ≥ 2 ist.

- Die Sequenz s_0 , die überhaupt keine Kontrolle einsetzt.

$$\begin{aligned} m_2(s_0) &= 0 \\ m_1(s_0) &= \frac{1}{v-1}(2\alpha + (k-2)\beta) \end{aligned}$$

- Die Sequenz s_{1*} , die keine Kontrolle auf den beiden Randfeldern und q Kontrollen im Mittelteil einsetzt.

$$m_2(s_{1*}) = \frac{1}{v}(q + q\lambda^2 + \gamma(q(1-\lambda))^2)$$

$$m_1(s_{1*}) = \frac{1}{v-1}(2\alpha + (k-2-q)\beta - m_2(s_{1*}))$$

- Die Sequenz s_{2*} , die eine Kontrolle auf einem der beiden Randfelder und $q-1$ Kontrollen im Mittelteil einsetzt.

$$m_2(s_{2*}) = \frac{1}{v}(q + q\lambda^2 - \lambda^2 + \gamma(q(1-\lambda) + \lambda)^2)$$

$$m_1(s_{2*}) = \frac{1}{v-1}(\alpha + (k-1-q)\beta - m_2(s_{2*}))$$

- Die Sequenz s_{3*} , die auf beiden Randfeldern die Kontrolle und $q-2$ Kontrollen im Mittelteil einsetzt.

$$m_2(s_{3*}) = \frac{1}{v}(q + q\lambda^2 - 2\lambda^2 + \gamma(q(1-\lambda) + 2\lambda)^2)$$

$$m_1(s_{3*}) = \frac{1}{v-1}((k-q)\beta - m_2(s_{3*}))$$

Die zu den vier Sequenzen gehörenden Polynome sollen nun an den folgenden drei Punkten bezüglich ihrer Lage verglichen werden:

$$x_0 = -1 \quad x_1 = -\frac{\sqrt{v+1}}{v-1+\sqrt{v+1}} \quad x_2 = 0$$

Die Funktionswerte an den drei Stellen sehen in Abhängigkeit von den zur jeweiligen Sequenz s gehörenden m_1 - bzw. m_2 -Werte wie folgt aus:

$$H_s(x_0) = (v-1)m_2(s)$$

$$H_s(x_1) = \frac{1}{(v-1+\sqrt{v+1})^2}((v-1)^2 m_1(s) + (v+1)(v-1)m_2(s))$$

$$H_s(x_2) = m_1(s)$$

Der Vergleich der Funktionswerte der vier Sequenzen an den drei genannten Punkten liefert folgende Ergebnisse:

- An der Stelle $x_0 = -1$ gilt: $H_{s_{1*}}(-1) \geq H_{s_{2*}}(-1) \geq H_{s_{3*}}(-1) \geq H_{s_0}(-1)$.
- An der Stelle $x_1 = -\frac{\sqrt{v+1}}{v-1+\sqrt{v+1}}$ gilt: $H_{s_0}(x_1) \geq H_{s_{1*}}(x_1) \geq H_{s_{2*}}(x_1) \geq H_{s_{3*}}(x_1)$.
- An der Stelle $x_2 = 0$ gilt: $H_{s_{1*}}(0) \leq H_{s_{2*}}(0) \leq H_{s_{3*}}(0) \leq H_{s_0}(0)$.

Dies sieht man wie folgt: Die Differenz der Polynome von s_{1*} und s_{2*} an der Stelle $x_0 = -1$ ist nichts anderes als

$$H_{s_{1*}}(-1) - H_{s_{2*}}(-1) = \frac{v-1}{v} \left(\underbrace{\lambda^2}_{\geq 0} - \underbrace{2\gamma}_{< 0} \underbrace{q\lambda(1-\lambda)}_{\geq 0} - \underbrace{\lambda^2 \gamma}_{< 0} \right) \geq 0.$$

Auch

$$H_{s_{2*}}(-1) - H_{s_{3*}}(-1) = \frac{v-1}{v} \left(\underbrace{\lambda^2}_{\geq 0} - \underbrace{2\gamma}_{< 0} \underbrace{q\lambda(1-\lambda)}_{\geq 0} - \underbrace{3\lambda^2 \gamma}_{< 0} \right) \geq 0.$$

ist größer oder gleich Null, bei

$$H_{s_{3*}}(-1) - H_{s_0}(-1) = m_2(s_{3*})$$

ergibt sich dies schon allein aus der Tatsache, dass m_2 -Werte nie negativ sein können. Somit ist die Anordnung an Punkt x_0 gezeigt.

Wir verfahren analog an der Stelle $x_1 = -\frac{\sqrt{v+1}}{v-1+\sqrt{v+1}}$ und betrachten die mit dem positiven Faktor $(v-1+\sqrt{v+1})^2$ multiplizierten Differenzen der Polynome:

$$\begin{aligned} (v-1+\sqrt{v+1})^2(H_{s_0}(x_1) - H_{s_{1*}}(x_1)) &= (v-1)(q\beta - q - q\lambda^2 - q^2\varepsilon) \\ &= (v-1) \underbrace{\cdot \varepsilon}_{<0} \cdot \underbrace{(q - q^2)}_{\leq 0} \geq 0 \end{aligned}$$

$$\begin{aligned} (v-1+\sqrt{v+1})^2(H_{s_{1*}}(x_1) - H_{s_{2*}}(x_1)) &= (v-1)(\alpha - \beta + \lambda^2 - 2q(1-\lambda)\lambda\gamma - \lambda^2\gamma) \\ &= 2(v-1)\gamma(\lambda - \lambda^2 - q(1-\lambda)\lambda) \\ &= 2(v-1) \underbrace{\cdot \gamma}_{<0} \cdot \underbrace{(\lambda - 1)}_{<0} \lambda(q-1) \geq 0 \text{ für } q \geq 1 \\ &\quad \underbrace{\hspace{10em}}_{\geq 0} \end{aligned}$$

$$\begin{aligned} (v-1+\sqrt{v+1})^2(H_{s_{2*}}(x_1) - H_{s_{3*}}(x_1)) &= (v-1)(\alpha - \beta + \lambda^2 - 2q(1-\lambda)\lambda\gamma - 3\lambda^2\gamma) \\ &= 2(v-1)\gamma(\lambda - 2\lambda^2 - q(1-\lambda)\lambda) \\ &= -2(v-1)\lambda \cdot \underbrace{\gamma}_{<0} \cdot (1 + (1-\lambda)(q-2)) \geq 0 \text{ für } q \geq 2 \\ &\quad \underbrace{\hspace{10em}}_{\geq 0} \end{aligned}$$

Da die Sequenzen s_{2*} und s_{3*} überhaupt nur existieren, wenn $q \geq 1$ bzw. ≥ 2 ist, können diese Untergrenzen von q für die letzten beiden Polynomdifferenzen vorausgesetzt werden. Somit ist dann die Anordnung an Punkt x_1 gezeigt.

An der Stelle $x_2 = 0$ ergibt sich

$$\begin{aligned} H_{s_0}(0) - H_{s_{3*}}(0) &= \frac{1}{v-1}(2\alpha + (q-2)\beta + m_2(s_{3*})) \\ &\geq \frac{1}{v-1} \underbrace{(q\alpha)}_{>0} + \underbrace{m_2(s_{3*})}_{\geq 0} > 0 \quad (\text{wegen } \beta \geq \alpha, \text{ siehe auch Behauptung 2.5}) \end{aligned}$$

$$\begin{aligned} H_{s_{3*}}(0) - H_{s_{2*}}(0) &= \frac{1}{v-1}(\beta - \alpha + m_2(s_{2*}) - m_2(s_{3*})) \\ &= \frac{1}{v-1}(\lambda^2 + \underbrace{(\lambda-2)\lambda}_{\geq 0} \cdot \underbrace{\gamma}_{<0}) + \frac{1}{v}(\lambda^2 - 2 \cdot \underbrace{\gamma}_{<0} \cdot q\lambda(1-\lambda) - 3\lambda^2 \cdot \underbrace{\gamma}_{<0}) \geq 0. \end{aligned}$$

$$\begin{aligned} H_{s_{2*}}(0) - H_{s_{1*}}(0) &= \frac{1}{v-1}(\beta - \alpha + m_2(s_{1*}) - m_2(s_{2*})) \\ &= \frac{1}{v-1}(\lambda^2 + \underbrace{(\lambda-2)\lambda}_{\geq 0} \cdot \underbrace{\gamma}_{<0}) + \frac{1}{v}(\lambda^2 - 2 \cdot \underbrace{\gamma}_{<0} \cdot q\lambda(1-\lambda) - \lambda^2 \cdot \underbrace{\gamma}_{<0}) \geq 0 \end{aligned}$$

Damit ist die Anordnung an Punkt x_2 gezeigt.

Aus diesen Anordnungen folgt aber unmittelbar, dass in der der Behauptung zugrundeliegenden Situation Sequenzen, die Kontrollen auch am Rand einsetzen, nicht weiter betrachtet werden müssen. Dies sieht man wie folgt: Da es sich bei all diesen Polynomen um nach oben geöffnete Parabeln handelt, können sich jeweils zwei dieser Polynome höchstens an zwei Stellen schneiden. An den Stellen $x_0 = -1$ und $x_1 = -\frac{\sqrt{v+1}}{v-1+\sqrt{v+1}}$ sind die H-Werte

von Sequenz s_{1*} nicht kleiner als die Funktionswerte zu den Sequenzen s_{2*} und s_{3*} , an der Stelle $x_2 = 0$ sind sie nicht größer. Daraus lässt sich folgern, dass auch überall in dem Intervall $[-1, -\frac{\sqrt{v+1}}{v-1+\sqrt{v+1}}]$ die zur Sequenz s_{1*} gehörende Funktion bessere oder jedenfalls nicht schlechtere Werte erzielt als die zu s_{2*} und s_{3*} gehörenden Funktionen. Gäbe es nämlich in diesem Intervall Stellen, an denen dies nicht der Fall wäre, so müssten sich $H_{s_{1*}}$ und $H_{s_{2*}}$ bzw. $H_{s_{3*}}$ in diesem Intervall an genau zwei Stellen schneiden, und ab dem zweiten dieser Schnittpunkte müsste die Funktion $H_{s_{1*}}$ sogar echt größere Werte liefern als jede der beiden anderen Funktionen. Da aber schon festgestellt wurde, dass an der Stelle $x_2 = 0$ die Funktion $H_{s_{1*}}$ nicht größer sein kann als die beiden anderen Funktionen, scheidet diese Möglichkeit aus.

Analog lässt sich zeigen, dass es im Intervall $[-\frac{\sqrt{v+1}}{v-1+\sqrt{v+1}}, 0]$ auch keine Stelle geben kann, an denen $H_{s_{2*}}$ oder $H_{s_{3*}}$ höhere Werte liefert als H_{s_0} . Die Funktionen $H_{s_{2*}}$ und $H_{s_{3*}}$ werden somit aber im gesamten relevanten Bereich $[-1, 0]$ von anderen Funktionen dominiert und deshalb können die entsprechenden Sequenzen bei den weiteren Betrachtungen außen vor gelassen werden. \square

Von den Behauptungen 3.4 und 3.5 nicht erfasst ist die Situation, in der genau $\frac{k}{2}$ Kontrollen in einer Sequenz eingesetzt werden. Diese kann natürlich überhaupt nur für gerade Werte von k eintreten. Der Grund dafür, dass die beiden Behauptungen diesen Fall unberücksichtigt ließen, ist, dass zwei der in den Beweisen genutzten Sequenzklassen, s_1 in Behauptung 3.4 und s_{1*} in Behauptung 3.5, in dieser Situation nicht bzw. nicht für jede Kombination von p und q existieren. Die Beweise lassen sich aber dennoch dazu nutzen, um Aussagen darüber zu treffen, wie eine Sequenz mit $\frac{k}{2}$ Kontrollen aussehen sollte:

Behauptung 3.6: *Wenn die Zahl der in einer Sequenz insgesamt eingesetzten Kontrollen genau $\frac{k}{2}$ betragen soll, so müssen nur Sequenzen beachtet werden, die genau eine der Kontrollen auf ein Randfeld setzen und bei denen keine zwei Kontrollen direkt benachbart sind.*

Beweis: Wegen Behauptung 3.1 ist klar, dass mindestens eine der zu verwendenden Kontrollen auf einem der Randfelder eingesetzt werden muss, dass also alle Sequenzen, die die Kontrollen ausschließlich im Mittelteil einsetzen würden, nicht mehr betrachtet werden müssen.

Es müssen also nur noch Sequenzen betrachtet werden, für die $p \in \{1, 2\}$ und $q \in \{\frac{k}{2} - 1, \frac{k}{2} - 2\}$ gilt. Für solche Parameterkombinationen existieren aber die im Beweis von Behauptung 3.4 benutzten Sequenzen s_1 und s_2 , der Beweis lässt sich also auch auf diese spezielle Situation anwenden und folglich können alle Sequenzen, die Kontrollen auf direkt benachbarten Positionen einsetzen, aussortiert werden.

Es bleibt also nur noch die Frage zu klären, ob eine der beiden Sequenzklassen $[p = 1, q = \frac{k}{2} - 1, n = 0]$ bzw. $[p = 2, q = \frac{k}{2} - 2, n = 0]$ der anderen überlegen ist. Hier hilft der Beweis zu Behauptung 3.5: In diesem wurde gezeigt, dass die Sequenz s_{2*} , die eben genannter

Sequenzklasse $[p = 1, q = \frac{k}{2} - 1, n = 0]$ entspricht, der Sequenz s_{3*} , die der Sequenzklasse $[p = 2, q = \frac{k}{2} - 2, n = 0]$ entspricht, im Intervall $[-1, -\frac{\sqrt{v+1}}{v-1+\sqrt{v+1}}]$ überlegen ist.

Da - wie ebenfalls in diesem Beweis gezeigt - auf dem Rest des relevanten Bereichs, dem Intervall $[-\frac{\sqrt{v+1}}{v-1+\sqrt{v+1}}, 0]$, die Sequenz s_0 , die keine Kontrolle einsetzt, die beiden Sequenzen s_{2*} und s_{3*} dominiert, ist die Sequenz s_{3*} also im gesamten relevanten Bereich einer anderen Sequenz unterlegen und deshalb für alle weiteren Überlegungen uninteressant. Es bleibt also letztlich nur die Sequenzklasse $[p = 1, q = \frac{k}{2} - 1, n = 0]$ für die weitere Betrachtung übrig. \square

Von den vormalig mehr als 2^{k-2} Sequenzen, die die Anwendung der Regeln aus [Kunert et al., 2010] überstanden haben, bleiben nach der weiteren Anwendungen der Behauptungen 3.2 bis 3.6 aus diesem Abschnitt damit insgesamt nur noch $(\frac{k}{2} + 1)$ zu betrachtende Sequenzklassen übrig: Für jede Anzahl verwendeter Kontrollen Z ($0 \leq Z \leq \frac{k}{2}$) gibt es nämlich genau eine einzige Sequenzklasse, die den Aussortiervorgang übersteht.

Für $0 \leq Z < \frac{k}{2}$ sieht diese Klasse und die zugehörigen m_1 - und m_2 -Werte wie folgt aus:

$$[p = 0, q = Z, n = 0] \stackrel{3.2}{\rightsquigarrow} \begin{cases} m_2(s) = \frac{1}{v}(Z + Z\lambda^2 + \gamma Z^2(1 - \lambda)^2) = \frac{1}{v}(Z + Z\lambda^2 + Z^2\varepsilon) \\ m_1(s) = \frac{1}{v-1}(2\alpha + (k - 2 - Z)\beta - m_2(s)) \end{cases}$$

Für $Z = \frac{k}{2}$ ergibt sich das folgende Aussehen der Sequenzklasse:

$$[p = 1, q = \frac{k}{2} - 1, n = 0] \stackrel{3.2}{\rightsquigarrow} \begin{cases} m_2(s) = \frac{1}{v}(\frac{k}{2} + (\frac{k}{2} - 1)\lambda^2 + \gamma((\frac{k}{2} - 1)(1 - \lambda) + 1)^2) \\ m_1(s) = \frac{1}{v-1}(\alpha + (\frac{k}{2} - 1)\beta - m_2(s)) \end{cases}$$

Da es für jedes Z jeweils nur eine einzige zu betrachtende Klasse gibt, kann in den kommenden Abschnitten deshalb zu folgender Kurzschreibweise übergegangen werden:

Definition 3.7: Mit $s_{(Z)}$ soll für $0 \leq Z \leq \frac{k}{2}$ jeweils ein Vertreter der Klasse von Sequenzen bezeichnet sein, für die gilt:

- In ihr werden genau Z Kontrollen eingesetzt.
- Kontrollen befinden sich in der Sequenz nicht an direkt benachbarten Stellen.
- Falls $Z < \frac{k}{2}$, so befindet sich keine Kontrolle auf einem der beiden Randfelder der Sequenz.
- Falls $Z = \frac{k}{2}$, so befindet sich genau eine Kontrolle auf einem der beiden Randfelder.

4 Reduzierung der Anzahl zu betrachtender Schnittpunkte

In den vorangegangenen Abschnitten wurden Sequenzen, die an keiner Stelle einer anderen Sequenz überlegen sind, aussortiert und die Zahl der zu betrachtenden Sequenzen somit reduziert. Nach diesem Aussortieren bleiben für gegebenes k noch genau $\lfloor \frac{k}{2} + 1 \rfloor$ zu betrachtende Sequenzen übrig.

Die Stelle x^* , für die $\max_s H_s(x^*)$ sein Minimum annimmt, kann wie in Behauptung 2.3 gezeigt wurde nur an dem Scheitelpunkt einer oder dem Schnittpunkt zweier oder mehrerer Sequenzen liegen. Bei der Suche nach ihr müssten demnach aber immer noch die Funktionswerte aller in Frage kommenden Sequenzen an all diesen Scheitel- und Schnittpunkten ermittelt werden: Dies würde aber bedeuten, dass jeweils $\lfloor \frac{k}{2} + 1 \rfloor$ Funktionswerte an $\lfloor \frac{k}{2} + 1 \rfloor$ Scheitel- und $2 \binom{\lfloor \frac{k}{2} + 1 \rfloor}{2}$ Schnittpunkte ausgewertet werden müssten.

Die Zahl dieser zu überprüfenden Schnitte würde demnach also quadratisch mit k wachsen. Es ist aber nicht nötig, wirklich alle Schnitte zu betrachten, da aufgrund von Behauptung 2.3 natürlich überhaupt nur Schnitte im Intervall $[-1, 0]$ von Interesse sein können und - wie in diesem Kapitel gezeigt wird - bei einem Großteil dieser Schnitte die H-Werte der sich schneidenden Sequenzen kleiner sind als die H-Werte anderer Sequenzen an dieser Stelle.

Dazu werden zunächst einige Aussagen über das Verhalten der Sequenzen $s_{(z)}$ im relevanten Intervall $[-1, 0]$ getroffen:

Behauptung 4.1: *Je mehr Kontrollen eine Sequenz $s_{(z)}$ einsetzt, desto höher ist der Wert der zu ihr gehörenden H-Funktion an der Stelle $x_1 = -1$ und desto niedriger ist er an der Stelle $x_2 = 0$.*

Beweis: Sei $z_1 = z$ und $z_2 = z + 1$, wobei $z \leq \frac{k-3}{2}$ ist. Sowohl bei $s_{(z_1)}$ als auch bei $s_{(z_2)}$ handelt es sich also um Sequenzen, die ohne Randbelegung auskommen. Die m_2 - und m_1 -Werte dieser Sequenzen sehen wie folgt aus:

$$\begin{aligned} m_2(s_{(z_1)}) &= \frac{1}{v}(z + z\lambda^2 + \gamma z^2(1 - \lambda)^2) \\ m_1(s_{(z_1)}) &= \frac{1}{v-1}(2\alpha + (k - 2 - z)\beta - m_2(s_{(z_1)})) \\ m_2(s_{(z_2)}) &= \frac{1}{v}(z + 1 + z\lambda^2 + \lambda^2 + \gamma z^2(1 - \lambda)^2 + 2\gamma z(1 - \lambda)^2 + \gamma(1 - \lambda)^2) \\ m_1(s_{(z_2)}) &= \frac{1}{v-1}(2\alpha + (k - 3 - z)\beta - m_2(s_{(z_2)})) \end{aligned}$$

Betrachte nun die Differenz der zu diesen Sequenzen gehörenden H-Funktionen. Mit

$$\begin{aligned} \Delta m_2 &:= m_2(s_{(z_2)}) - m_2(s_{(z_1)}) = \frac{1}{v}(1 + \lambda^2 + 2z \underbrace{\gamma(1 - \lambda)^2}_{\varepsilon} + \underbrace{\gamma(1 - \lambda)^2}_{\varepsilon}) = \frac{1}{v}(\beta + 2z\varepsilon) \\ \Delta m_1 &:= m_1(s_{(z_2)}) - m_1(s_{(z_1)}) = \frac{1}{v-1}(-\beta - \frac{1}{v}\beta - \frac{1}{v}2z\varepsilon) = -\frac{1}{v(v-1)}((v+1)\beta + 2z\varepsilon) \end{aligned}$$

lässt sich diese Differenz schreiben als

$$\Delta H(x) := H_{s_{(z_2)}}(x) - H_{s_{(z_1)}}(x) = (1 + x)^2 \Delta m_1 + (v - 1)x^2 \Delta m_2.$$

Die Ausdrücke $(\beta + 2z\varepsilon)$ und $((v+1)\beta + 2z\varepsilon)$ sind wegen Behauptung 2.7 aber immer positiv: Da $z \leq \frac{k-3}{2}$ ist, lassen sie sich nämlich nach unten durch $(\beta + (k-3)\varepsilon)$ und $((v+1)\beta + (k-3)\varepsilon)$

abschätzen und darauf dann besagte Behauptung anwenden. Somit ist klar, dass für alle Werte von $z \in [0, \lfloor \frac{k-3}{2} \rfloor]$ das oben definierte Δm_2 positiv und die Größe Δm_1 immer negativ ist.

Da weiter aber $\Delta H(-1) = (v-1)\Delta m_2$ und $\Delta H(0) = \Delta m_1$ ist, ist die Differenz der beiden Funktionen an der Stelle $x_1 = -1$ also positiv, an der Stelle $x_2 = 0$ aber negativ. Je mehr Kontrollen eine Sequenz $s_{(z)}$ für $z < \frac{k}{2}$ also einsetzt, desto höher ist ihr Funktionswert an der Stelle -1 und desto niedriger ist er an der Stelle 0 .

Noch ausgelassen wurde bei den letzten Überlegungen die Sequenz $s_{(\frac{k}{2})}$. Sei deshalb nun $z_{1*} = \frac{k}{2} - 1$ und $z_{2*} = \frac{k}{2}$. Während die Sequenz $s_{(z_{1*})}$ keine Kontrolle auf eines der beiden Randfelder legt, setzt Sequenz $s_{(z_{2*})}$ dort eine Kontrolle ein. Indem man sich zu Nutze macht, dass $\alpha = 1 + \gamma$ und folglich $1 = \alpha - \gamma$ ist, erhält man für die Differenzen der m_1 - und m_2 -Werte für diese neuen Sequenzen:

$$\begin{aligned}\Delta m_{2*} &:= m_2(s_{(z_{2*})}) - m_2(s_{(z_{1*})}) = \frac{1}{v}(\alpha + (k-2)\delta) \\ \Delta m_{1*} &:= m_1(s_{(z_{2*})}) - m_1(s_{(z_{1*})}) = -\frac{1}{v(v-1)}((v+1)\alpha + (k-2)\delta)\end{aligned}$$

In völliger Analogie zum Beweis zu Behauptung 2.7 lässt sich zeigen, dass $\alpha + (k-2)\delta$ und somit auch $(v+1)\alpha + (k-2)\delta$ echt größer Null ist: Da sowohl α als auch δ nach Behauptung 2.4 monoton in λ wachsend sind, lassen sich die Ausdrücke nämlich nach unten abschätzen, indem man ihren Wert für $\lambda = 0$ ermittelt.

Auch für diese Sequenzen gilt also wieder, dass Δm_{2*} positiv und Δm_{1*} negativ ist und - nach erneuter vollkommen analoger Schlußfolgerung - dass Sequenz $s_{(z_{2*})}$, die eine Kontrolle mehr verwendet als Sequenz $s_{(z_{1*})}$, an der Stelle -1 also auch höhere H-Werte erreicht, an der Stelle 0 dafür aber niedrigere.

Somit ist für alle Sequenzen $s_{(z)}$ die Behauptung gezeigt. \square

Behauptung 4.2: *Im Intervall $[-1, 0]$ schneiden sich zwei Sequenzen $s_{(z_1)}$ und $s_{(z_2)}$ mit z_1 bzw. $z_2 \leq \frac{k}{2}$ und $z_1 \neq z_2$ genau einmal.*

Beweis: Bei den zu den Sequenzen gehörenden H-Funktionen handelt es sich um Parabeln, siehe Behauptung 2.2. Diese können sich also an höchstens zwei Stellen schneiden.

Wie in Behauptung 4.1 gezeigt, ist der Funktionswert einer Sequenz $s_{(z)}$ an der Stelle $x_1 = -1$ umso größer, je mehr Kontrollen sie einsetzt, und an der Stelle $x_2 = 0$ umso kleiner. Folglich muss es es in dem Intervall zwischen diesen beiden Stellen aber eine ungerade Anzahl an Schnittpunkten geben: Gäbe es nämlich keinen oder eine gerade Anzahl an Schnitten, so könnte die Anordnung der Funktionswerte an den beiden Stellen nicht gegenläufig sein. Da es aber wie zu Beginn des Beweises erläutert höchstens zwei Schnittpunkte geben kann, ist die Behauptung somit gezeigt. \square

Da sich nach dieser Behauptung zwei zu betrachtende Sequenzen im relevanten Intervall nur einmal schneiden, kann die folgende Notation für den Schnittpunkt der Sequenzen verwendet werden.

Definition 4.3: Mit der Schreibweise $x_{i|j}$ wird der Schnittpunkt der zu den Sequenzen $s_{(i)}$ und $s_{(j)}$ gehörenden H-Funktionen im Intervall $[-1, 0]$ bezeichnet.

Über die Anordnung der Schnittpunkte im Intervall lassen sich nun folgende Aussagen treffen:

Behauptung 4.4: Für $0 \leq \lambda < 1$ gilt für die Schnittpunkte der Sequenzen:

$$x_{z+1|z+2} < x_{z|z+1}.$$

Beweis: Betrachte zunächst nur die Schnittpunkte der Sequenzen $s_{(z)}$ und $s_{(z+1)}$ mit $z \leq \frac{k-3}{2}$, d.h. nur die Schnittpunkte von Sequenzen, die sich nur um eine Kontrolle unterscheiden und die beide keine Kontrolle auf einem Randfeld einsetzen.

Die Differenz der H-Werte dieser Funktionen lässt sich schreiben als

$$\Delta H(x) := H_{s_{(z+1)}}(x) - H_{s_{(z)}}(x) = (1+x)^2 \Delta m_1 + (v-1)x^2 \Delta m_2,$$

wobei

$$\begin{aligned} \Delta m_2 &:= m_2(s_{(z+1)}) - m_2(s_{(z)}) = \frac{1}{v}(\beta + 2z\varepsilon) \\ \Delta m_1 &:= m_1(s_{(z+1)}) - m_1(s_{(z)}) = -\frac{1}{v(v-1)}((v+1)\beta + 2z\varepsilon) \end{aligned}$$

ist (vergleiche Beweis zu Behauptung 4.1).

Durch Nullsetzen von ΔH und anschließendem Auflösen nach x erhält man nun die folgenden beiden Schnittpunktgleichungen:

$$x_{s_1} = \frac{-\Delta m_1 - \sqrt{-(v-1)\Delta m_1 \Delta m_2}}{\Delta m_1 + (v-1)\Delta m_2} \quad \text{und} \quad x_{s_2} = \frac{-\Delta m_1 + \sqrt{-(v-1)\Delta m_1 \Delta m_2}}{\Delta m_1 + (v-1)\Delta m_2}.$$

Der Nenner dieser Ausdrücke, $\Delta m_1 + (v-1)\Delta m_2 = \frac{1}{v-1}(2\varepsilon z(v-2) + \beta(v-3))$, ist immer positiv. Beachte dazu, dass $z \leq \frac{k-3}{2}$ gelten muss und er ganz offensichtlich monoton in ε und β wächst und in z fällt. Durch Einsetzen der jeweils ungünstigsten Werte erhält man dann die folgende Abschätzung:

$$\Delta m_1 + (v-1)\Delta m_2 = \frac{2\varepsilon z(v-2) + \beta(v-3)}{v-1} \geq \frac{(v-k) + (v-3)}{k(v-1)} \underset{v \geq k > 3}{>} 0$$

Da aber nach Behauptung 2.7 Δm_2 immer positiv und Δm_1 immer negativ ist (siehe hierzu auch Beweis zu Behauptung 4.1), gilt somit für den zweiten Schnittpunkt x_{s_2} :

$$x_{s_2} = \frac{\overbrace{-\Delta m_1}^{>0} + \overbrace{\sqrt{-(v-1)\Delta m_1 \Delta m_2}}^{>0}}{\underbrace{\Delta m_1 + (v-1)\Delta m_2}_{>0}} > 0.$$

Der im relevanten Intervall $[-1, 0]$ liegende und damit für die weitere Betrachtung relevante Schnittpunkt ist also nicht der Punkt x_{s_2} , sondern der Punkt x_{s_1} .

Setzt man in die Schnittpunktgleichung die konkreten Werte für Δm_1 und Δm_2 ein, so ergibt sich die folgende Darstellung von x_{s_1} :

$$x_{s_1} = \frac{2\varepsilon z + (v+1)\beta - (v-1)\sqrt{(2\varepsilon z + \beta)(2\varepsilon z + (v+1)\beta)}}{2\varepsilon z v(v-2) + \beta v(v-3)}$$

Interessant zu untersuchen ist, wie sich die Lage des Schnittpunkts verändert, wenn man z , also die Anzahl eingesetzter Kontrollen in Sequenz $s_{(z)}$, variiert. Dazu wird z nicht als diskrete, sondern als stetige Einflussgröße betrachtet und der ganze Ausdruck danach abgeleitet.

Nach der Quotientenregel ergibt sich, dass $\frac{\partial x_{s_1}}{\partial z} = \frac{Z'N - ZN'}{N^2}$, wobei Z und N Zähler bzw. Nenner von x_{s_1} und Z' und N' die jeweiligen Ableitungen von Zähler und Nenner nach z bezeichnen.

Der Nenner N^2 dieser Ableitung ist auf Grund des Quadrats aber immer positiv. Im Folgenden ist für uns nur das Vorzeichen der Ableitung von Bedeutung, nicht aber ihr genauer Wert. Das Vorzeichen von $\frac{Z'N - ZN'}{N^2}$ entspricht aber dem Vorzeichen von $Z'N - ZN'$, insofern genügt es für die weitere Betrachtung also, sich auf letzteren Ausdruck zu konzentrieren. Mit

$$\begin{aligned} Z &= 2\epsilon z + (v+1)\beta - (v-1)\sqrt{(2\epsilon z + \beta)(2\epsilon z + (v+1)\beta)} & Z' &= 2\epsilon - \frac{(v-1)\epsilon(4\epsilon z + (v+2)\beta)}{\sqrt{(2\epsilon z + \beta)(2\epsilon z + (v+1)\beta)}} \\ N &= 2\epsilon z v(v-2) + \beta v(v-3) & N' &= 2v(v-2)\epsilon \end{aligned}$$

erhält man also

$$\begin{aligned} Z'N - ZN' &= (2\epsilon z v(v-2) + \beta v(v-3))\left(2\epsilon - \frac{(v-1)\epsilon(4\epsilon z + (v+2)\beta)}{\sqrt{(2\epsilon z + \beta)(2\epsilon z + (v+1)\beta)}}\right) \\ &\quad - (2v(v-2)\epsilon)((2\epsilon z + (v+1)\beta)) - (v-1)\sqrt{(2\epsilon z + \beta)(2\epsilon z + (v+1)\beta)} \\ &= \underbrace{4\epsilon^2 z v(v-2) + 2\beta\epsilon v(v-3)}_{(I)} + 2v(v-2)\epsilon(v-1)\sqrt{(2\epsilon z + \beta)(2\epsilon z + (v+1)\beta)} \\ &\quad - \underbrace{4v(v-2)\epsilon^2 z - 2v(v-2)(v+1)\beta\epsilon}_{(II)} - \frac{(2\epsilon z v(v-2) + \beta v(v-3))(v-1)\epsilon(4\epsilon z + (v+2)\beta)}{\sqrt{(2\epsilon z + \beta)(2\epsilon z + (v+1)\beta)}} \\ &= \underbrace{-2\beta\epsilon v(v-1)^2}_{(I)+(II)} - \frac{(2\epsilon z v(v-2) + \beta v(v-3))(v-1)\epsilon(4\epsilon z + (v+2)\beta)}{\sqrt{(2\epsilon z + \beta)(2\epsilon z + (v+1)\beta)}} \\ &\quad + 2v(v-2)\epsilon(v-1)\sqrt{(2\epsilon z + \beta)(2\epsilon z + (v+1)\beta)} \end{aligned}$$

Dieser gesamte Ausdruck lässt sich durch $\epsilon v(v-1)$ teilen. Da ϵ negativ ist, hat das Resultat dieser Division ein anderes Vorzeichen als der Originalausdruck. Dementsprechend wäre die Ableitung von x_{s_1} nach z genau dann negativ, wenn $Z'N - ZN'$ kleiner bzw. wenn $\frac{Z'N - ZN'}{\epsilon v(v-1)}$ größer Null ist. Diese letzte Ungleichung lässt sich also äquivalent weiterumformen:

$$\begin{aligned} \frac{Z'N - ZN'}{\epsilon v(v-1)} &> 0 \\ \Leftrightarrow -2\beta(v-1) - \frac{(2\epsilon z(v-2) + \beta(v-3))(4\epsilon z + (v+2)\beta)}{\sqrt{(2\epsilon z + \beta)(2\epsilon z + (v+1)\beta)}} + 2v(v-2)\epsilon(v-1)\sqrt{(2\epsilon z + \beta)(2\epsilon z + (v+1)\beta)} &> 0 \end{aligned}$$

Nun multipliziert man beide Seiten der Ungleichung mit $\sqrt{(2\epsilon z + \beta)(2\epsilon z + (v+1)\beta)}$. Da dies positiv ist, ändert diese Multiplikation nichts an der Richtung der Ungleichung.

$$\begin{aligned} \dots &\Leftrightarrow -2\beta(v-1)\sqrt{(2\epsilon z + \beta)(2\epsilon z + (v+1)\beta)} - (2\epsilon z(v-2) + \beta(v-3))(4\epsilon z + (v+2)\beta) \\ &\quad + 2v(v-2)\epsilon(v-1)(2\epsilon z + \beta)(2\epsilon z + (v+1)\beta) > 0 \\ &\Leftrightarrow 2(v-2)(2\epsilon z + \beta)(2\epsilon z + \beta + (v+1)\beta) - (2\epsilon z(v-2) + \beta(v-3))(4\epsilon z + (v+2)\beta) \\ &\quad > 2\beta(v-1)\sqrt{(2\epsilon z + \beta)(2\epsilon z + (v+1)\beta)} \\ &\Leftrightarrow 2\beta\epsilon v^2 z - 4\beta\epsilon v z + 4\beta\epsilon z + \beta^2 v^2 - \beta^2 v + 2\beta^2 \\ &\quad > 2\beta(v-1)\sqrt{(2\epsilon z + \beta)(2\epsilon z + (v+1)\beta)} \end{aligned}$$

$$\begin{aligned} &\Leftrightarrow 2\beta\epsilon z((v-1)^2+1) + \beta^2((v-1)^2+(v+1)) > 2\beta(v-1)\sqrt{(2\epsilon z+\beta)(2\epsilon z+(v+1)\beta)} \\ &\Leftrightarrow 2\epsilon z((v-1)^2+1) + \beta((v-1)^2+(v+1)) > 2(v-1)\sqrt{(2\epsilon z+\beta)(2\epsilon z+(v+1)\beta)} \end{aligned}$$

Die rechte Seite dieser Ungleichung ist ohnehin größer als Null, aber auch die linke Seite ist hier nicht negativ: Da $v > 0$ und wie schon erwähnt wegen 2.7 auch $b + 2\epsilon z > 0$ ist, ist nämlich auch

$$\begin{aligned} 2\epsilon z((v-1)^2+1) + \beta((v-1)^2+(v+1)) &\stackrel{v>0}{>} 2\epsilon z((v-1)^2+1) + \beta((v-1)^2+1) \\ &= \underbrace{(\beta+2\epsilon z)}_{>0} \underbrace{((v-1)^2+1)}_{>0 \forall v} \end{aligned}$$

nicht negativ. Da beide Seiten der Ungleichung größer Null sind, bleibt diese auch dann erhalten, wenn man beide Seiten quadriert. Man kann also auch schreiben:

$$\begin{aligned} \dots &\Leftrightarrow (2\epsilon z((v-1)^2+1) + \beta((v-1)^2+(v+1)))^2 > 4(v-1)^2(2\epsilon z+\beta)(2\epsilon z+(v+1)\beta) \\ &\Leftrightarrow 4\epsilon^2 z^2((v-1)^4 + 2(v-1)^2+1) + 4\beta\epsilon z((v-1)^4 + (v-1)^2(v+1) + (v-1)^2+(v+1)) \\ &\quad + \beta^2((v-1)^4 + 2(v-1)^2(v+1) + (v+1)^2) > 4(v-1)^2(2\epsilon z+\beta)(2\epsilon z+(v+1)\beta) \\ &\Leftrightarrow 4\epsilon^2 v^2 z^2(v^2-4v+4) + 4\beta\epsilon v^2 z(v^2-5v+6) + \beta^2 v^2(v^2-6v+9) > 0 \\ &\Leftrightarrow 4\epsilon^2 v^2 z^2(v-2)^2 + 4\beta\epsilon v^2 z(v-2)(v-3) + \beta^2 v^2(v-3)^2 > 0 \\ &\Leftrightarrow (2\epsilon v z(v-2) + \beta v(v-3))^2 > 0 \end{aligned}$$

Der Ausdruck in Klammern in der letzten Ungleichung ist nichts anderes als

$$v(v-1)(\Delta m_1 + (v-1)\Delta m_2),$$

was - wie schon gezeigt - für $v \geq k > 3$ immer echt positiv ist. Die letzte Ungleichung gilt also immer, dementsprechend ist auch $\frac{Z'N-ZN'}{\epsilon v(v-1)}$ immer größer oder gleich Null und $Z'N - ZN'$ kleiner oder gleich Null. Die Ableitung von x_{s_1} nach z ist also nie positiv: Dies bedeutet anders ausgedrückt, dass mit wachsendem z die Schnittpunkte der Sequenzen $s_{(z)}$ und $s_{(z+1)}$ jeweils ein Stück nach links wandern.

Bislang wurden in diesem Beweis jedoch nur Sequenzen ohne Randbelegung berücksichtigt. Wenn k gerade ist, gibt es mit der Sequenz $s_{(\frac{k}{2})}$ aber auch eine Sequenz von Interesse, die eine Kontrolle auf einem der Randfelder einsetzt. Es bleibt also für die Behauptung noch zu zeigen, dass der Schnittpunkt $x_{\frac{k}{2}-1|\frac{k}{2}}$ kleiner ist als der Schnittpunkt $x_{\frac{k}{2}-2|\frac{k}{2}-1}$.

Dies wird über einen Umweg gezeigt. Berechne dazu

$$\widetilde{\Delta H}(x) := H_{s_{(\frac{k}{2})}}(x) - H_{s_{(\frac{k}{2}-1)}}(x)$$

für den Schnittpunkt $x_{\frac{k}{2}-2|\frac{k}{2}-1}$.

Durch Einsetzen von $z = \frac{k}{2} - 2$ in die Schnittpunktgleichung lässt sich ermitteln, dass

$$x_{\frac{k}{2}-2|\frac{k}{2}-1} = \frac{(k-4)\epsilon+(v+1)\beta-(v-1)\sqrt{((k-4)\epsilon+\beta)((k-4)\epsilon+(v+1)\beta)}}{(k-4)\epsilon v(v-2)+\beta v(v-3)}$$

ist. Mit

$$\begin{aligned} \widetilde{\Delta m}_2 &:= m_2(s_{(\frac{k}{2})}) - m_2(s_{(\frac{k}{2}-1)}) = \frac{1}{v}(\alpha + (k-2)\delta) \\ \widetilde{\Delta m}_1 &:= m_1(s_{(\frac{k}{2})}) - m_1(s_{(\frac{k}{2}-1)}) = -\frac{1}{v(v-1)}((v+1)\alpha + (k-2)\delta) \end{aligned}$$

(vergleiche Beweis zu Behauptung 4.1) ergibt sich dann:

$$\begin{aligned}\widetilde{\Delta H}(x_{\frac{k}{2}-2|\frac{k}{2}-1}) &= (1 + x_{\frac{k}{2}-2|\frac{k}{2}-1})^2 \widetilde{\Delta m}_1 + (v-1)x_{\frac{k}{2}-2|\frac{k}{2}-1}^2 \widetilde{\Delta m}_2 \\ &= \left(\frac{(k-4)(v-1)^2 \varepsilon + (v-1)^2 \beta - (v-1) \sqrt{((k-4)\varepsilon + \beta)((k-4)\varepsilon + (v+1)\beta)}}{(k-4)\varepsilon v(v-2) + \beta v(v-3)} \right)^2 \\ &\quad \cdot \left(-\frac{1}{v(v-1)} ((v+1)\alpha + (k-2)\delta) \right) \\ &\quad + \left(\frac{(k-4)\varepsilon + (v+1)\beta - (v-1) \sqrt{((k-4)\varepsilon + \beta)((k-4)\varepsilon + (v+1)\beta)}}{(k-4)\varepsilon v(v-2) + \beta v(v-3)} \right)^2 \\ &\quad \cdot (v-1) \left(\frac{1}{v} (\alpha + (k-2)\delta) \right)\end{aligned}$$

Da der in beiden Summanden des Ausdrucks auftauchende Faktor $\frac{1}{((k-4)\varepsilon v(v-2) + \beta v(v-3))^2}$ in jedem Fall positiv ist, aber wieder nur das Vorzeichen des ganzen Ausdrucks von Interesse ist, kann der Faktor einfach weggelassen werden. Durch Ausmultiplizieren und Wiedereinfachen lässt sich der Ausdruck zudem noch etwas „vereinfachen“:

$$\begin{aligned}\widetilde{\Delta H}(x_{\frac{k}{2}-2|\frac{k}{2}-1}) \cdot ((k-4)\varepsilon v(v-2) + \beta v(v-3))^2 &= \\ &= -2(v-1)^2 ((k-2)\beta\delta - (k-4)\alpha\varepsilon) \sqrt{((v+1)\beta + (k-4)\varepsilon)(\beta + (k-4)\varepsilon)} \\ &\quad + \varepsilon((v-1)^3 + (v-1))(k-4)((k-2)\beta\delta - (k-4)\alpha\varepsilon) \\ &\quad + \beta((v-1)^3 + (v+1)(v-1))((k-2)\beta\delta - (k-4)\alpha\varepsilon)\end{aligned}$$

Aus jedem der Summanden lässt sich der Faktor $(v-1)$ herausziehen, am Vorzeichen des gesamten Ausdrucks ändert sich dadurch nichts. Zudem steckt in jedem Summand auch noch der Faktor $((k-2)\beta\delta - (k-4)\alpha\varepsilon)$. Dieser lässt sich schreiben als $(k-2)(\beta\delta - \alpha\varepsilon) + 2\alpha\varepsilon$. Für diesen Ausdruck gilt aufgrund von Behauptung 2.4 aber:

$$(k-2)(\beta\delta - \alpha\varepsilon) + 2\alpha\varepsilon = (k-2) \underbrace{(-\lambda(1-\lambda)^2\gamma^2)}_{\leq 0} + \underbrace{\gamma\lambda(1-\lambda^2)}_{\leq 0} + \underbrace{2\alpha \cdot \varepsilon}_{>0 <0} < 0$$

Der Faktor $((k-2)\beta\delta - (k-4)\alpha\varepsilon)$ ist also echt kleiner als Null, ein Teilen des gesamten Ausdrucks durch ihn führt dann dazu, dass sich sein Vorzeichen umdreht. Der Ausdruck $\widetilde{\Delta H}(x_{\frac{k}{2}-2|\frac{k}{2}-1})$ ist dementsprechend genau dann negativ, wenn

$$\frac{1}{(v-1)((k-2)\beta\delta - (k-4)\alpha\varepsilon)} \cdot \widetilde{\Delta H}(x_{\frac{k}{2}-2|\frac{k}{2}-1}) \cdot ((k-4)\varepsilon v(v-2) + \beta v(v-3))^2 > 0.$$

Es soll also gelten, dass

$$\begin{aligned}-2(v-1) \sqrt{((v+1)\beta + (k-4)\varepsilon)(\beta + (k-4)\varepsilon)} + \beta((v-1)^2 + (v+1)) + \varepsilon((v-1)^2 + 1)(k-4) &> 0 \\ \Leftrightarrow \beta((v-1)^2 + (v+1)) + \varepsilon((v-1)^2 + 1)(k-4) &> 2(v-1) \sqrt{((v+1)\beta + (k-4)\varepsilon)(\beta + (k-4)\varepsilon)}\end{aligned}$$

Beide Seiten dieser Ungleichung sind aber positiv. Bei der rechten Seite ist dies ganz offensichtlich, bei der linken folgt dies aus folgender Abschätzung:

$$\beta((v-1)^2 + (v+1)) + \varepsilon((v-1)^2 + 1)(k-4) = \underbrace{\beta v}_{>0} + \underbrace{(\beta + (k-4)\varepsilon)}_{>0 \text{ wegen Beh. 2.7}} \cdot \underbrace{((v-1)^2 + 1)}_{\geq 0 \forall v} > 0$$

Das Quadrieren beider Seiten lässt die Ungleichung also bestehen.

Mit einigen weiteren Umformungen wird aus der quadrierten Ungleichung so schließlich:

$$\dots \Leftrightarrow \left(\beta((v-1)^2 - (v+1)) + (k-4)\varepsilon((v-1)^2 - 1) \right)^2 > 0$$

Kleiner Null kann der Ausdruck auf der linken Seite aufgrund des Quadrats offensichtlich nie werden. Dass er aber sogar echt größer als Null ist, lässt sich durch die folgende Abschätzung zeigen, bei der die Monotonie von β und ε benutzt wird (siehe erneut Behauptung 2.4):

$$\begin{aligned} & \beta((v-1)^2 - (v+1)) + (k-4)\varepsilon((v-1)^2 - 1) \\ &= v((v-3) \cdot \underbrace{\beta}_{\geq \frac{k-1}{k}} + (k-4)(v-2) \cdot \underbrace{\varepsilon}_{\geq -\frac{1}{k}}) \geq (v-3) - \frac{v-3}{k} - \frac{(v-2)(k-4)}{k} = \frac{(v-k)+(2v-5)}{k} \underset{v \geq k \geq 3}{>} 0 \end{aligned}$$

Damit ist gezeigt, dass die Sequenz $s_{(\frac{k}{2})}$ an der Stelle $x_{\frac{k}{2}-2|\frac{k}{2}-1}$ für $0 \leq \lambda < 1$ schon echt kleinere Funktionswerte annimmt als die Sequenzen $s_{(\frac{k}{2}-2)}$ und $s_{(\frac{k}{2}-1)}$. Da aber - wie in Behauptung 4.1 gezeigt wurde - die Sequenz $s_{(\frac{k}{2})}$ an der Stelle $x = -1$ höhere Funktionswerte erreichte als die beiden anderen Sequenzen, muss es folglich irgendwo in dem Intervall $(-1, x_{\frac{k}{2}-2|\frac{k}{2}-1})$ schon zu einem Schnitt gekommen sein. \square

Mit den Behauptungen 4.1, 4.2 und 4.4 lässt sich nun zeigen, dass es nicht nötig ist, alle Schnittpunkte zweier Sequenzen $s_{(i)}$ und $s_{(j)}$ zu betrachten, sondern dass es genügt sich auf eine wesentlich kleinere Untermenge zu beschränken:

Behauptung 4.5: *Es müssen nur Schnittpunkte $x_{z|z+1}$ für $z \leq \frac{k}{2} - 1$ betrachtet werden.*

Beweis: In dem Intervall $[-1, 0]$ schneiden sich nach Behauptung 4.2 zwei Sequenzen genau einmal. Zwei Sequenzen $s_{(i)}$ und $s_{(i+1)}$, die sich um genau eine Kontrolle unterscheiden, weisen dabei aber wegen Behauptung 4.1 immer das folgende Verhalten auf: Vor ihrem Schnittpunkt $x_{i|i+1}$ erzielt die Sequenz mit der zusätzlichen Kontrolle, $s_{(i+1)}$, die höheren H-Werte, nach dem Schnitt die Sequenz $s_{(i)}$, die eine Kontrolle weniger einsetzt.

Betrachte nun den Schnittpunkt $x_{i|i+D}$ der Sequenzen mit i bzw. mit $i+D$ Kontrollen ($D > 1$). Egal wo dieser Schnittpunkt genau liegt, eine andere Sequenz erreicht an dieser Stelle immer höhere (bzw. zumindest nicht niedrigere) Funktionswerte: Läge dieser Schnittpunkt nämlich vor dem Schnittpunkt $x_{i|i+1}$ der Sequenzen mit i bzw. $i+1$ Kontrollen, so würde die Sequenz mit $i+1$ Kontrollen hier wegen den Ausführungen im ersten Absatz dieses Beweises höhere Funktionswerte erreichen als die Sequenz mit nur i Kontrollen, und der Schnittpunkt dementsprechend kein Kandidat mehr sein können. Läge der Schnittpunkt $x_{i|i+D}$ aber rechts von $x_{i|i+1}$, so würde er wegen Behauptung 4.4 auch rechts vom Schnittpunkt der Sequenzen mit $i+D$ bzw. $i+D-1$ Kontrollen liegen. Dementsprechend würde die Sequenz mit $i+D-1$ Kontrollen also hier besser abschneiden als die Sequenz mit $i+D$ Kontrollen, und der Schnittpunkt also wieder nicht von Interesse sein können. \square

Die Zahl der zu überprüfenden Schnittpunkte wurde somit in diesem Kapitel erheblich gesenkt: Wie gezeigt werden konnte, reicht die Betrachtung von höchstens $\lfloor \frac{k}{2} \rfloor$ Schnittpunkten für die Suche nach der Optimalstelle aus.

5 Das Vorgehen bei der Suche nach optimalen Sequenzen

In den letzten Kapiteln wurde zunächst die Zahl der zu betrachtenden Sequenzen und dann die Zahl der zu überprüfenden Schnittpunkte beträchtlich reduziert. Nach den in diesen Kapiteln vorgestellten Ergebnissen würden somit für gegebenes k nur noch die H -Funktionen von genau $\lfloor \frac{k}{2} + 1 \rfloor$ Sequenzen zur Betrachtung übrigbleiben, die an bis zu $\lfloor \frac{k}{2} + 1 \rfloor$ Scheitel- und $\lfloor \frac{k}{2} + 1 \rfloor$ Schnittpunkten ausgewertet werden müssten, um sicher $\max_s H_s(x^*) = \min_x \max_s H_s(x)$ zu finden.

Auch wenn das Ziel der Argumentation im letzten Kapitel nur war, irrelevante Schnittpunkte aus dem Kreis der Kandidaten zu entfernen, wurde dabei doch implizit noch einiges mehr gezeigt. Diese zusätzlichen Erkenntnisse, die sich aus den Behauptungen des letzten Kapitels ergeben, sollen hier noch einmal explizit wiedergegeben und daraus ein konstruktives Vorgehen für die Suche nach den optimalen Sequenzen abgeleitet werden.

Definition 5.1: Mit der Schreibweise $x_{(z)}$ wird der Scheitelpunkt der zu Sequenz $s_{(z)}$ gehörenden H -Funktion bezeichnet.

Behauptung 5.2: Für die Scheitelpunkte gilt: $-1 \leq x_{(z)} < x_{(z+1)} < 0$ für $z \leq \lfloor \frac{k}{2} - 1 \rfloor$.

Beweis: Für den Scheitelpunkt der Sequenz $s_{(z)}$ ergibt sich:

$$\begin{aligned} H_{s_{(z)}}(x) &= (1+x)^2 m_1(s_{(z)}) + (v-1)x^2 m_2(s_{(z)}) \\ \Rightarrow H'_{s_{(z)}}(x) &= 2(1+x) \cdot m_1(s_{(z)}) + 2(v-1)x \cdot m_2(s_{(z)}) = 0 \\ \Rightarrow x_{(z)} &= \frac{-m_1(s_{(z)})}{m_1(s_{(z)}) + (v-1)m_2(s_{(z)})} = \frac{-1}{1 + \frac{(v-1)m_2(s_{(z)})}{m_1(s_{(z)})}} \end{aligned}$$

Da aber $H_{s_{(z)}}(-1) = (v-1)m_2(s_{(z)})$ und $H_{s_{(z)}}(0) = m_1(s_{(z)})$ ist (vergleiche auch Beweis zu Behauptung 3.5), lässt sich der Scheitelpunkt $x_{(z)}$ also auch so darstellen:

$$x_{(z)} = \frac{-1}{1 + \frac{H_{s_{(z)}}(-1)}{H_{s_{(z)}}(0)}}$$

Da die H -Funktionen in jedem Fall nicht-negativ sind (siehe auch Behauptung 2.2), lassen sich die Untergrenze für die Scheitelpunkte, $x = -1$, und die für sie nicht erreichbare Obergrenze $x = 0$ unmittelbar dieser Darstellung entnehmen. Wegen Behauptung 4.1 wird der Quotient $\frac{H_{s_{(z)}}(-1)}{H_{s_{(z)}}(0)}$ nun offensichtlich aber auch noch umso größer, je größer z ist. Die Behauptung folgt nun aus der Tatsache, dass der Quotient im Nenner steht, der Zähler aber mit negativem Vorzeichen versehen ist. \square

Behauptung 5.3: Für die Ableitungen der H -Funktionen nach x gelten die folgenden beiden logischen Schlüsse:

$$\bullet H'_{s_{(z+1)}}(x) > 0 \Rightarrow H'_{s_{(z)}}(x) > 0 \quad \bullet H'_{s_{(z)}}(x) < 0 \Rightarrow H'_{s_{(z+1)}}(x) < 0$$

Beweis: Bei den verbleibenden H -Funktionen handelt es sich um nach oben geöffnete Parabeln, siehe Behauptung 2.2. Als solche gilt für sie natürlich:

$$H'_{s_{(z)}}(x) < 0, \text{ wenn } x < x_{(z)}, \quad \text{und} \quad H'_{s_{(z)}}(x) > 0, \text{ wenn } x > x_{(z)}.$$

Da wegen Behauptung 5.2 nun aber auch $x_{(z)} < x_{(z+1)}$ ist, ergibt sich weiter:

$$\begin{aligned} H'_{s_{(z+1)}}(x) > 0 &\Rightarrow x > x_{(z+1)} \Rightarrow x > x_{(z)} \Rightarrow H'_{s_{(z)}}(x) > 0 \\ H'_{s_{(z)}}(x) < 0 &\Rightarrow x < x_{(z)} \Rightarrow x < x_{(z+1)} \Rightarrow H'_{s_{(z+1)}}(x) < 0 \end{aligned}$$

Somit sind beide Teile der Behauptung also gezeigt. \square

Behauptung 5.4: *Der Scheitelpunkt der Sequenz $s_{(0)}$ kann nicht die gesuchte Stelle x^* darstellen.*

Beweis: Die Lage dieses Scheitelpunkts hängt nicht von λ oder den anderen Parametern ab, sondern er liegt immer an einer festen Stelle. Für diese Sequenz gilt nämlich z.B. nach Behauptung 3.2:

$$\begin{aligned} m_2(s_{(0)}) &= 0 \\ m_1(s_{(0)}) &= \frac{1}{v-1}(2\alpha + (k-2)\beta) \end{aligned}$$

Da $m_2(s_{(0)}) = 0$ ist, ergibt sich aus der Darstellung des Scheitelpunkts wie im Beweis zu Behauptung 5.2, $x_{(z)} = (1 + \frac{(v-1)m_2(s_{(z)})}{m_1(s_{(z)})})^{-1}$, dann aber direkt, dass dieser immer an der Stelle $x = -1$ liegen muss.

Nun gilt aber weiter nach Behauptung 4.1 für eben diese Stelle $x = -1$, dass die Sequenzen hier höhere H-Werte erzielen, je mehr Kontrollen sie einsetzen. Somit kann dieser Scheitelpunkt aber nicht der gesuchte Punkt sein und man kann ihn bei der Überprüfung übergehen. \square

Behauptung 5.5: *Sei $1 \leq z \leq \lfloor \frac{k}{2} \rfloor - 1$: Dann erreicht eine Sequenz $s_{(z)}$ im offenen Intervall $(x_{z|z+1}, x_{z-1|z})$ höhere H-Werte als die anderen Sequenzen und ist an den Schnittpunkten $x_{z|z+1}$ und $x_{z-1|z}$ nur ihren „Schnittpartnern“ $s_{(z+1)}$ bzw. $s_{(z-1)}$ nicht überlegen.*

Analog dazu sind die Sequenzen $s_{(\lfloor \frac{k}{2} \rfloor)}$ und $s_{(0)}$ in den Intervallen $[-1, x_{\lfloor \frac{k}{2} \rfloor - 1 | \lfloor \frac{k}{2} \rfloor})$ bzw. $(x_{0|1}, 0]$ bezüglich ihrer H-Werte den anderen Sequenzen überlegen und an den Schnittpunkten $x_{\lfloor \frac{k}{2} \rfloor - 1 | \lfloor \frac{k}{2} \rfloor}$ bzw. $x_{0|1}$ nur den Sequenzen $s_{(\lfloor \frac{k}{2} \rfloor - 1)}$ bzw. $s_{(1)}$ nicht überlegen.

Beweis: Diese Behauptung hängt mehr als eng mit Behauptung 4.5 zusammen und stellt sie quasi nur in einer Form dar, in der nun nicht darauf abgezielt wird, bestimmte Schnittpunkte aus dem Kandidatenkreis auszuschließen, sondern Erkenntnisse für die konkrete Suche nach optimalen Sequenzen zu gewinnen. Der Beweisgang ist dabei aber natürlich eigentlich identisch, daher wird die grundlegende Beweisidee hier auch nur noch in Kurzform präsentiert:

Wie schon im Beweis zu jener Behauptung anhand der Behauptungen 4.1 und 4.2 dargelegt wurde, gilt: Im Intervall $[-1, x_{i|i+1})$ ist $H_{s_{(i+1)}}(x) > H_{s_{(i)}}(x)$, im Intervall $(x_{i|i+1}, 0]$ dagegen $H_{s_{(i+1)}}(x) < H_{s_{(i)}}(x)$. Zudem gilt wegen Behauptung 4.4 aber auch, dass $x_{i|i+1} < x_{i-1|i}$ ist.

Zusammengenommen folgt aus diesen beiden Feststellungen aufgrund der Transitivität der Relationen aber direkt die Behauptung. Andere Szenarien, in denen beispielsweise davon ausgegangen würde, dass eine Sequenz $s_{(y)}$ im Intervall $[x_{z|z+1}, x_{z-1|z}]$ höhere H-Werte als $s_{(z)}$ erreicht, würden also sofort - wie im Beweis zu Behauptung 4.5 gesehen - zu einem Widerspruch führen. \square

In Abbildung 1 auf Seite 35 ist noch einmal das Verhalten der Sequenzen im relevanten Bereich $[-1, 0]$ - und somit auch die Aussagen der Behauptungen 4.1 bis 4.5 sowie 5.3 und 5.5 - anschaulich dargestellt.

Darüber, wie man anhand der zuletzt vorgestellten Erkenntnisse nun nach der Stelle $\operatorname{argmin}_x \max_s H_s(x)$ und somit nach den optimalen Sequenzen suchen kann, gibt die folgende Behauptung Aufschluß:

Behauptung 5.6: Die gesuchte Stelle findet man, indem man die Schnittpunkte $x_{i|i+1}$ der Reihe nach überprüft.

Die gesuchte Stelle ist dabei der Scheitelpunkt $x_{(\lfloor \frac{k}{2} \rfloor)}$ der Sequenz $s_{(\lfloor \frac{k}{2} \rfloor)}$, falls

$$H'_{s_{(\lfloor \frac{k}{2} \rfloor)}}(x_{\lfloor \frac{k}{2} \rfloor - 1 | \lfloor \frac{k}{2} \rfloor}) > 0$$

ist.

Die gesuchte Stelle ist der Scheitelpunkt $x_{(i)}$ der Sequenz $s_{(i)}$ (für $\lfloor \frac{k}{2} \rfloor - 1 \geq i \geq 1$), falls

$$H'_{s_{(i)}}(x_{i|i+1}) < 0 \text{ und } H'_{s_{(i)}}(x_{i-1|i}) > 0$$

ist.

Die gesuchte Stelle ist der Schnittpunkt $x_{i|i+1}$ der Sequenzen $s_{(i)}$ und $s_{(i+1)}$ (für $\lfloor \frac{k}{2} \rfloor - 1 \geq i \geq 0$), falls

$$H'_{s_{(i)}}(x_{i|i+1}) \geq 0 \text{ und } H'_{s_{(i+1)}}(x_{i|i+1}) \leq 0$$

ist.

Beweis: Wie in Behauptung 5.5 gezeigt wurde gibt es für jede Sequenz $s_{(z)}$ ein Intervall, an der die zu ihr gehörende H-Funktion allen anderen überlegen ist. Dieses Intervall ist durch die Schnittpunkte $x_{z|z+1}$ und $x_{z-1|z}$ begrenzt, wenn $1 \leq z \leq \lfloor \frac{k}{2} \rfloor - 1$ ist, und durch die Punkte -1 und $x_{\lfloor \frac{k}{2} \rfloor - 1 | \lfloor \frac{k}{2} \rfloor}$ bzw. $x_{0|1}$ und 0 , wenn $z = \lfloor \frac{k}{2} \rfloor$ bzw. $= 0$ ist.

Wenn nun an der Intervallgrenze $x_{z|z+1}$ $H'_{s_{(z)}}(x_{z|z+1}) < 0$ ist und an der anderen Grenze $x_{z-1|z}$ $H'_{s_{(z)}}(x_{z-1|z}) > 0$ ist, so muss es in dem dazwischen liegenden Intervall aufgrund der Stetigkeit der Funktionen eine Stelle x geben, an der $H'_{s_{(z)}}(x) = 0$ ist. Das Minimum der Maximumfunktion muss also irgendwo in diesem offenen Intervall liegen, und da die Sequenz $s_{(z)}$ hier echt größere H-Werte als alle anderen Sequenzen erzielt, kann nur ihr Scheitelpunkt die gesuchte Stelle sein.

Für das Randintervall $(-1, x_{\lfloor \frac{k}{2} \rfloor - 1 | \lfloor \frac{k}{2} \rfloor})$ gilt die gleiche Argumentation. Beachte, dass hier keine Bedingung an die Ableitung an der Intervallgrenze $x = -1$ gestellt werden muss: Wegen Behauptung 4.1 erreicht hier Sequenz $s_{\lfloor \frac{k}{2} \rfloor}$ die höchsten H-Werte, wegen Behauptung 5.2 liegt ihr Scheitelpunkt jedoch auf jeden Fall weiter rechts, d.h. die zugehörige H-Funktion fällt an dieser Stelle auf jeden Fall. Beachte weiter, dass das Intervall $(x_{0|1}, 0]$ nicht überprüft werden muss, da der Scheitelpunkt der Sequenz $s_{(0)}$ wegen Behauptung 5.4 schon ausgeschlossen werden konnte.

Auch die die Schnittpunkte betreffende Aussage in der Behauptung ist schnell gezeigt: Da das Maximum einer Schar streng konvexer Funktionen selbst wieder eine streng konvexe Funk-

tion darstellt, ist klar, dass wenn an einer Stelle $x_{z|z+1}$ die eine Funktion fällt und die andere wächst, diese Stelle die gesuchte sein muss. Schließlich gibt es an dieser Stelle nach Behauptung 5.5 auch keine Funktionen, die höhere Werte erzielen als $H_{s(z)}$ und $H_{s(z+1)}$, wäre die Stelle also nicht die gesuchte, so ergäbe sich ein Widerspruch zur Konvexität der Maximumfunktion. \square

Da der Arbeitsaufwand der Suche nach dieser Minimalstelle schon in den letzten beiden Kapiteln durch den Ausschluss zahlreicher Sequenzen und in Frage kommender Schnittpunkte sehr stark reduziert wurde, bringt Behauptung 5.6 für dieses Vorhaben keine wesentliche Arbeitserleichterung mehr mit sich. Dennoch ist sie aber von großer Wichtigkeit: Anhand von ihr lassen sich - wie im folgenden Kapitel zu sehen sein wird - allgemeine Aussagen über das Aussehen optimaler Sequenzen herleiten, die nicht nur für eine konkrete Einstellung der Parameter Gültigkeit haben.

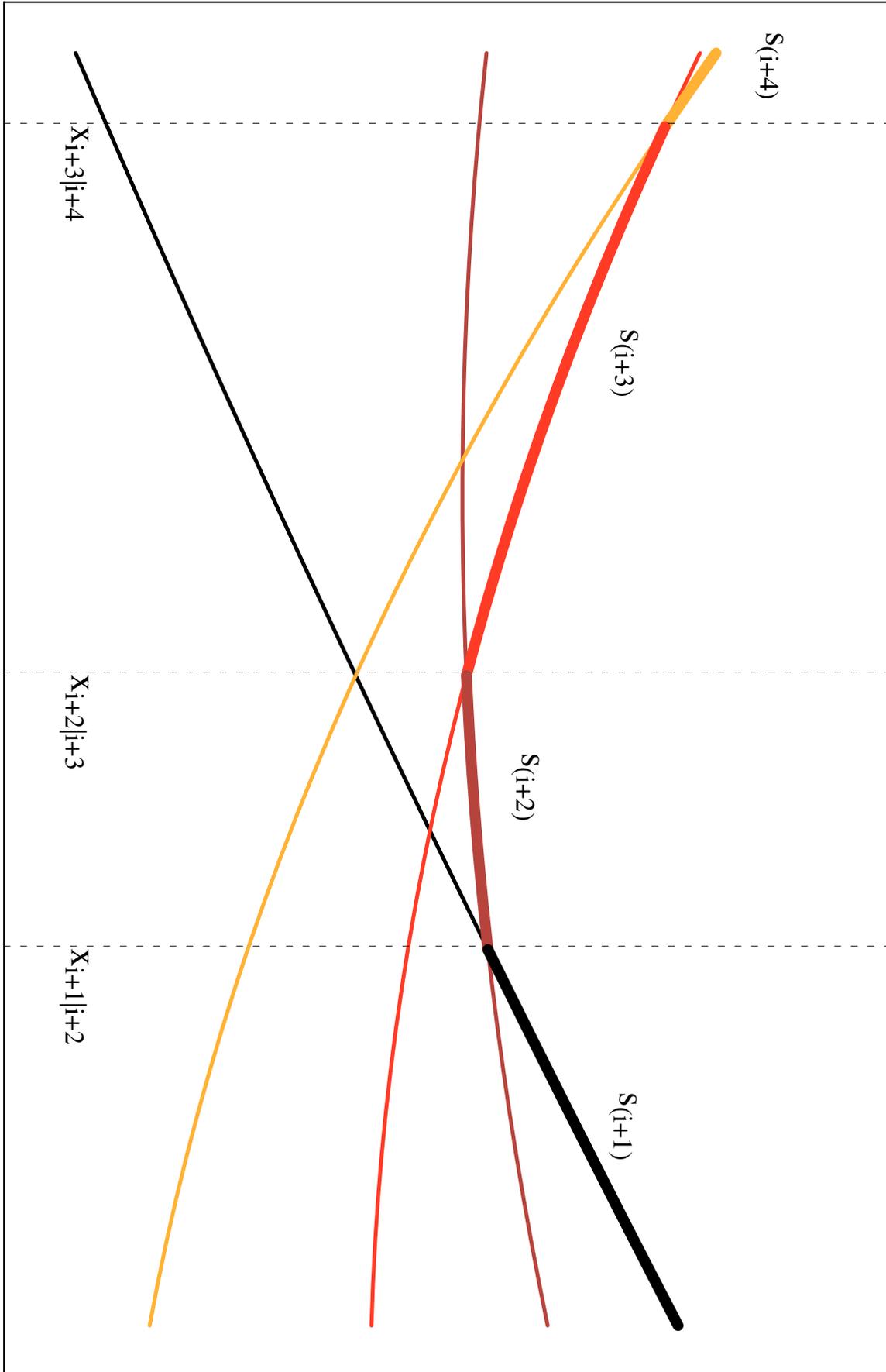


Abbildung 1: Das Verhalten der Sequenzen im Intervall $[-1, 0]$

6 Allgemeine Aussagen über die optimalen Sequenzen

Im letzten Kapitel war das Ziel der Überlegungen, eine konstruktive Vorgehensweise herzuleiten, mit der sich für gegebene Parametereinstellungen die gesuchte Stelle einfach und ohne allzu viele Funktionsauswertungen ermitteln lässt. In diesem Kapitel soll dieses und andere der in dieser Arbeit vorgestellten Ergebnisse dazu benutzt werden, einige allgemeine Aussagen über das Aussehen optimaler Sequenzen und ihre Abhängigkeit von den Parameterwerten zu treffen.

Die erste Aussage dieses Kapitels betrifft das Aussehen optimaler Sequenzen für hinreichend große Werte von v :

Behauptung 6.1: *Der Fall, dass $x_{0|1}$ die gesuchte Stelle darstellt, kann nur dann eintreten, wenn $v > (k-2)^2$ ist. Ist v sogar größer als $(k-1)^2$, so tritt er immer ein, egal welchen Wert λ annimmt.*

Beweis: Wie auch schon im Beweis zu Behauptung 5.6 zu sehen war, schneiden sich die Sequenzen $s_{(0)}$ und $s_{(1)}$, egal welche Werte k und λ annehmen, immer an der Stelle

$$x_{0|1} = -\frac{\sqrt{v+1}}{v-1+\sqrt{v+1}}.$$

Wegen Behauptung 5.5 ist auch klar, dass alle anderen Sequenzen an dieser Stelle niedrigere H-Werte aufweisen als die beiden Sequenzen $s_{(0)}$ und $s_{(1)}$.

Da der Scheitelpunkt der Sequenz $s_{(0)}$ wie im Beweis zu Behauptung 5.4 gezeigt immer bei $x_{(0)} = -1$ (und damit auf jeden Fall links vom Schnittpunkt $x_{0|1}$) liegt, ergibt sich weiter, dass die zu $s_{(0)}$ gehörende H-Funktion hier bereits monoton wachsend ist.

Für den Beweis dieser Behauptung bleibt also nach Behauptung 5.6 noch zu untersuchen, wann die zur Sequenz $s_{(1)}$ gehörende H-Funktion an dieser Stelle monoton fallend ist, die Ableitung ihrer H-Funktion nach x hier also kleiner oder gleich Null ist.

Für $H'_{s_{(1)}}$ ergibt sich an diesem Schnittpunkt:

$$H'_{s_{(1)}}\left(-\frac{\sqrt{v+1}}{v-1+\sqrt{v+1}}\right) = 2\left(\frac{v-1}{v-1+\sqrt{v+1}}\right)\frac{1}{v-1}(2\alpha + (k-3)\beta - \frac{\beta}{v}) + 2\left(-\frac{\sqrt{v+1}}{v-1+\sqrt{v+1}}\right)(v-1)\frac{\beta}{v}$$

Diese Ableitung ist aber genau dann kleiner oder gleich Null, wenn gilt:

$$\begin{aligned} H'_{s_{(1)}}\left(-\frac{\sqrt{v+1}}{v-1+\sqrt{v+1}}\right) \leq 0 &\Leftrightarrow (2\alpha + (k-3)\beta - \frac{\beta}{v}) - \sqrt{v+1}\left(\frac{v-1}{v}\right)\beta \leq 0 \\ &\Leftrightarrow \underbrace{2 + (k-3 - \frac{1}{v})\frac{\beta}{\alpha}}_{\alpha > 0} - \frac{\sqrt{v+1}(v-1)}{v}\frac{\beta}{\alpha} \leq 0 \\ &\Leftrightarrow (k-3 - \frac{1}{v} - \frac{\sqrt{v+1}(v-1)}{v})\frac{\beta}{\alpha} \leq -2 \end{aligned}$$

Über den hier auftretenden Quotienten $\frac{\beta}{\alpha}$ wurde in Behauptung 2.5 eine Aussage getroffen: Danach ist er monoton in λ wachsend und ist nach unten durch 1 und nach oben durch 2 begrenzt. Basierend auf dieser Erkenntnis lässt sich nun berechnen, was für gegebenes k das kleinstmögliche v ist, für das die Ableitung $H'_{s_{(1)}}$ am Schnittpunkt nicht-positiv ist: Dazu setzt man in der Ungleichung die „günstigste“ wenngleich eigentlich ausgeschlossene Wertebelegung $\lambda = 1$ bzw. $\frac{\beta}{\alpha} = 2$ ein und löst die Ungleichung daraufhin nach v auf.

Man erhält auf diesem Wege:

$$\begin{aligned} k - 2 - \frac{1}{v} - \frac{\sqrt{v+1}(v-1)}{v} = 0 &\Leftrightarrow \frac{v(k-2)-1}{v-1} = \sqrt{v+1} \\ &\Rightarrow \frac{(v(k-2)-1)^2}{(v-1)^2} = v+1 \\ &\Leftrightarrow v = \frac{(k-3)\sqrt{(k-1)^2+4}+(k-2)^2+1}{2} \end{aligned}$$

Da v aber eigentlich natürlich nur ganzzahlige Werte annehmen kann, macht es Sinn zu untersuchen, welches das kleinste ganzzahlige v^* ist, das über dem eben berechneten Wert für v liegt. Um diesen Wert zu bestimmen, schätzen wir den letzten Ausdruck zunächst nach unten ab:

$$\frac{(k-3)\sqrt{(k-1)^2+4}+(k-2)^2+1}{2} \geq \frac{(k-3)\sqrt{(k-1)^2}+(k-2)^2+1}{2} = \frac{(k-3)(k-1)+(k-2)^2+1}{2} = (k-2)^2$$

Für $v = (k-2)^2$ ist am Schnittpunkt $x_{0|1}$ die Ableitung der Sequenz $s_{(1)}$ also noch nicht fallend, für $v = (k-2)^2 + 1$ für die uns interessierenden Werte von k mit $k \geq 4$ hingegen schon. Es gilt nämlich für $v = (k-2)^2 + 1$:

$$\begin{aligned} (k-2)^2 + 1 &\geq \frac{(k-3)\sqrt{(k-1)^2+4}+(k-2)^2+1}{2} \Leftrightarrow (k-2)^2 + 1 \geq (k-3)\sqrt{(k-1)^2+4} \\ &\Rightarrow k^4 - 8k^3 + 26k^2 - 40k + 25 \geq k^4 - 8k^3 + 26k^2 - 48k + 45 \\ &\Leftrightarrow 8k \geq 20 \Leftrightarrow k \geq \frac{5}{2}. \end{aligned}$$

Da die von uns betrachteten Werte von k immer größer sind als in dieser Schranke gefordert, ist dieser Teil der Behauptung also gezeigt.

Analog zur Untersuchung, ab welchen v -Werten der Fall, dass der Schnittpunkt $x_{0|1}$ der gesuchte Punkt ist, frühestens auftreten kann, lässt sich auch untersuchen, ob es eine Grenze gibt, ab der er immer auftreten muss. Für den Quotienten $\frac{b}{a}$ wird dazu der „ungünstigste“ Fall $\lambda = 0$ bzw. $\frac{b}{a} = 1$ betrachtet. Analog zum zuvor gezeigten muss dazu der Ausdruck $k - 3 - \frac{1}{v} - \frac{\sqrt{v+1}(v-1)}{v}$ kleiner oder gleich -2 sein. Umformen und Auflösen liefert als Lösung für v dann schließlich:

$$v = \frac{(k-2)\sqrt{k^2+4}+k^2-2k+2}{2}$$

Dieser Wert liegt für $k \geq \frac{3}{2}$ zwischen den beiden folgenden ganzzahligen Schranken $(k-1)^2$ und $(k-1)^2 + 1$, wie man durch folgende Abschätzungen bzw. Umformungen zeigen kann:

$$\begin{aligned} \frac{(k-2)\sqrt{k^2+4}+k^2-2k+2}{2} &\stackrel{\substack{\geq \\ \sqrt{k^2+4} \geq k}}{\geq} \frac{(k-2)k+k^2-2k+2}{2} = \frac{2k^2-4k+2}{2} = (k-1)^2 \\ \frac{(k-2)\sqrt{k^2+4}+k^2-2k+2}{2} &\leq (k-1)^2 + 1 \Leftrightarrow (k-2)\sqrt{k^2+4} \leq k^2 - 2k + 2 \\ &\Rightarrow (k-2)^2(k^2+4) \leq (k^2 - 2k + 2)^2 \\ &\Leftrightarrow k^4 - 4k^3 + 8k^2 - 16k + 16 \leq k^4 - 4k^3 + 8k^2 - 8k + 4 \\ &\Leftrightarrow -8k \leq -12 \Leftrightarrow k \geq \frac{3}{2} \end{aligned}$$

Damit ist die obere Schwelle, ab der besagter Fall immer eintritt, also gefunden und der Beweis somit abgeschlossen. \square

Wie in der letzten Behauptung gezeigt wurde, ist es für sehr große Werte von v optimal, eine Mischung der Sequenzen $s_{(0)}$ und $s_{(1)}$ zu verwenden, also gerade der Sequenzen, die die Kon-

trolle seltenst bis gar nicht einsetzen. Dass sich diese Erkenntnis nun ausweiten und es sich ganz allgemein zeigen lässt, dass mit wachsendem Wert von v generell immer weniger Kontrollen eingesetzt werden sollten, zeigen die nächsten Behauptungen. Die erste von ihnen ist dabei eine sehr technische. Ihre Bedeutung wird schließlich mit Behauptung 6.3 klar.

Behauptung 6.2: Bezeichne $x_{(z)}^{(v)}$ den Scheitelpunkt der Sequenz $s_{(z)}$ für gegebenes v und $x_{(z)}^{(v+1)}$ den Scheitelpunkt derselben Sequenz, wenn nicht v , sondern $v + 1$ Behandlungen mit der Kontrolle verglichen werden sollten.

Analog dazu seien mit $x_{z|z+1}^{(v)}$ und $x_{z|z+1}^{(v+1)}$ die Schnittpunkte der Sequenzen $s_{(z)}$ und $s_{(z+1)}$ bezeichnet, wenn einmal v und einmal $v + 1$ Behandlungen mit der Kontrolle verglichen werden sollten.

Weiter seien als relative Wachstumsgeschwindigkeit die Quotienten

$$\frac{x_{(z)}^{(v+1)}}{x_{(z)}^{(v)}} \text{ bzw. } \frac{x_{z|z+1}^{(v+1)}}{x_{z|z+1}^{(v)}}$$

bezeichnet, wobei aufgrund des negativen Vorzeichens aller relevanten Schnitt- und Scheitelpunkte gilt, dass das Wachstum umso schneller ist, je niedriger der Wert des Quotienten.

Dann gilt für $1 \leq z \leq \lfloor \frac{k-1}{2} \rfloor$ und $0 \leq i \leq \lfloor \frac{k-3}{2} \rfloor$: Ein Scheitelpunkt $x_{(z)}$ wächst mit wachsendem v mit höherer relativer Geschwindigkeit als jeder der Schnittpunkte $x_{i|i+1}$.

Beweis: In der Behauptung werden nur die Schnitt- und Scheitelpunkte der Sequenzen betrachtet, die auf die Verwendung einer Kontrolle auf einem der Randfelder verzichten. Auch umfasst die Behauptung nicht den Scheitelpunkt $x_{(0)}$, der ja - wie in Behauptung 5.4 gezeigt - immer an der Stelle -1 liegt und somit also gar nicht wächst.

Wegen des Hauptsatzes der Integralrechnung (siehe z.B. [Hildebrandt, 2006], Seite 289) gilt für negative $x^{(v)}$ und $x^{(v+1)}$ der folgende Zusammenhang:

$$\ln\left(\frac{x^{(v+1)}}{x^{(v)}}\right) = \ln(-x^{(v+1)}) - \ln(-x^{(v)}) = \int_v^{v+1} \frac{\partial x(u)}{x(u)} du$$

Aufgrund der Monotonie des natürlichen Logarithmus lässt sich weiter folgern: Gilt immer

$$\frac{\frac{\partial x_{j_1}^{(v)}}{\partial v}}{x_{j_1}^{(v)}} > \frac{\frac{\partial x_{j_2}^{(v)}}{\partial v}}{x_{j_2}^{(v)}},$$

so gilt auch

$$\frac{x_{j_1}^{(v+1)}}{x_{j_1}^{(v)}} > \frac{x_{j_2}^{(v+1)}}{x_{j_2}^{(v)}}.$$

Anstatt als Geschwindigkeitsmaß für das Wachstum die in der Behauptung dargestellten Quotienten zu verwenden, kann man also auch die Quotienten

$$Q_1 = \frac{\frac{\partial x_{(z)}^{(v)}}{\partial v}}{x_{(z)}^{(v)}} \text{ bzw. } Q_2 = \frac{\frac{\partial x_{z|z+1}^{(v)}}{\partial v}}{x_{z|z+1}^{(v)}}$$

betrachten und, wenn für dieses Geschwindigkeitsmaß die Behauptung gilt, auf das andere Maß schließen. Durch dieses „Auswechseln“ des Geschwindigkeitsmaßes erleichtern sich die für den Beweis erforderlichen Berechnungen nämlich erheblich.

Um zu zeigen, dass jeder der betrachteten Scheitel- schneller in v wächst als jeder der betrachteten Schnittpunkte, reicht es aus, die Wachstumsgeschwindigkeit des langsamsten Scheitel- mit der des schnellsten Schnittpunkts zu vergleichen: Ist ersterer immer noch schneller als letzterer, so gilt dies erst recht für jeden anderen Scheitel- und jeden anderen Schnittpunkt.

Der langsamste Scheitelpunkt lässt sich durch Ableitung des Quotienten Q_1 nach z ermitteln:

$$\frac{\partial Q_1}{\partial z} = \frac{-(v-1) \cdot A \cdot B \cdot C}{(z(1+\lambda^2+z\varepsilon)+\beta v(z-k)+2v(\beta-\alpha))^2 (vz(1+\lambda^2+\varepsilon z)-z(2+\beta+2\lambda^2+2\varepsilon z)+(k-2)\beta+2\alpha)^2}$$

mit

$$A = z + z\lambda^2 + z^2\varepsilon - (k-2-z)\beta - 2\alpha,$$

$$B = -((k-2)\beta + 2\alpha)(1 + \lambda^2 + 2\varepsilon z) + \beta\varepsilon z^2 \quad \text{und}$$

$$C = (v-3)\varepsilon z^2 + (\lambda^2 - \beta + 1)vz - z(3\lambda^2 + \beta + 3) + (v+1)(2\alpha + (k-2)\beta).$$

Der Nenner dieser Ableitung ist aufgrund der Quadrate natürlich immer positiv. Da für die betrachteten Scheitelpunkte gilt, dass $1 \leq z \leq \lfloor \frac{k-1}{2} \rfloor$, lassen sich die im Zähler vorkommenden Faktoren A , B und C wie folgt abschätzen:

$$\begin{aligned} A &= z + z\lambda^2 + \underbrace{z^2\varepsilon}_{\leq z\varepsilon} - (k-2-z)\beta - 2\alpha \leq \underbrace{z + z\lambda^2 + z\varepsilon}_{=z\beta} - (k-2-z)\beta - 2\alpha \\ &= \underbrace{(2z-k)\beta}_{\leq -1} + 2(\beta - \alpha) \leq \beta - 2\alpha < 0 \end{aligned} \quad (\text{a})$$

$$B = -((k-2)\beta + 2\alpha)(1 + \lambda^2 + 2\varepsilon z) + \underbrace{\beta\varepsilon z^2}_{< 0} < -\underbrace{((k-2)\beta + 2\alpha)}_{> 0} \underbrace{(1 + \lambda^2 + 2\varepsilon z)}_{> 0 \text{ wegen (b)}} < 0$$

$$\begin{aligned} C &= (v-3)\varepsilon z^2 + \underbrace{(\lambda^2 - \beta + 1)}_{=\varepsilon} vz - \underbrace{z(3\lambda^2 + \beta + 3)}_{\leq 8} + (v+1) \underbrace{(2\alpha + (k-2)\beta)}_{\geq k\alpha} \\ &\geq (v-3)\varepsilon z^2 + \varepsilon vz - 8z + (v+1)k\alpha \\ &\geq (v-3)\varepsilon \left(\frac{k-1}{2}\right)^2 + \varepsilon v \frac{k-1}{2} - 8 \frac{k-1}{2} + (v+1)k\alpha \quad (\text{c}) \\ &\geq -(v-3) \frac{1}{k} \left(\frac{k-1}{2}\right)^2 - v \frac{1}{k} \frac{k-1}{2} - 8 \frac{k-1}{2} + (v+1)(k-1) \quad (\text{d}) \\ &= \frac{(k-1)(v(3k-1)-9k-3)}{4k} \\ &= \frac{(k-1)(k(3k-1)-9k-3)}{4k} \quad (\text{e}) \\ &= \frac{(k-1)(3(k-3)^2+8k-30)}{4k} > 0 \quad \text{für } k \geq 4. \end{aligned}$$

Bei den Abschätzungen wurden dabei an den mit Buchstaben markierten Stellen die folgenden Begründungen gebraucht:

- Bei (a) wurde das Resultat von Behauptung 2.5 ausgenutzt.
- Bei (b) wurde ausgenutzt, dass der Ausdruck offensichtlich monoton wachsend in λ (und somit auch in ε) ist und ebenso offensichtlich monoton fallend in z . Durch Einsetzen der jeweils ungünstigsten Werte $\lambda = 0$ und $z = \frac{k-1}{2}$ gelangt man so zu einer Abschätzung nach unten, die immer noch positiv ist.
- Auch bei (c) wurde die offensichtliche Monotonie in z ausgenutzt und z wieder durch den ungünstigsten Fall, $z = \frac{k-1}{2}$, ersetzt. Diese Monotonie liegt vor, da die Summanden, in denen z auftaucht, alle kleiner oder gleich Null sind.

- Bei (d) wurde analog dazu die Monotonie in λ benutzt, um α und ε - und damit natürlich auch den gesamten Ausdruck - nach unten abzuschätzen.
- An der mit (e) gekennzeichneten Stelle folgt die Ungleichung schließlich daraus, dass der Ausdruck offensichtlich in v wächst und in dieser Arbeit nur Fälle berücksichtigt werden, in denen $v \geq k$ gilt, v also kleinstenfalls den Wert $v = k$ annehmen kann.

Da die Abschätzungen also zeigen, dass A und B beide kleiner oder gleich Null sind, C aber für die in der Arbeit betrachteten Werte von k positiv ist, ist auch das Produkt $A \cdot B \cdot C$ positiv. Da der ebenfalls im Zähler vorkommende Vorfaktor $-(v-1)$ in jedem Fall negativ ist, ist der ganze Zähler - und somit auch der ganze Bruch - für $k \geq 4$, $v \geq k$, $0 \leq \lambda < 1$ negativ. Der Quotient Q_1 ist somit monoton fallend in z . Mit anderen Worten: Der Scheitelpunkt, der am langsamsten wandert, ist der, den wir für $z = 1$ erhalten, also der Scheitelpunkt $x_{(1)}$ (der Scheitelpunkt $x_{(0)}$ wurde in dieser Behauptung ja bewusst außer Acht gelassen, da er gar nicht wandert).

Auf dem selben Wege wird nun nach dem am schnellsten wachsenden Schnittpunkt gesucht. Die Ableitung von Q_2 nach z ergibt dabei:

$$\frac{\partial Q_2}{\partial z} = \frac{-\beta\varepsilon(2\varepsilon z + \beta)(D+E)}{2(2\varepsilon z + (v+1)\beta)^2(2\varepsilon z(v-2) + \beta(v-3))^2 \sqrt{(2\varepsilon z + \beta)(2\varepsilon z + \beta(v+1))}(\sqrt{(2\varepsilon z + \beta)(2\varepsilon z + \beta(v+1))} - (\beta + 2\varepsilon z)(v-1))^2},$$

wobei

$$\begin{aligned} D = & 2(v-1)\sqrt{(2\varepsilon z + \beta)(2\varepsilon z + (v+1)\beta)}(8\varepsilon^3 v^3 z^3 - 40\varepsilon^3 v^2 z^3 + 64\varepsilon^3 v z^3 - 64\varepsilon^3 z^3 + 12\beta\varepsilon^2 v^3 z^2 \\ & - 76\beta\varepsilon^2 v^2 z^2 + 112\beta\varepsilon^2 v z^2 - 144\beta\varepsilon^2 z^2 + 2\beta^2\varepsilon v^3 z - 42\beta^2\varepsilon v^2 z + 46\beta^2\varepsilon v z - 102\beta^2\varepsilon z \\ & - \beta^3 v^3 - 9\beta^3 v^2 + \beta^3 v - 23\beta^3) \end{aligned}$$

und

$$\begin{aligned} E = & -((v+1)\beta + 2\varepsilon z) \cdot \\ & (v^3(2\varepsilon z + \beta)((16\varepsilon^2 z^2 + 22\beta\varepsilon z + \beta^2) - v\beta(2\varepsilon z + \beta)) - 2(2\varepsilon z + \beta)(32\varepsilon^2 z^2 + 56\beta\varepsilon z + 23\beta^2) \\ & + v(256\varepsilon^3 z^3 + 544\beta\varepsilon^2 z^2 + 352\beta^2\varepsilon z + 71\beta^3) - v^2(176\varepsilon^3 z^3 + 360\beta\varepsilon^2 z^2 + 216\beta^2\varepsilon z + 41\beta^3)) \end{aligned}$$

ist.

Das Vorzeichen der Summe $D + E$ zu ermitteln, ist aufgrund der recht komplexen Gestalt der Summanden und vor allem aufgrund des störenden Wurzelterms in D sehr mühsam. Behelfen kann man sich hier jedoch mit einem Trick: Anstatt wirklich die Summe $D + E$ zu betrachten, betrachtet man zunächst nur die Differenz $D^2 - E^2$, die sich - auch wenn es die Einzelbestandteile nicht vermuten lassen würden - nämlich noch recht kompakt darstellen lässt. Wenn für diese Differenz nun gelten sollte, dass sie größer Null ist, so würde das Vorzeichen der Summe ($D + E$) dem Vorzeichen von D entsprechen. Sollte die Differenz $D^2 - E^2$ hingegen kleiner als Null sein, so würde das Vorzeichen von E das Vorzeichen der Summe bestimmen. Gleichbedeutend mit der Ermittlung des Vorzeichens von $D + E$ ist also die Überprüfung des Vorzeichens von $D^2 - E^2$ und der anschließenden Ermittlung des Vorzeichens des bestimmten Summanden.

Für $D^2 - E^2$ erhält man:

$$D^2 - E^2 = v^2(2\varepsilon z + (v+1)\beta)(2\varepsilon z(v-2) + \beta(v-3))^4 \cdot F$$

mit

$$F = 48\beta\varepsilon^2v^2z^2 - 204\beta^2\varepsilon v z - \beta^3v^3 - 7\beta^3v^2 - 47\beta^3v + 23\beta^3 \\ + 32\varepsilon^3v^2z^3 - 128\varepsilon^3vz^3 + 64\varepsilon^3z^3 + 102\beta^2\varepsilon z + 144\beta\varepsilon^2z^2 - 288\beta\varepsilon^2vz^2 + 6\beta^2\varepsilon v^2z$$

Die Faktoren v^2 , $(2\varepsilon z(v-2) + \beta(v-3))^4$ und $(2\varepsilon z + (v+1)\beta)$ sind alle größer oder gleich Null (letzterer wegen Behauptung 2.7). Der Faktor F ist negativ. Dies zeigt man durch eine Abschätzung des Ausdrucks nach oben:

$$F = 48\beta\varepsilon^2v^2z^2 - 204\beta^2\varepsilon v z - \beta^3v^3 - 7\beta^3v^2 - 47\beta^3v + 23\beta^3 \\ + \underbrace{32\varepsilon^3z^3}_{\leq 0} \underbrace{((v-3)^2 + (2v-7))}_{> 0, \text{ wenn } v \geq 4} + \underbrace{\beta\varepsilon z}_{< 0} \underbrace{(102b + 144\varepsilon z)}_{> 0 \text{ wegen 2.7}} - \underbrace{288\beta\varepsilon^2vz^2}_{< 0} + \underbrace{6\beta^2\varepsilon v^2z}_{< 0} \\ < 48\beta\varepsilon^2v^2z^2 - 204\beta^2\varepsilon v z - \beta^3v^3 - 7\beta^3v^2 - 47\beta^3v + 23\beta^3 \\ \propto 48\varepsilon^2v^2z^2 - 204\beta\varepsilon v z - \beta^2v^3 - 7\beta^2v^2 - 47\beta^2v + \underbrace{23\beta^2}_{< 92} \quad (f) \\ < 48\varepsilon^2v^2z^2 - 204\beta\varepsilon v z - \beta^2v^3 - 7\beta^2v^2 - 47\beta^2v + 92 \\ \leq 48\frac{v^2z^2}{k^2} + 204\frac{(k-1)vz}{k^2} - \frac{(k-1)^2v^3}{k^2} - 7\frac{(k-1)^2v^2}{k^2} - 47\frac{(k-1)^2v}{k^2} + 92 \quad (g) \\ \leq 12\frac{v^2(k-3)^2}{k^2} + 102\frac{(k-1)v(k-3)}{k^2} - \frac{(k-1)^2v^3}{k^2} - 7\frac{(k-1)^2v^2}{k^2} - 47\frac{(k-1)^2v}{k^2} + 92 \quad (h) \\ \leq -\frac{(k-3)^4 + 5(k-4)^3 + 10(k-1)^2 + 3(3k-10)}{k} \quad (i) \\ < 0 \text{ für } k \geq 4.$$

An den mit Buchstaben gekennzeichneten Stellen gilt die Abschätzung jeweils wegen einem der folgenden Argumente:

- Bei (f) wurde ausgenutzt, dass β immer positiv ist (siehe auch Behauptung 2.4), und das Vorzeichen des Ausdrucks sich also nicht ändert, wenn man durch β teilt.
- Bei (g) wurde der Umstand genutzt, dass alle Summanden monoton in λ fallen (siehe hierzu auch Behauptung 2.6). Für $\lambda = 0$ erreichen alle Summanden also ihren höchsten Wert und die Summe dieser Maximalwerte kann als Abschätzung verwendet werden.
- Auch bei (h) wurde eine offensichtliche Monotonie ausgenutzt, diesmal jedoch die in z . Da nur Schnittpunkte $x_{z|z+1}$ mit $z \leq \frac{k-3}{2}$ betrachtet werden, lässt sich also durch Einsetzen von $z = \frac{k-3}{2}$ eine Obergrenze finden.
- Bei (i) wurde schließlich benutzt, dass der Ausdruck monoton in v fallend ist und der kleinstmögliche Wert für v genau $v = k$ entspricht. Dementsprechend erreicht der Ausdruck also auch für $v = k$ sein Maximum und diese Ersetzung kann zur Abschätzung nach oben benutzt werden.

Die angesprochene Monotonie in v lässt sich durch die Betrachtung der Ableitungen des bei (h) gezeigten Ausdrucks nachweisen:

$$3. \text{ Ableitung des Ausdrucks nach } v: \frac{\partial(h)}{\partial^3 v} \leq -\frac{6(k-1)^2}{k^2} < 0$$

↪ 2. Ableitung des Ausdrucks nach v wird maximal für $v = k$:

$$\frac{\partial(h)}{\partial^2 v} \leq -\frac{6(k-2)^3 + 14(k-2)^2 + (106k-210)}{k^2} < 0$$

↪ 1. Ableitung des Ausdrucks nach v wird maximal für $v = k$:

$$\frac{\partial(h)}{\partial v} \leq -\frac{3(k-2)^4 + 8k^2(k-1) + (208k-307)}{k^2} < 0$$

Da in dieser Arbeit nur die Situation $k \geq 4$ untersucht wird, ist somit gezeigt, dass F und damit auch die Differenz $D^2 - E^2$ negativ ist. Es genügt daher, das Vorzeichen von E zu bestimmen, um das Vorzeichen des Ausdrucks $D + E$ zu erfahren.

Der Ausdruck E selbst ist aber das Produkt der folgenden beiden Faktoren:

$$E = -((v+1)\beta + 2\varepsilon z) \cdot \tilde{E},$$

wobei

$$\tilde{E} = \overbrace{v^3(2\varepsilon z + \beta)((16\varepsilon^2 z^2 + 22\beta\varepsilon z + \beta^2) - v\beta(2\varepsilon z + \beta))}^{(I)} \overbrace{-2(2\varepsilon z + \beta)(32\varepsilon^2 z^2 + 56\beta\varepsilon z + 23\beta^2)}^{(II)} \\ + \underbrace{v(256\varepsilon^3 z^3 + 544\beta\varepsilon^2 z^2 + 352\beta^2\varepsilon z + 71\beta^3) - v^2(176\varepsilon^3 z^3 + 360\beta\varepsilon^2 z^2 + 216\beta^2\varepsilon z + 41\beta^3)}_{(III)}$$

ist.

Der erste Faktor dieses Produkts, $-((v+1)\beta + 2\varepsilon z)$, ist wegen Behauptung 2.7 und des Minuszeichens vor dem Ausdruck offensichtlich negativ. Dass aber auch der zweite Faktor, \tilde{E} , negativ sein muss, lässt sich über seine Aufteilung in die drei Summanden (I), (II) und (III) zeigen. Für diese Summanden gilt nämlich jeweils:

$$(I) = \overbrace{v^3(2\varepsilon z + \beta)}^{>0 \text{ wegen 2.7}} ((16\varepsilon^2 z^2 + 22\beta\varepsilon z + \beta^2) - v\beta(2\varepsilon z + \beta)) \\ \propto 16\varepsilon^2 z^2 + 22\beta\varepsilon z + \beta^2 - v \underbrace{\beta(2\varepsilon z + \beta)}_{>0 \text{ wegen 2.7}} \\ \stackrel{v \geq 4}{\leq} 16\varepsilon^2 z^2 + 22\beta\varepsilon z + \beta^2 - 4\beta(2\varepsilon z + \beta) = 16\varepsilon^2 z^2 + 14\beta\varepsilon z - 3\beta^2 \\ = \underbrace{14\varepsilon z}_{<0} \underbrace{(\beta + \frac{8}{7}\varepsilon z)}_{>0 \text{ wegen 2.7}} - \underbrace{3\beta^2}_{>0} < 0$$

$$(II) = -2(2\varepsilon z + \beta)(32\varepsilon^2 z^2 + 56\beta\varepsilon z + 23\beta^2) \\ = -2(2\varepsilon z + \beta)(16(\beta + \frac{11}{8}\varepsilon z)^2 + 7\beta^2 + 12\beta\varepsilon z + \frac{7}{4}\varepsilon^2 z^2) \\ = -2 \underbrace{(2\varepsilon z + \beta)}_{>0 \text{ wegen 2.7}} \overbrace{(16(\beta + \frac{11}{8}\varepsilon z)^2 + 7\beta^2 + 12\beta\varepsilon z + \frac{7}{4}\varepsilon^2 z^2)}^{>0} \leq 0 \\ \underbrace{\hspace{10em}}_{>0 \text{ wegen 2.7}} \underbrace{\hspace{2em}}_{>0 \text{ wegen 2.7}} \underbrace{\hspace{2em}}_{>0}$$

$$(III) = v(256\varepsilon^3 z^3 + 544\beta\varepsilon^2 z^2 + 352\beta^2\varepsilon z + 71\beta^3) - v^2(176\varepsilon^3 z^3 + 360\beta\varepsilon^2 z^2 + 216\beta^2\varepsilon z + 41\beta^3) \\ \propto 256\varepsilon^3 z^3 + 544\beta\varepsilon^2 z^2 + 352\beta^2\varepsilon z + 71\beta^3 - v(176\varepsilon^3 z^3 + 360\beta\varepsilon^2 z^2 + 216\beta^2\varepsilon z + 41\beta^3) \\ = 256\varepsilon^3 z^3 + 544\beta\varepsilon^2 z^2 + 352\beta^2\varepsilon z + 71\beta^3 \\ - v(41 \underbrace{(\beta + 2\varepsilon z)^3}_{>0 \text{ wegen 2.7}} - \underbrace{\varepsilon z}_{<0} \underbrace{(25(\beta + 2\varepsilon z)^2 + (\beta + 4\varepsilon z)^2 + 4(\beta + 3\varepsilon z)^2)}_{\geq 0}) \\ \underbrace{\hspace{15em}}_{\geq 0}$$

$$\begin{aligned}
& \overset{v>3}{<} 256\varepsilon^3 z^3 + 544\beta\varepsilon^2 z^2 + 352\beta^2\varepsilon z + 71\beta^3 \\
& \quad - 3(41(\beta + 2\varepsilon z)^3 - \varepsilon z(25(\beta + 2\varepsilon z)^2 + (\beta + 4\varepsilon z)^2 + 4(\beta + 3\varepsilon z)^2)) \\
& = -4(13\beta^3 + 74\beta^2\varepsilon z + 134\beta\varepsilon^2 z^2 + 68\varepsilon^3 z^3) \\
& = -4 \underbrace{(13(\beta + 2\varepsilon z)^3)}_{>0 \text{ wegen 2.7}} \underbrace{-2\varepsilon z(2(\beta + 3\varepsilon z)^2)}_{>0} \underbrace{-\beta\varepsilon z)}_{\geq 0} < 0
\end{aligned}$$

Da alle drei Teilsummanden kleiner oder gleich Null sind, ist ihre Summe, \tilde{E} , natürlich erst recht negativ. Wie bereits begründet folgt daraus nun direkt, dass der Ausdruck E und somit auch die Summe $D + E$ positiv sind.

Für den Zähler der Ableitung $\frac{\partial Q_2}{\partial z}$ gilt also zusammengefasst:

$$\underbrace{-\beta\varepsilon}_{>0} \cdot \underbrace{(2\varepsilon z + b)}_{>0 \text{ wegen 2.7}} \cdot \underbrace{(D + E)}_{>0} > 0$$

Da aber auch der Nenner der Ableitung in jedem Fall positiv ist, ist die gesamte Ableitung größer Null und Q_2 (und damit auch der Quotient $\frac{x_{z|z+1}^{(v+1)}}{x_{z|z+1}^{(v)}}$) folglich monoton wachsend in z . Der schnellst wachsende Schnittpunkt ist also der, den man für $z = 0$ erhält: Der Schnittpunkt $x_{0|1}$.

Wenn also das relative Veränderungsverhalten des schnellst wachsenden Schnittpunkts mit dem des langsamst wachsenden Scheitelpunkts verglichen werden soll, müssen die folgenden beiden Punkte betrachtet werden:

$$x_{0|1} = -\frac{\sqrt{v+1}}{v-1+\sqrt{v+1}} \quad \text{und} \quad x_{(1)} = -\frac{(k-3-\frac{1}{v})\beta+2\alpha}{(k+v-5)\beta+2\alpha}$$

(Siehe zur Ermittlung dieser Darstellung der Punkte auch die Beweise zu den Behauptungen 6.1 und 5.2.)

Die relativen Ableitungen dieser beiden Punkte bezüglich v sehen wie folgt aus:

$$r_{0|1} = \frac{\frac{\partial x_{0|1}}{\partial v}}{x_{0|1}} = -\frac{v+3}{2(v+1)(\sqrt{v+1}+v-1)} \quad \text{und} \quad r_1 = \frac{\frac{\partial x_{(1)}}{\partial v}}{x_{(1)}} = -\frac{\beta(v-1)(\beta v(k-3)+2\alpha(v+1)+\beta(k-5))}{v((k+v-5)\beta+2\alpha)(\beta v(k-3-\frac{1}{v})+2\alpha v)}$$

Sowohl r_1 als auch $r_{0|1}$ sind ganz offensichtlich negativ. Für den Quotienten dieser relativen Ableitungen ergibt sich:

$$\frac{r_{0|1}}{r_1} = \frac{v(v+3)((k+v-5)\beta+2\alpha)}{2\beta(v+1)(v-1)(\sqrt{v+1}+v-1)} \cdot \frac{(\beta v(k-3)+2\alpha v-\beta)}{(\beta v(k-3)+2\alpha v+2\alpha+\beta(k-5))}$$

Am Wert dieses Quotienten lässt sich nun ablesen, wer mit wachsendem v schneller wächst, ob Scheitel- oder Schnittpunkt. Für Werte größer 1 wäre es der Schnittpunkt, für Werte kleiner 1 der Scheitelpunkt.

In der obigen Darstellung wurde der Quotient bereits in zwei offensichtlich positive Faktoren G_1 und G_2 zerlegt. Im Folgenden soll gezeigt werden, dass beide Teilfaktoren jeweils kleiner 1 sind und es der Gesamtausdruck dementsprechend erst recht ist.

Für G_2 ist dies schnell gezeigt: Der Term $\beta v(k-3) + 2\alpha v$ taucht hier in Zähler und Nenner auf. Der Summand $(k-5)\beta$, der im Nenner steht, ist für $k \geq 4$ aber größer oder gleich dem $-\beta$, das im Zähler steht. Da auch der Summand 2α aus dem Nenner größer als Null ist, ist klar, dass der Zähler auf jeden Fall kleiner ist als der Nenner, und damit der gesamte Bruch kleiner als 1 sein muss.

Der Faktor G_1 lässt sich nach oben abschätzen, indem man sich zunutze macht, dass nach Voraussetzung $v \geq k$ und nach Behauptung 2.5 für $\lambda \in [0, 1)$ auch $\beta \geq \alpha$ ist. Durch Ersetzen des im Zähler auftauchenden k durch v bzw. α durch β ergibt sich dann die folgende Oberschranke:

$$G_1 \underset{v \geq k}{\leq} \frac{v(v+3)((2v-5)\beta+2\alpha)}{2\beta(v+1)(v-1)(\sqrt{v+1}+v-1)} \underset{\alpha \leq \beta}{\leq} \frac{v(v+3)((2v-5)\beta+2\beta)}{2\beta(v+1)(v-1)(\sqrt{v+1}+v-1)} = \frac{v(v+3)(v-\frac{3}{2})}{(v+1)(v-1)(\sqrt{v+1}+v-1)}$$

Durch einfache Äquivalenzumformungen lässt sich nun zeigen, dass diese Obergrenze für alle relevanten Werte von v mit $v \geq 4$ kleiner als 1 ist:

$$\begin{aligned} & \frac{v(v+3)(v-\frac{3}{2})}{(v+1)(v-1)(\sqrt{v+1}+v-1)} < 1 \\ \Leftrightarrow & v(v+3)(v-\frac{3}{2}) < (v+1)(v-1)\sqrt{v+1} + (v+1)(v-1)^2 \\ \Leftrightarrow & \frac{5v^2-7v-2}{2} < (v+1)(v-1)\sqrt{v+1} \\ \Rightarrow & \frac{(5v^2-7v-2)^2}{4} < (v+1)^3(v-1)^2 \\ \Leftrightarrow & \frac{v}{4}(4v^4 - 21v^3 + 62v^2 - 37v - 24) > 0 \\ \Leftrightarrow & 4v^4 - 21v^3 + 62v^2 - 37v - 24 > 0 \\ \Leftrightarrow & 4(v-2)^4 + 11(v-2)^3 + v(32v-41) > 0 \end{aligned}$$

Mit G_1 und G_2 ist aber auch der Quotient $\frac{r_{01}}{r_1}$ kleiner als 1. Anders ausgedrückt bedeutet dies nun: Selbst der „langsamste“ der betrachteten Scheitelpunkte wächst also im Verhältnis immer noch schneller als der „schnellste“ der betrachteten Schnittpunkte. Damit ist also die Behauptung gezeigt. \square

Die Aussage der Behauptung lässt sich folgendermaßen zusammenfassen: Scheitelpunkte können mit wachsendem v Schnittpunkte überholen, Schnittpunkte aber keine Scheitelpunkte. Mit dieser recht technischen Behauptung kann nun eine ganz praktische Aussage bewiesen werden.

Behauptung 6.3: Sei im Folgenden $i < \frac{k}{2} - 1$:

- Wenn dann für gegebene Werte von λ und v die Optimalstelle bei $x_{i|i+1}$ liegt, so kann sie für größere Werte von v nur an den Stellen $x_{j|j+1}$ und $x_{(j)}$ mit $j \leq i$ liegen.
- Liegt die Optimalstelle für gegebenes λ und v bei $x_{(i)}$, so kann sie für größere Werte von v nur an den Stellen $x_{j_1|j_1+1}$ und $x_{(j_2)}$ mit $j_1 < i$ und $j_2 \leq i$ liegen.

Beweis: Damit für gegebene Parameterwerte $x_{i|i+1}$ die gesuchte Stelle sein kann, muss nach Behauptung 5.6 hierfür gelten, dass

$$H'_{s(i)}(x_{i|i+1}^{(v)}) \geq 0 \text{ und } H'_{s(i+1)}(x_{i|i+1}^{(v)}) \leq 0$$

ist. Es gelten also die folgenden beiden Aussagen (siehe dazu auch den Beweis zu Behauptung 5.3 und bedenke, dass all diese Scheitel- und Schnittpunkte negativ sind, sich Ungleichungen beim Dividieren bzw. Multiplizieren mit ihnen also umdrehen):

$$x_{(i)}^{(v)} \leq x_{i|i+1}^{(v)} \Leftrightarrow \frac{x_{(i)}^{(v)}}{x_{i|i+1}^{(v)}} \geq 1 \quad \text{und} \quad x_{(i+1)}^{(v)} \geq x_{i|i+1}^{(v)} \Leftrightarrow \frac{x_{(i+1)}^{(v)}}{x_{i|i+1}^{(v)}} \leq 1.$$

Da durch die zu Beginn der Behauptung gestellte Forderung $i < \frac{k}{2} - 1$ sichergestellt ist, dass für den Scheitelpunkt $x_{(i)}$ und den Schnittpunkt $x_{i-1|i}$ die Bedingung aus Behauptung 6.2 erfüllt ist, kann diese Behauptung nun entsprechend verwendet werden. Wegen ihr gilt

$$\frac{x_{(i)}^{(v+1)}}{x_{(i)}^{(v)}} < \frac{x_{i|i+1}^{(v+1)}}{x_{i|i+1}^{(v)}} \Leftrightarrow \frac{x_{(i)}^{(v+1)}}{x_{i|i+1}^{(v+1)}} < \frac{x_{(i)}^{(v)}}{x_{i|i+1}^{(v)}}$$

und folglich auch für $v^* > v$ (durch mehrmaliges Anwenden der Behauptung)

$$\frac{x_{(i)}^{(v^*)}}{x_{(i)}^{(v)}} < \frac{x_{i|i+1}^{(v^*)}}{x_{i|i+1}^{(v)}} \Leftrightarrow \frac{x_{(i)}^{(v^*)}}{x_{i|i+1}^{(v^*)}} < \frac{x_{(i)}^{(v)}}{x_{i|i+1}^{(v)}}.$$

Kombiniert man diese Abschätzungen mit den obigen Aussagen, so gelangt man zu den folgenden beiden zu betrachtenden Szenarien: Entweder gilt auch hier, dass

$$x_{(i)}^{(v^*)} \leq x_{i|i+1}^{(v^*)} \Leftrightarrow \frac{x_{(i)}^{(v^*)}}{x_{i|i+1}^{(v^*)}} \geq 1 \quad \text{und} \quad x_{(i+1)}^{(v^*)} \geq x_{i|i+1}^{(v^*)} \Leftrightarrow \frac{x_{(i+1)}^{(v^*)}}{x_{i|i+1}^{(v^*)}} \leq 1$$

ist, oder es tritt der Fall ein, dass

$$x_{(i)}^{(v^*)} \geq x_{i|i+1}^{(v^*)} \Leftrightarrow \frac{x_{(i)}^{(v^*)}}{x_{i|i+1}^{(v^*)}} \leq 1 \quad \text{und} \quad x_{(i+1)}^{(v^*)} \geq x_{i|i+1}^{(v^*)} \Leftrightarrow \frac{x_{(i+1)}^{(v^*)}}{x_{i|i+1}^{(v^*)}} \leq 1.$$

In ersterem Fall gilt äquivalent zur obigen Szenariobeschreibung

$$H'_{s_{(i)}^{(v^*)}}(x_{i|i+1}^{(v^*)}) \geq 0 \quad \text{und} \quad H'_{s_{(i+1)}^{(v^*)}}(x_{i|i+1}^{(v^*)}) \leq 0,$$

d.h. die Optimalstelle liegt nach Behauptung 5.6 für v^* am gleichen Schnittpunkt wie für v .

Im zweiten Fall gilt

$$H'_{s_{(i)}^{(v^*)}}(x_{i|i+1}^{(v^*)}) \leq 0 \quad \text{und} \quad H'_{s_{(i+1)}^{(v^*)}}(x_{i|i+1}^{(v^*)}) \leq 0,$$

d.h. die Optimalstelle muss nun rechts von dem Schnittpunkt $x_{i|i+1}$ liegen, da links davon die Funktionswerte ja noch höher sind. Aufgrund von Behauptung 5.5 können rechts des Schnittpunkts aber überhaupt nur Sequenzen $s_{(z)}$ mit $z \leq i$ beziehungsweise ihre Scheitel- und Schnittpunkte in Frage kommen. Damit ist der erste Teil der Behauptung also gezeigt.

Der zweite Teil der Behauptung folgt ganz analog: Wenn für ein v die gesuchte Optimalstelle bei $x_{(i)}$ liegt, so gilt nach Behauptung 5.6 auch

$$H'_{s_{(i)}^{(v)}}(x_{i|i+1}^{(v)}) < 0 \quad \text{und} \quad H'_{s_{(i-1|i)}^{(v)}}(x_{i-1|i}^{(v)}) > 0$$

und damit

$$x_{(i)}^{(v)} > x_{i|i+1}^{(v)} \Leftrightarrow \frac{x_{(i)}^{(v)}}{x_{i|i+1}^{(v)}} < 1 \quad \text{und} \quad x_{(i)}^{(v)} < x_{i-1|i}^{(v)} \Leftrightarrow \frac{x_{(i)}^{(v)}}{x_{i-1|i}^{(v)}} > 1.$$

Da wegen Behauptung 6.2 nun aber wiederum gilt, dass

$$\frac{x_{(i)}^{(v+1)}}{x_{i|i+1}^{(v+1)}} < \frac{x_{(i)}^{(v)}}{x_{i|i+1}^{(v)}} \quad \text{und} \quad \frac{x_{(i)}^{(v+1)}}{x_{i-1|i}^{(v+1)}} < \frac{x_{(i)}^{(v)}}{x_{i-1|i}^{(v)}},$$

folgen für $v^* > v$ auch hier wieder zwei Szenarien. Beim ersten der beiden Szenarien ist

$$H'_{s(i)}(x_{i|i+1}^{(v^*)}) < 0 \quad \text{und} \quad H'_{s(i)}(x_{i-1|i}^{(v^*)}) > 0$$

und die gesuchte Stelle dementsprechend nach 5.6 wieder der Scheitelpunkt der Sequenz $s(i)$.

Beim Alternativszenario muss gelten

$$H'_{s(i)}(x_{i|i+1}^{(v^*)}) < 0 \quad \text{und} \quad H'_{s(i)}(x_{i-1|i}^{(v^*)}) < 0,$$

was zusammengenommen aufgrund von Behauptung 4.4 einfach der zweiten Teilbedingung

$$H'_{s(i)}(x_{i-1|i}^{(v^*)}) < 0$$

entspricht. Wenn die Ableitung der Sequenz $s(i)$ am Schnittpunkt $x_{i-1|i}$ aber fällt, so kann die Optimalstelle nicht links dieses Schnittpunkts liegen. In Frage kommen also nur die Schnitt- und Scheitelpunkte größer oder gleich $x_{i-1|i}$ und das sind nach Behauptung 5.5 genau die im zweiten Teil der Behauptung genannten mit Ausnahme von $x(i)$. Fasst man beide Szenarien zusammen, so ist damit auch der zweite Teil der Behauptung gezeigt und der Beweis abgeschlossen. \square

Umgangssprachlich ausgedrückt lässt sich die Aussage von Behauptung 6.3 also als „Monotonie in v “ verstehen: Je größer der Wert von v , desto niedriger die Anzahl eingesetzter Kontrollen in den zugehörigen optimalen Sequenzen. Die in der Behauptung geforderte Beschränkung auf Punkte $x(i)$ bzw. $x_{i|i+1}$ mit $i < \frac{k}{2} - 1$ stellt dabei nur scheinbar eine Einschränkung dar: Wie nämlich in den letzten beiden Behauptungen dieser Arbeit, 6.4 und 6.5, zu sehen sein wird, sind für $k \geq 5$ und $v = k$ überhaupt nur solche Stellen relevant, d.h. die Behauptung gilt ab dieser Mindestblocklänge von 5 immer.

Ein vergleichbares Monotonieresultat wie das in Behauptung 6.3 für v gezeigte existiert übrigens für steigende Werte von λ nicht, wie auch Abbildung 2 auf Seite 55 beweist. Zur Erstellung dieser Abbildung wurde für festes k und verschiedene Parameterkombinationen aus λ und v mit dem in Behauptung 5.6 vorgestellten Verfahren nach den optimalen Sequenzen gesucht. Welche Stelle $\operatorname{argmin}_x \max_s H_s(x)$ für welche Kombination der Parameter jeweils optimal war, ist in der Abbildung vermerkt.

Wie man an dieser Grafik erkennt, scheint es für die meisten Werte von v nun zwar tatsächlich so etwas wie eine „Monotonie in λ “ in dem Sinne zu geben, dass mit wachsenden λ meist auch die Anzahl eingesetzter Kontrollen bei den jeweiligen optimalen Sequenzen sinkt, nur gilt dies aber eben nicht für alle Werte von v : Bei dem hier betrachteten $k = 5$ ist für $v = 11$ der Scheitelpunkt der Sequenz $s(1)$ sowohl für sehr kleine λ -Werte als auch für Werte um 0.65 herum die Optimalstelle, nicht aber für dazwischenliegende λ -Werte wie zum Beispiel $\lambda = 0.4$.

Im Folgenden soll nun mit Hilfe von Behauptung 6.2 untersucht werden, wie viele Kontrollen die optimalen Sequenzen für gegebenes k höchstens einsetzen. Diese Untersuchung ist auf zwei Behauptungen aufgeteilt: In Behauptung 6.4 wird ein allgemeines Ergebnis für Blocklängen k

von mindestens 13 hergeleitet. In Behauptung 6.5 werden dann in Einzelbetrachtung die noch zu untersuchenden, aber von Behauptung 6.4 nicht erfassten Fälle $k = 5$ bis $k = 12$ überprüft.

Behauptung 6.4: Für $i \geq 4$ und $k \in [(i-1)^2 + i, i^2 + i]$ kommen als gesuchte Stelle nur Schnittpunkte $x_{j_1-1|j_1}$ mit $j_1 \leq i$ und Scheitelpunkte $x_{(j_2)}$ mit $j_2 < i$ in Frage. Die Obergrenze für die Anzahl einzusetzender Kontrollen liegt demnach bei i .

Beweis: Notwendige Bedingung dafür, dass es sich bei einem Scheitelpunkt $x_{(i)}$ oder einem Schnittpunkt $x_{i|i+1}$ um die gesuchte Stelle handeln kann, ist, dass am Schnittpunkt $x_{i-1|i}$ die zu Sequenz $s_{(i)}$ gehörende H-Funktion eine nicht-negative Ableitung hat. Dies sieht man, indem man sich überlegt, welche Auswirkungen die Alternative hätte. Es gilt nämlich:

$$H_{s_{(i)}}(x_{i-1|i}) < 0 \implies H_{s_{(i)}}(x_{i|i+1}) < 0,$$

da wegen Behauptung 4.4 $x_{i|i+1} < x_{i-1|i}$ gilt und es sich bei allen noch betrachteten H-Funktionen um nach oben geöffnete Parabeln handelt (siehe Behauptung 2.2).

Alleine die Tatsache, dass $H_{s_{(i)}}(x_{i-1|i}) < 0$ ist, reicht aber nun schon aus, um den Scheitelpunkt $x_{(i)}$ aus dem Kreis der Kandidatenpunkte auszuschließen, da er damit nämlich nicht die Bedingung aus Behauptung 5.6 erfüllen kann. Aufgrund der aus der Ungleichung gezogenen Folgerung kann aber auch der Schnittpunkt $x_{i|i+1}$ die an ihn nach Behauptung 5.6 gestellten Bedingungen nicht erfüllen und scheidet ebenfalls aus.

Um zu zeigen, dass die in der Behauptung genannten Obergrenzen tatsächlich zutreffen, genügt es also zu zeigen, dass

$$H_{s_{(i)}}(x_{i-1|i}) < 0$$

ist, falls $i \geq 4$ und $k \in [(i-1)^2 + i, i^2 + i]$.

Diese Untersuchung kann man sich nun aber unter Berücksichtigung der Ergebnisse aus Behauptung 6.2 noch ein wenig erleichtern. Für einen Schnittpunkt $x_{i-1|i}$ und einen Scheitelpunkt $x_{(i)}$ gilt für $i < \frac{k}{2}$ nach dieser Behauptung nämlich

$$\frac{x_{(i)}^{(v+1)}}{x_{(i)}^{(v)}} < \frac{x_{i-1|i}^{(v+1)}}{x_{i-1|i}^{(v)}} \quad \text{bzw. äquivalent dazu} \quad \frac{x_{(i)}^{(v+1)}}{x_{i-1|i}^{(v+1)}} < \frac{x_{(i)}^{(v)}}{x_{i-1|i}^{(v)}}.$$

Dass in der hier betrachteten Situation die Bedingung erfüllt ist und $i < \frac{k}{2}$ gelten muss, sieht man anhand von folgender Umformung, bei der man für k den kleinstmöglichen vorgesehenen (und damit ungünstigsten) Wert von $(i-1)^2 + i$ einsetzt:

$$i < \frac{k}{2} \Leftrightarrow i < \frac{(i-1)^2 + i}{2} \Leftrightarrow 0 < \frac{1}{2}i^2 - \frac{3}{2}i + \frac{1}{2} \Leftrightarrow 0 < \frac{1}{2}i(i-3) + \frac{1}{2}.$$

Beachte dabei, dass im letzten Ausdruck für $i \geq 3$ nur positive Summanden stehen. Obige Ungleichung lässt sich also anwenden und man kann damit analog zum Beweis zu Behauptung 6.3 weiter folgern, dass

$$H_{s_{(i)}^{(v)}}(x_{i-1|i}^{(v)}) < 0 \Leftrightarrow x_{i-1|i}^{(v)} < x_{(i)}^{(v)} \Leftrightarrow \frac{x_{(i)}^{(v)}}{x_{i-1|i}^{(v)}} < 1 \Leftrightarrow \frac{x_{(i)}^{(v+1)}}{x_{i-1|i}^{(v+1)}} < 1 \Leftrightarrow H_{s_{(i)}^{(v+1)}}(x_{i-1|i}^{(v+1)}) < 0.$$

Gilt die Aussage also für ein v , so gilt sie folglich erst recht für alle nachfolgenden $v^{(*)} > v$. Die Gültigkeit der Aussage muss also nur noch für den „Worst Case“ $v = k$ gezeigt werden. Für diesen erhält man nun mit

$$m_2(s_{(i)}) = \frac{1}{k}(i + i\lambda^2 + i^2\varepsilon) \quad m_1(s_{(i)}) = \frac{1}{k-1}(2\alpha + (k-2-i)\beta - m_2(s_{(i)}))$$

und

$$x_{i-1|i} = \frac{-(k-1)\sqrt{(2\varepsilon(i-1)+\beta)(\beta(k+1)+2\varepsilon(i-1))+\beta(k+1)+2\varepsilon(i-1)}}{2\varepsilon(i-1)(k-2)k+\beta(k-3)k}$$

(siehe dazu auch die Beweise zu den Behauptungen 4.1 und 4.4) für $H_{s_{(i)}}(x_{i-1|i})$:

$$H_{s_{(i)}}(x_{i-1|i}) = -\frac{2}{k(\beta(k-3)+2\varepsilon(i-1)(k-2))} \cdot (h_1 - h_2),$$

wobei

$$h_1 = \sqrt{(\beta + 2\varepsilon(i-1))((k+1)\beta + 2\varepsilon(i-1))(2\alpha + (k-2-i)\beta + i(k-2)(\beta + (i-1)\varepsilon))}$$

$$h_2 = (k-1)((\beta + 2\varepsilon(i-1))(2\alpha + \beta(k-2)) - \beta i \varepsilon(i-1)).$$

Der Faktor $\frac{2}{k(\beta(k-3)+2\varepsilon(i-1)(k-2))}$ ist immer positiv. Dies sieht man, indem man den Ausdruck im Nenner näher betrachtet. Dabei kann man zunächst seine offensichtliche Monotonie in λ ausnutzen und dann verwenden, dass k in dieser Behauptung $\geq (i-1)^2 + i$ ist:

$$\beta(k-3) + 2\varepsilon(i-1)(k-2) \stackrel{\text{Beh. 2.4}}{\geq} \frac{(k-1)(k-3)}{k} - \frac{2(i-1)(k-2)}{k} > \frac{(k-2)(k-3)}{k} - \frac{2(i-1)(k-2)}{k}$$

$$\propto k-3-2(i-1) \geq (i-1)^2 + i - 3 - 2i + 2 = i(i-3)$$

Diese Untergrenze ist für $i \geq 4$ aber positiv und dementsprechend auch der gesamte Faktor.

Weniger einfach ist es, das Vorzeichen der Differenz $h_1 - h_2$ zu bestimmen. Aufgrund der Behauptungen 2.4 und 2.7 ist klar, dass beide Terme größer oder gleich Null sind. Dadurch ist das Problem der Bestimmung des Vorzeichens ihrer Differenz zwar nicht aus dem Weg geräumt, man kann nun allerdings benutzen, dass wenn $h_1 - h_2 \geq 0$ ist, auch die Differenz $h_1^2 - h_2^2 \geq 0$ sein muss, und sich auf dieser Weise des Wurzelausdrucks entledigen. Durch Quadrieren und Aufteilen des Ergebnisses in zwei Summanden ergibt sich:

$$h_1^2 - h_2^2 = t_1 + t_2$$

mit

$$t_1 = \frac{1}{3}(\beta + 2\varepsilon(i-1))((k+1)\beta + 2\varepsilon(i-1))(2\alpha + (k-2-i)\beta)^2 - (k-1)^2\beta^2i^2\varepsilon^2(i-1)^2$$

und

$$t_2 = (\beta + 2\varepsilon(i-1)) \cdot [((k+1)\beta + 2\varepsilon(i-1)) \cdot$$

$$\left(\frac{2}{3}(2\alpha + (k-2-i)\beta)^2 + 2(2\alpha + (k-2-i)\beta)(k-2)i(\beta + \varepsilon(i-1)) + i^2(k-2)^2(\beta + (i-1)\varepsilon)^2\right)$$

$$- (k-1)^2(2\alpha + \beta(k-2))((\beta + 2\varepsilon(i-1))(2\alpha + \beta(k-2)) - 2\beta i \varepsilon(i-1))]$$

Im Folgenden wird nun gezeigt werden, dass beide Summanden t_1 und t_2 positiv sind. Begonnen wird dabei mit einer Abschätzung des Summanden t_1 .

$$\begin{aligned}
& + \underbrace{\varepsilon^2(k-2)^2i^2(i-1)^2}_{>0} + \underbrace{\alpha\beta(4i(k-2) - 2(k-1)^2)}_{\leq\beta^2} + \underbrace{\alpha\varepsilon(i-1)(4i(k-2) - 4(k-1)^2)}_{\leq 0} \\
> & \beta\varepsilon(i-1)(2(k-2)^2i - 2(k-2)i^2 + 2(k-2)^2i^2 - 2(k-1)^2(k-2) + 2(k-1)^2i) \\
& + \beta^2(2(k-2)^2i - 2(k-2)i^2 + (k-2)^2i^2 - (k-1)^2(k-2)) + \frac{2}{3}(2\alpha + (k-2-i)\beta)^2 \\
& + \beta^2(4i(k-2) - 2(k-1)^2) + \underbrace{\alpha\varepsilon(i-1)(4i(k-2) - 4(k-1)^2)}_{<\frac{1}{2}\beta\varepsilon<0} \\
\geq & \beta\varepsilon(i-1)(2(k-2)^2i - 2(k-2)i^2 + 2(k-2)^2i^2 - 2(k-1)^2(k-2) + 2(k-1)^2i) \\
& + \beta^2(2(k-2)^2i - 2(k-2)i^2 + (k-2)^2i^2 - (k-1)^2(k-2)) + \frac{2}{3}(2\alpha + (k-2-i)\beta)^2 \\
& + \beta^2(4i(k-2) - 2(k-1)^2) + \beta\varepsilon(i-1)(2i(k-2) - 2(k-1)^2) \\
= & \underbrace{\beta^2(2(k-2)^2i + 4i(k-2) - 2(k-2)i^2 + (k-2)^2i^2 - (k-1)^2k)}_{t_{2,1}} \\
& + \underbrace{\beta\varepsilon(i-1)2((k-2)^2i - (k-2)i^2 + (k-2)^2i^2 + (k-1)^2i + i(k-2) - (k-1)^3)}_{t_{2,2}} \\
& + \underbrace{\frac{2}{3}(2\alpha + (k-2-i)\beta)^2}_{t_{2,3}}
\end{aligned}$$

Dabei wurde an einigen Stellen das Resultat von Behauptung 2.5 verwendet: β ist immer größer oder gleich α , aber nie größer als 2α . Berücksichtigt man, dass zudem ε nach Behauptung 2.4 < 0 ist, so ergibt sich daraus auch $\alpha > \frac{1}{2}\beta > 0 \Rightarrow \alpha\varepsilon < \frac{1}{2}\beta\varepsilon < 0$.

Alle drei Teilsummanden von t_2 sind monoton wachsend in λ . Dies sieht man wie folgt:

- Bei $t_{2,1}$ hat der Vorfaktor von β^2 die folgende zweite Ableitung nach k :

$$-6k + 2i^2 + 4i + 4 \underset{k \geq (i-1)^2 + i}{\leq} -4((i-2)^2 + \frac{1}{2}(3i-7)) < 0 \text{ für } i \geq 3.$$

Der Vorfaktor ist also konkav in k und nimmt sein Minimum daher am Rand des zu betrachtenden Bereichs an, d.h. entweder für $k = (i-1)^2 + i$ oder $k = i^2 + 1$. Für die beiden Randstellen ergibt sich aber:

$$\begin{aligned}
k = (i-1)^2 + i & \rightsquigarrow i(3(i-2)^4 + 13(i-2)^3 + 15(i-2)^2 + 2(i-3)) > 0 \text{ für } i \geq 3 \\
k = i^2 + 1 & \rightsquigarrow i((i-2)^4 + 6(i-2)^3 + 7(i-4)^2 + (21i-81)) > 0 \text{ für } i \geq 4
\end{aligned}$$

Da der Vorfaktor also an beiden Rändern nicht-negativ ist, kann er aufgrund seiner Konkavität auch dazwischen nicht negativ sein. Da nach Behauptung 2.6 β^2 monoton wachsend in λ ist, folgt somit auch die Monotonie des ganzen Teilsummanden $t_{2,1}$.

- Der Vorfaktor von $\beta\varepsilon$ im zweiten Teilsummanden $t_{2,2}$ hat als zweite Ableitung nach k :

$$-4(i-1)(3k - i^2 - 2i - 3) \underset{k \geq (i-1)^2 + i}{\leq} -4(i-1)i(2i-5) < 0 \text{ für } i \geq 3$$

Auch dieser Vorfaktor ist also konkav in k und nimmt sein Minimum daher für einen der Randwerte von k an. Für diese Randstellen erhält man:

$$\begin{aligned}
k = (i-1)^2 + i & \rightsquigarrow 2(i-1)i^2(3i^2(i-3) + 5i + 3) > 0 \text{ für } i \geq 3 \\
k = i^2 + 1 & \rightsquigarrow 2(i-1)^2((i-1)^4 + 5(i-1)^3 + 7(i-1)^2 + 2(i-2)) > 0 \text{ für } i \geq 2
\end{aligned}$$

Da der Ausdruck an beiden Rändern aber immer noch positiv ist, ist er es auch im gesamten dazwischenliegenden Bereich. Aufgrund der ebenfalls in Behauptung 2.7 gezeigten Monotonie von $\beta\varepsilon$ in λ ist somit auch $t_{2,2}$ monoton in λ .

- Bei $t_{2,3}$ folgt die Monotonie direkt aus der Tatsache, dass nach Behauptung 2.4 α und β beide positiv und monoton in λ sind.

Der ganze Ausdruck $t_2 = t_{2,2} + t_{2,2} + t_{2,3}$ lässt sich also weiter nach unten abschätzen, indem man für α , β und ε die Werte einsetzt, die sich für sie für $\lambda = 0$ ergeben. Daraus folgt:

$$t_2 \geq -\frac{k-1}{3k^2} \cdot \overbrace{(3k^4 - 3i^2k^3 - 12ik^3 - 5k^3 + 6i^3k^2 + 27i^2k^2 + 28ik^2 - 7k^2 - 30i^3k - 44i^2k - 4ik + 15k + 36i^3 + 8i^2 - 12i - 6)}^{(*)}$$

Der erste Faktor dieser Untergrenze für t_2 , $-\frac{k-1}{3k^2}$, ist auf jeden Fall negativ. Dass aber auch der Ausdruck (*) kleiner Null sein muss, sieht man wie folgt:

$$\begin{aligned} \frac{\partial(*)}{\partial^3 k} &= 72k - (18i^2 + 72i + 30) \stackrel{k \geq (i-1)^2 + i}{\geq} 6(3i-7)(3i-1) \geq 0 \text{ für } i \geq \frac{7}{3} \\ \Rightarrow \frac{\partial(*)}{\partial^2 k} &\text{ monoton wachsend in } k \\ \Rightarrow \frac{\partial(*)}{\partial^2 k} &= 2(18k^2 - 9i^2k - 36ik - 15k + 6i^3 + 27i^2 + 28i - 7) \stackrel{k \geq (i-1)^2 + i}{\geq} 2(9i^4 - 57i^3 + 93i^2 - 29i - 4) \\ &= 2(9(i-4)^4 + 87(i-4)^3 + 273(i-4)^2 + (283i - 1108)) \geq 0 \text{ für } i \geq 4 \\ \Rightarrow (*) &\text{ ist konvex in } k \text{ und nimmt das Maximum an einer der beiden Randstellen an.} \end{aligned}$$

An beiden Stellen sind die Werte aber negativ:

(*) für $k = (i-1)^2 + i$:

$$-(i-1)i(9(i-2)^5 + 41(i-2)^4 + 41(i-3)^3 + 93(i-4)^2 + 197i - 747) < 0 \text{ für } i \geq 4$$

(*) für $k = i^2 + i$:

$$\begin{aligned} &-(3(i-1)^7 + 14(i-2)^6 + 89(i-3)^5 + 622(i-3)^4 + 1592(i-4)^3 + 6237(i-4)^2 \\ &+ 7275i - 27546) < 0 \text{ für } i \geq 4 \end{aligned}$$

Da der Ausdruck (*) selbst an seinen Maximalstellen für $i \geq 4$ negativ ist, ist er es auch an allen anderen Stellen. Somit ist gezeigt, dass t_2 selbst aber immer positiv ist und damit auch $h_1^2 - h_2^2$. Aufgrund des negativen Vorfaktors ist $H_{s(i)}(x_{i-1|i}) = -\frac{2}{k(\beta(k-3) + 2\varepsilon(i-1)(k-2))} \cdot (h_1 - h_2)$ dementsprechend immer negativ und der Beweis abgeschlossen. \square

Die Obergrenze aus Behauptung 6.4 ließe sich sogar noch weiter verschärfen. Da der Beweis dann allerdings noch unangenehmer werden würde, weil nicht mehr ganz so großzügig abgeschätzt werden könnte, und der Zusatznutzen aber nur sehr gering wäre, wird hier bewusst auf den Beweis dieser verschärften Variante verzichtet und diese Variante nur als Vermutung geäußert. In ihrer schärfer formulierten Fassung müsste Behauptung 6.4 demnach wie folgt lauten:

Für $i \geq 4$ und $k \in [i^2 - 1, i^2 + 2i - 1]$ kommen als gesuchte Stelle nur Schnittpunkte $x_{j_1-1|j_1}$ mit $j_1 \leq i$ und Scheitelpunkte $x_{(j_2)}$ mit $j_2 < i$ in Frage. Die Obergrenze für die Anzahl einzusetzender Kontrollen liegt demnach bei i .

In Tabelle 2 ist für Blocklängen bis knapp über $k = 100$ gegenübergestellt, wie die Obergrenzen der Anzahl verwendeter Kontrollen nach den beiden Varianten von Behauptung 6.4 aussehen.

Behauptung 6.4 greift wie gesehen erst für Blocklängen $k \geq 13$. Die Fälle $k = 3$ und $k = 4$ wurden bei [Kunert et al., 2010] sehr intensiv behandelt und bedürfen somit keiner weiteren Betrachtung

Tabelle 2: Obergrenzen der Zahl zu verwendender Kontrollen für $k \geq 13$

Obergrenze	nach Behauptung 6.4 für k von...	nach verschärfter Variante für k von...
4	13 bis 20	15 bis 23
5	21 bis 30	24 bis 34
6	31 bis 42	35 bis 47
7	43 bis 56	48 bis 62
8	57 bis 72	63 bis 79
9	73 bis 90	80 bis 98
10	91 bis 110	99 bis 119

mehr. Der Vollständigkeit halber sollten allerdings natürlich auch die Fälle $k = 5$ bis $k = 12$ untersucht werden. Darüber, wie in diesen die jeweiligen Obergrenzen aussehen, gibt die folgende Behauptung Aufschluß. Durch diese Behauptung ist dann auch der Nachweis erbracht, dass für $k \geq 5$ Sequenzen mit Randbelegung tatsächlich keine Rolle mehr spielen und Behauptung 6.3 entsprechend uneingeschränkt gültig ist.

Behauptung 6.5: *Auch für $k < 13$ lassen sich Obergrenzen für die Zahl der Kontrollen in den optimalen Sequenzen angeben:*

- Für $k \in [5, 7]$ kommen als gesuchte Stelle nur Schnittpunkte $x_{j_1-1|j_1}$ mit $j_1 \leq 2$ und Scheitelpunkte $x_{(j_2)}$ mit $j_2 < 2$ in Frage. Die Obergrenze für die Anzahl einzusetzender Kontrollen liegt demnach bei 2.
- Für $k \in [8, 12]$ kommen als gesuchte Stelle nur Schnittpunkte $x_{j_1-1|j_1}$ mit $j_1 \leq 3$ und Scheitelpunkte $x_{(j_2)}$ mit $j_2 < 3$ in Frage. Die Obergrenze für die Anzahl einzusetzender Kontrollen liegt demnach bei 3.

Beweis: In diesem Beweis wird die Argumentation von Behauptung 6.4 aufgegriffen, die acht von ihr noch nicht erfassten Fälle $k = 5$ bis $k = 12$ aber als separate Einzelfälle betrachtet.

Wie in Behauptung 6.4 dargelegt, reicht es zu zeigen, dass für $v = k$

$$H_{s(i)}(x_{i-1|i}) < 0$$

ist, um die Gültigkeit der angegebenen Obergrenzen nachzuweisen. Behauptung 6.2 ist nämlich auf den Scheitelpunkt $x_{(i)}$ und den Schnittpunkt $x_{i-1|i}$ wieder anwendbar, da die in der Behauptung genannten Bedingungen für die hier genannten Kombinationen aus i und k jeweils erfüllt sind. Wie ebenfalls in Behauptung 6.4 zu sehen war, genügt es dazu wiederum das Vorzeichen der Differenz $h_1 - h_2$ zu ermitteln, wobei h_1 und h_2 wie folgt aussehen:

$$h_1 = \sqrt{(\beta + 2\varepsilon(i-1))((k+1)\beta + 2\varepsilon(i-1))(2\alpha + (k-2-i)\beta + i(k-2)(\beta + (i-1)\varepsilon))}$$

$$h_2 = (k-1)((\beta + 2\varepsilon(i-1))(2\alpha + \beta(k-2)) - \beta i \varepsilon(i-1)).$$

In dieser Situation lässt sich dies durch Einsetzen der Definition von α , β und ε und der konkreten Werte von i und k erledigen. Es ergibt sich nach geschicktem Zusammenfassen dann für $h_1 - h_2$ jeweils:

Für $k = 5$ und $i = 2$:

$$-\frac{16(-\lambda^3+7\lambda^2-9\lambda-1)}{(3\lambda-5)^4}(171(1-\lambda)^9 + 1962\lambda(1-\lambda)^8 + 10519\lambda^2(1-\lambda)^7 + 32194\lambda^3(1-\lambda)^6 + 60433\lambda^4(1-\lambda)^5 + 71262\lambda^5(1-\lambda)^4 + 52221\lambda^6(1-\lambda)^3 + 22710\lambda^7(1-\lambda)^2 + (5312\lambda^8 - 4800\lambda^9))$$

Für $k = 6$ und $i = 2$:

$$-\frac{(\lambda^3+15\lambda^2-21\lambda-7)}{2(2\lambda-3)^4}(316(1-\lambda)^9 + 3332\lambda(1-\lambda)^8 + 17346\lambda^2(1-\lambda)^7 + 52522\lambda^3(1-\lambda)^6 + 97895\lambda^4(1-\lambda)^5 + 113894\lambda^5(1-\lambda)^4 + 80900\lambda^6(1-\lambda)^3 + 32956\lambda^7(1-\lambda)^2 + (6976\lambda^8 - 6384\lambda^9))$$

Für $k = 7$ und $i = 2$:

$$-\frac{16(3\lambda^3+7\lambda^2-11\lambda-7)}{(5\lambda-7)^4}(100(1-\lambda)^9 + 1440\lambda(1-\lambda)^8 + 11864\lambda^2(1-\lambda)^7 + 46616\lambda^3(1-\lambda)^6 + 100590\lambda^4(1-\lambda)^5 + 126596\lambda^5(1-\lambda)^4 + 92425\lambda^6(1-\lambda)^3 + 36442\lambda^7(1-\lambda)^2 + (7136\lambda^8 - 6592\lambda^9))$$

Für $k = 8$ und $i = 3$:

$$-\frac{(\lambda^3+47\lambda^2-57\lambda-11)}{16(3\lambda-4)^4}(54975(1-\lambda)^9 + 524670\lambda(1-\lambda)^8 + 2358956\lambda^2(1-\lambda)^7 + 6190172\lambda^3(1-\lambda)^6 + 10123200\lambda^4(1-\lambda)^5 + 10375440\lambda^5(1-\lambda)^4 + 6406416\lambda^6(1-\lambda)^3 + 2182944\lambda^7(1-\lambda)^2 + (373504\lambda^8 - 348480\lambda^9))$$

Für $k = 9$ und $i = 3$:

$$-\frac{96(2\lambda^3+12\lambda^2-15\lambda-5)}{(7\lambda-9)^4}(4122(1-\lambda)^9 + 36882\lambda(1-\lambda)^8 + 158419\lambda^2(1-\lambda)^7 + 401910\lambda^3(1-\lambda)^6 + 640360\lambda^4(1-\lambda)^5 + 640390\lambda^5(1-\lambda)^4 + 383119\lambda^6(1-\lambda)^3 + 123074\lambda^7(1-\lambda)^2 + (19424\lambda^8 - 18240\lambda^9))$$

Für $k = 10$ und $i = 3$:

$$-\frac{(17\lambda^3+47\lambda^2-61\lambda-31)}{16(4\lambda-5)^4}(153909(1-\lambda)^9 + 1308258\lambda(1-\lambda)^8 + 5449488\lambda^2(1-\lambda)^7 + 13539516\lambda^3(1-\lambda)^6 + 21244792\lambda^4(1-\lambda)^5 + 20915072\lambda^5(1-\lambda)^4 + 12222672\lambda^6(1-\lambda)^3 + 3723968\lambda^7(1-\lambda)^2 + (545856\lambda^8 - 515328\lambda^9))$$

Für $k = 11$ und $i = 3$:

$$-\frac{32(7\lambda^3+11\lambda^2-15\lambda-11)}{(9\lambda-11)^4}(25968(1-\lambda)^9 + 211248\lambda(1-\lambda)^8 + 864916\lambda^2(1-\lambda)^7 + 2132836\lambda^3(1-\lambda)^6 + 3335715\lambda^4(1-\lambda)^5 + 3265800\lambda^5(1-\lambda)^4 + 1881129\lambda^6(1-\lambda)^3 + 546834\lambda^7(1-\lambda)^2 + (74912\lambda^8 - 71040\lambda^9))$$

Für $k = 12$ und $i = 3$:

$$-\frac{3(41\lambda^3+39\lambda^2-57\lambda-59)}{16(5\lambda-6)^4}(78489(1-\lambda)^9 + 611586\lambda(1-\lambda)^8 + 2507260\lambda^2(1-\lambda)^7 + 6265716\lambda^3(1-\lambda)^6 + 9954896\lambda^4(1-\lambda)^5 + 9843152\lambda^5(1-\lambda)^4 + 5659120\lambda^6(1-\lambda)^3 + 1580704\lambda^7(1-\lambda)^2 + (203648\lambda^8 - 193856\lambda^9))$$

Anhand dieser Darstellungen sieht man, dass all diese Differenzen $h_1 - h_2$ positiv sind: Die Differenz lässt sich wie zuvor zu sehen nämlich immer als Produkt zweier Faktoren schreiben.

Der erste ist jeweils ein mit einem Minuszeichen versehener Bruch, dessen Nenner aufgrund des Exponenten in jedem Fall positiv ist, und dessen Zähler immer negativ ist. Letzteres sieht man durch einen Blick auf die Koeffizienten: Die zu λ^1 und λ^0 gehörenden Koeffizienten haben alle ein negatives Vorzeichen und sind vom Betrag her ihren zu λ^2 und λ^3 gehörenden Pendanten überlegen. Da wegen $0 \leq \lambda < 1$ aber gilt, dass $1 > \lambda > \lambda^2 > \lambda^3$ ist, folgt daraus, dass dieser Zähler kleiner Null sein muss. Da der Zähler aber negativ ist und der Nenner positiv, ist auch der gesamte Bruch negativ und aufgrund des jeweils vor dem Bruch stehenden Minuszeichen der erste Faktor folglich positiv. Auch der zweite Faktor ist aber positiv: Da wegen der gleichen Argumentation wie zuvor natürlich auch gelten muss, dass $\lambda^8 > \lambda^9$, ist klar, dass dieser jeweils aus lauter positiven Summanden besteht und somit gar kein anderes Vorzeichen haben kann.

Da $h_1 - h_2$ als Produkt zweier positiver Faktoren selbst natürlich auch positiv ist, greift dieselbe Argumentation wie in Behauptung 6.4 und der Beweis ist abgeschlossen. \square

Die in Behauptung 6.5 nachgewiesenen Obergrenzen werden mit einer Ausnahme auch tatsächlich angenommen. Die Ausnahme betrifft die Blocklänge $k = 8$: In diesem Fall spielt zwar der Scheitelpunkt $x_{(2)}$ für bestimmte λ -Werte eine Rolle, nicht aber der Schnittpunkt $x_{2|3}$. Verschärfen ließe sich die Behauptung in der gewählten Formulierung aber nicht.

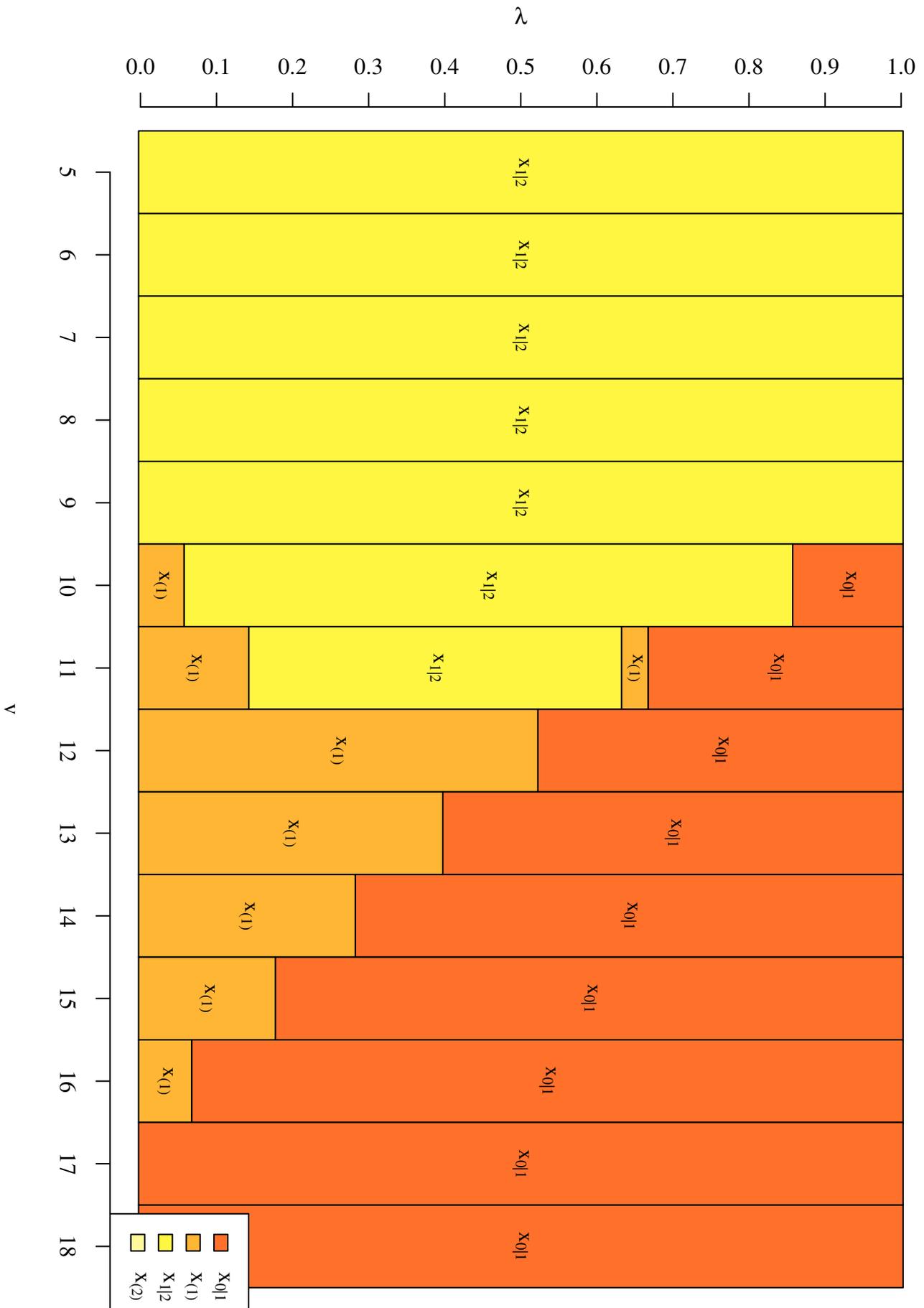


Abbildung 2: Das Aussehen der optimalen Sequenzen für $k = 5$ (λ und v variierend)

7 Zusammenfassung

In dieser Arbeit wurde eine Vorarbeit von [Kunert et al., 2010] aufgegriffen und für den Vergleich mehrerer Behandlungen mit einer Kontrolle im einfachen Blockmodell mit AR(1)-korrelierten Fehlern untersucht, wo sich die Stelle befindet, an der das Maximum über alle H-Funktionen seinen kleinsten Wert annimmt. Das Auffinden dieser Stelle ist deshalb so bedeutsam, da sich durch sie zumindest eine Untergrenze für das A_{tc} -Kriterium ermitteln lässt und sie in manchen Fällen sogar über die Gestalt des optimalen Versuchplans Auskunft gibt. Bei dieser Untersuchung wurde dabei von denselben grundlegenden Annahmen ausgegangen, die auch [Kunert et al., 2010] in ihrem Artikel voraussetzten: So wurde insbesondere nur die Situation betrachtet, in der die Anzahl der Behandlungen v mindestens so groß ist wie die Blocklänge k und die Fehler zudem nicht-negativ und auch nicht perfekt korreliert sind, der Korrelationsparameter λ des autoregressiven Prozesses also im Bereich $[0, 1)$ liegt.

In Kapitel 3 konnte mit einer alternativen Darstellung der Sequenzen schnell gezeigt werden, dass nur solche Sequenzen für diese Suche überhaupt von Relevanz sein können, die Kontrollen nie in direkter Folge einsetzen und den Einsatz einer Kontrolle auf einem der beiden Randfelder der Sequenz nach Möglichkeit vermeiden. Durch diese Erkenntnis und die Vorarbeit von [Kunert et al., 2010] ließ sich die Zahl der zu betrachtenden Sequenzen erheblich reduzieren und es konnte nachgewiesen werden, dass für jede bestimmte Anzahl von Kontrollen z mit $z \leq \frac{k}{2}$ jeweils nur eine einzige Sequenz $s_{(z)}$ als Repräsentant der ganzen Klasse betrachtet werden muss.

Bei der anschließenden näheren Betrachtung des Verhaltens dieser verbleibenden Sequenzen in den Kapiteln 4 und 5 stellte sich heraus, dass die zu ihnen gehörenden H-Funktionen im Intervall $[-1, 0]$, das - wie Vorüberlegungen zeigten - überhaupt den einzig interessanten Bereich abdeckt, besonderen Gesetzmäßigkeiten folgen: Die Scheitelpunkte $x_{(z)}$ dieser Funktionen liegen alle in diesem Bereich und zwar in der immer gleichen Reihenfolge $x_{(0)}, x_{(1)}, x_{(2)}$ und so weiter. Je zwei dieser Funktionen schneiden sich in diesem Intervall genau einmal. Für die Schnittpunkte $x_{j|j+1}$ zweier Sequenzen $s_{(j)}$ und $s_{(j+1)}$ konnte dabei gezeigt werden, dass diese Schnitte umso weiter links im Intervall liegen, je größer j ist, also die Reihenfolge $x_{0|1} > x_{1|2} > x_{2|3}$ usw. gilt. Diese Erkenntnisse vereinend und dabei die Tatsache berücksichtigend, dass es sich bei allen H-Funktionen um nach oben geöffnete Parabeln handelt, ließ sich damit weiter zeigen, dass es für all diese Sequenzen jeweils ein Teilintervall gibt, auf dem die jeweilige Sequenz allen anderen überlegen ist, und dass diese Intervalle umso weiter rechts liegen, je kleiner die Zahl der eingesetzten Kontrollen in der Sequenz.

Mit diesem neu gewonnenen Wissen konnte daraufhin ein konstruktives Verfahren hergeleitet werden, mit dem sich schnell und mit nur wenigen Funktionsauswertungen die gesuchte Optimalstelle finden lässt. Welche Auswirkungen eine Veränderung der Parameter auf die Position dieser Stelle bzw. die Zahl der in den dort optimalen Sequenzen eingesetzten Kontrollen hat, wurde schließlich zum Abschluß dieser Arbeit in Kapitel 6 erforscht. Dabei konnte unter anderem nachgewiesen werden, dass mit wachsendem v die Zahl der in diesen optimalen Sequenzen

verwendeten Kontrollen monoton sinkt und sich Schwellenwerte finden lassen, ab denen die gesuchte Optimalstelle am Schnittpunkt $x_{0|1}$ liegen kann bzw. ab denen sie dort sogar immer liegen muss. Weiter ließ sich so ermitteln, wie viele Kontrollen die optimalen Sequenzen für gegebenes k höchstens einsetzen.

Literatur

[Hildebrandt, 2006]

S. Hildebrandt (2006): Analysis I. Springer-Verlag Berlin, 2. Auflage.

[Kunert und Martin, 2000]

J. Kunert und R.J. Martin (2000): On the determination of optimal designs for an interference model. The Annals of Statistics 28, Seite 1728 - 1742.

[Kunert et al., 2010]

J. Kunert, R.J. Martin und J. Eccleston (2010): A-optimal block designs for the comparison with a control for correlated errors and analysis with the weighted least squares estimate. Journal of Statistical Planning and Inference 140, Seite 2719 - 2738.

[Kushner, 1997]

H. Kushner (1997): Optimal repeated measurements designs: the linear optimality equations. The Annals of Statistics 25, Seite 2328 - 2344.

[Majumdar und Notz, 1983]

D. Majumdar und W. Notz (1983): Optimal incomplete block designs for comparing treatments with a control, The Annals of Statistics 11, Seite 258 - 266.

[Maxima, 2010]

The maxima developers (2010): Maxima, a Computer Algebra System. Version 5.20.1, URL <http://maxima.sourceforge.net>

[R, 2009]

R Development Core Team (2009): R - A language and environment for statistical computing. Version 2.10.1, R Foundation for Statistical Computing, Wien, Österreich. ISBN 3-900051-07-0, URL <http://www.R-project.org>