1 Testing for a rectangular region of interest

1.1 Definitions

Definition 1.1. Let $M, N \in \mathbb{N}$ and $V \in \mathbb{R}^{M\times N}$ be a matrix. Assume there are two pairs of indices $(i_1, j_1), (i_2, j_2)$ with $1 \leq i_1 \leq i_2 \leq M$ and $1 \leq j_1 \leq j_2 \leq N$, such that

$$V(i,j) \neq 0 \text{ if and only if } (i,j) \in \{i_1, \dots, i_2\} \times \{j_1, \dots, j_2\} \tag{1}$$

We call $R = \{i_1, \dots, i_2\} \times \{j_1, \dots, j_2\}$ a *rectangular region of interest (rROI)* and say that V contains the rROI R.
Furthermore, we call (i_1, j_1) the *top left corner* and (i_2, j_2) the *bottom right corner* of R.

Definition 1.2. Let $M, N \in \mathbb{N}$ and $c \in \mathbb{R} \setminus \{0\}$. Let $V \in \mathbb{R}^{M\times N}$ be a matrix, that only takes values in the set $\{0, \pm c\}$ and that contains a rectangular region of interest R. We say that R has a *checkerboard pattern*, if one of the following relations is true:

$$\text{For all } (i,j) \in R : V(i,j) = c \text{ if and only if } i + j \text{ is odd.} \tag{2a}$$

$$\text{For all } (i,j) \in R : V(i,j) = c \text{ if and only if } i + j \text{ is even.} \tag{2b}$$

Remark 1.3. If the first relation in the definition is true, that immediately implies, that

$$\text{for all } (i,j) \in R : V(i,j) = -c \text{ if and only if } i + j \text{ is even,}$$

since V takes only values in $\{0, \pm c\}$, but for $(i,j) \in R$ we have $V(i,j) \neq 0$. Similarly, if the second relation is true, it implies that

$$\text{for all } (i,j) \in R : V(i,j) = -c \text{ if and only if } i + j \text{ is odd.}$$

In both cases the values of V alternate between $+c$ and $-c$ along the rows and columns of R. This is similar to the classical checkerboard pattern.

Definition 1.4. Let $M, N \in \mathbb{N}$ and $c \in \mathbb{R} \setminus \{0\}$. We define $\mathcal{V}_c^{M,N}$ to be the set of all matrices $V \in \mathbb{R}^{M \times N}$, that only take values in the set $\{0, \pm c\}$ and that contain a rectangular region of interest with a checkerboard pattern.

1.2 Statistical model

Let $M, N \in \mathbb{N}$, $c \in \mathbb{R} \setminus \{0\}$ and $G = \{1, \ldots, M\} \times \{1, \ldots, N\}$. Assume we are given noisy data

$$F(i,j) = c + V(i,j) + \varepsilon_{i,j} \tag{3}$$

where $(i,j) \in G$, $V \in \mathcal{V}_c^{M,N}$ and $\varepsilon_{i,j} \sim \mathcal{N}(0, \sigma^2)$ are i.i.d. normal distributed random variables for some $\sigma > 0$ and for all $(i,j) \in G$.

Although defined as a matrix, we often refer to F and V as images and to (i,j) as a pixel. Let R be the rectangular region of interest contained in V and let (i_1, j_1) and (i_2, j_2) be the top left and bottom right corner of R, respectively. We aim to find a statistical test to determine for each individual pixel whether $(i,j) \in R$ or $(i,j) \notin R$.

To proceed we define for each pixel $(i,j) \in G$ the four values

$$D_1^{\pm}(i,j) = V(i \pm 1, j) - V(i,j) \tag{4}$$
$$D_2^{\pm}(i,j) = V(i, j \pm 1) - V(i,j) \tag{5}$$

where we set

$$\begin{aligned} V(i,0) &= V(i,N) \\ V(0,j) &= V(M,j) \\ V(i,N+1) &= V(i,1) \\ V(M+1,j) &= V(1,j) \end{aligned}$$

to adjust to boundary issues. We now combine these values into two new values and assign to each pair $(i,j) \in G$ the values

$$D^{\pm}(i,j) = \sqrt{D_1^{\pm}(i,j)^2 + D_2^{\pm}(i,j)^2} \tag{6}$$

Since we have assumed that $V \in \mathcal{V}_c^{M,N}$, we know that $V(i,j) = 0$ if and only if $(i,j) \notin R$. Now if $(i,j) \notin R$ it follows that $i \notin \{i_1, \ldots, i_2\}$ or $j \notin \{j_1, \ldots, j_2\}$.

We have to distinguish four cases here:

$$\begin{aligned}
i < i_1 &\Rightarrow (i, j-1) \notin R \text{ and } (i-1, j) \notin R \\
&\Rightarrow V(i, j-1) = V(i-1, j) = 0 \\
j < j_1 &\Rightarrow (i-1, j) \notin R \text{ and } (i, j-1) \notin R \\
&\Rightarrow V(i-1, j) = V(i, j-1) = 0 \\
i > i_2 &\Rightarrow (i, j+1) \notin R \text{ and } (i+1, j) \notin R \\
&\Rightarrow V(i, j+1) = V(i+1, j) = 0 \\
j > j_2 &\Rightarrow (i+1, j) \notin R \text{ and } (i, j+1) \notin R \\
&\Rightarrow V(i+1, j) = V(i, j+1) = 0
\end{aligned}$$

We see, that in the first two cases, we have $D_1^-(i,j) = D_2^-(i,j) = 0$, which yields $D^-(i,j) = 0$. In the latter two cases, we get $D_1^+(i,j) = D_2^+(i,j) = 0$ and thus $D^+(i,j) = 0$. Thus, if $(i,j) \notin R$ it follows that $\min\{D^+(i,j), D^-(i,j)\} = 0$.
On the other hand, we have assumed, that R has a checkerboard pattern, thus $D^\pm(i,j) \neq 0$ for $(i,j) \in R$. This yields the equivalence

$$(i,j) \notin R \Leftrightarrow \min\{D^+(i,j), D^-(i,j)\} = 0$$

Since our goal is to test for $(i,j) \in R$, we define a null hypothesis for each individual pixel:

$$H_0(i,j) : \min\{D^+(i,j), D^-(i,j)\} = 0 \tag{7}$$

Unfortunately, we do not the actual values of V, which makes D^+ and D^- non-observable. We are given noisy data though. Based on the noisy data, we define four observable values for each pixel

$$\tilde{D}_1^\pm(i,j) = F(i \pm 1, j) - F(i,j) \tag{8}$$

$$\tilde{D}_2^\pm(i,j) = F(i, j \pm 1) - F(i,j) \tag{9}$$

where we again define

$$\begin{aligned} F(i,0) &= F(i,N) \\ F(0,j) &= F(M,j) \\ F(i,N+1) &= F(i,1) \\ F(M+1,j) &= F(1,j) \end{aligned}$$

to adjust to boundary issues. Again we combine these values into two new values

$$\tilde{D}^{\pm}(i,j) = \sqrt{\tilde{D}_1^{\pm}(i,j)^2 + \tilde{D}_2^{\pm}(i,j)^2} \tag{10}$$

We use these values for our test statistic and test for each individual pixel (i,j) the null hypothesis

$$H_0(i,j) : \min\{D^+(i,j), D^-(i,j)\} = 0 \tag{11}$$

against the alternative hypothesis

$$H_1(i,j) : \min\{D^+(i,j), D^-(i,j)\} \neq 0 \tag{12}$$

using the test statistic

$$T(i,j) = \min\{\tilde{D}^+(i,j), \tilde{D}^-(i,j)\} \tag{13}$$

After having established our hypotheses and test statistic, we want to show, that we can ensure a given statistical significance α. This will be done in two steps: In a nutshell, the testing procedure is the combination of two testing procedures (D^+ and D^-) and we will see, that we can bound our combination by either of the individual ones. In a second step, we will bound the individual procedures by computing their cumulative distribution function.
If the null hypothesis for a given pixel (i,j) is true, we can distiguish the two cases: $D^+(i,j) = 0$ or $D^-(i,j) = 0$. Note that these two cases are not exclusive. Without loss of generality we assume the first case, i.e. $D^+(i,j) = 0$. For $t \in \mathbb{R}^+$ we then have

$$\begin{aligned}\mathbb{P}(T(i,j) \geq t \mid H_0(i,j)) &= \mathbb{P}(\min\{\tilde{D}^+(i,j), \tilde{D}^-(i,j)\} \geq t \mid D^+(i,j) = 0)\\ &= \mathbb{P}(\{\tilde{D}^+(i,j) \geq t\} \cap \{\tilde{D}^-(i,j) \geq t\} \mid D^+(i,j) = 0)\\ &\leq \mathbb{P}(\tilde{D}^+(i,j) \geq t \mid D^+(i,j) = 0)\end{aligned} \tag{14}$$

In the second case we can actually get the same inequality by using the fact that $\mathbb{P}(\tilde{D}^-(i,j) \geq t \mid D^-(i,j) = 0) = \mathbb{P}(\tilde{D}^+(i,j) \geq t \mid D^+(i,j) = 0)$.
We want to take a look at the distribution of $\tilde{D}^+(i,j)$ given $D^+(i,j) = 0$:

$$\begin{aligned}p &:= \mathbb{P}(\tilde{D}^+(i,j) \leq t \mid D^+(i,j) = 0)\\ &= \mathbb{P}((c + V(i+1,j) + \varepsilon_{i+1,j} - c - V(i,j) - \varepsilon_{i,j})^2\\ &+ (c + V(i,j+1) + \varepsilon_{i,j+1} - c - V(i,j) - \varepsilon_{i,j})^2 \leq t^2 \mid D^\pm(i,j) = 0)\\ &= \mathbb{P}\left((\varepsilon_{i+1,j} - \varepsilon_{i,j})^2 + (\varepsilon_{i,j+1} - \varepsilon_{i,j})^2 \leq t^2\right)\end{aligned} \tag{15}$$

Assuming the common term $\varepsilon_{i,j}$ to be constant and by defining

$$X_1 = \varepsilon_{i+1,j} - \varepsilon_{i,j} \sim \mathcal{N}(\varepsilon_{i,j}, \sigma^2)$$
$$X_2 = \varepsilon_{i,j+1} - \varepsilon_{i,j} \sim \mathcal{N}(\varepsilon_{i,j}, \sigma^2)$$

we obtain

$$p = \mathbb{P}\left(\sqrt{\left(\frac{X_1}{\sigma}\right)^2 + \left(\frac{X_2}{\sigma}\right)^2} \leq \frac{t}{\sigma}\right) \tag{16}$$

This shows, that the square root inside has a non-central Chi distribution with two degrees of freedom and non-centrality parameter

$$\lambda = \sqrt{\left(\frac{\varepsilon_{i,j}}{\sigma}\right)^2 + \left(\frac{\varepsilon_{i,j}}{\sigma}\right)^2} = \frac{\sqrt{2}|\varepsilon_{i,j}|}{\sigma} \tag{17}$$

which shows that

$$\sqrt{\left(\frac{X_1}{\sigma}\right)^2 + \left(\frac{X_2}{\sigma}\right)^2} \sim \chi_2\left(\frac{\sqrt{2}|\varepsilon_{i,j}|}{\sigma}\right) \tag{18}$$

Up until this point we assumed $\varepsilon_{i,j}$ to be constant, but it is a normal distributed random variable with zero mean and standard deviation σ. Thus, we have a compound probability distribution:

$$\begin{aligned} p &= \mathbb{P}\left(\sqrt{\left(\frac{X_1}{\sigma}\right)^2 + \left(\frac{X_2}{\sigma}\right)^2} \leq \frac{t}{\sigma}\right) \\ &= \int_0^{\frac{t}{\sigma}} \int_0^\infty \underbrace{x \exp\left(-\frac{x^2}{2} - \frac{\eta^2}{\sigma^2}\right) I_0\left(\frac{\sqrt{2}\eta}{\sigma}x\right)}_{\text{pdf of } \chi_2\left(\frac{\sqrt{2}\eta}{\sigma}\right) \text{ for fixed } \eta} \underbrace{\frac{2}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{\eta^2}{2\sigma^2}\right)}_{\text{pdf of absolute value of } \mathcal{N}(0,\sigma^2)} d\eta dx \\ &= \frac{2}{\sqrt{2\pi\sigma^2}} \int_0^{\frac{t}{\sigma}} x \exp\left(-\frac{x^2}{2}\right) \int_0^\infty \exp\left(-\frac{3}{2\sigma^2}\eta^2\right) I_0\left(\frac{\sqrt{2}x}{\sigma}\eta\right) d\eta dx \end{aligned}$$

where I_0 is the modified Bessel function of the first kind. We can solve the inner integral first. For Re $\nu > -1$, Re $\alpha > 0$ the following equality holds CITE!!!:

$$\int_0^\infty \exp\left(-\alpha x^2\right) I_\nu(\beta x) dx = \frac{\sqrt{\pi}}{2\sqrt{\alpha}} \exp\left(\frac{\beta^2}{8\alpha}\right) I_{\frac{1}{2}\nu}\left(\frac{\beta^2}{8\alpha}\right)$$

In our case we have $\nu = 0$, $\alpha = \frac{3}{2\sigma^2}$ and $\beta = \frac{\sqrt{2}x}{\sigma}$, which yields

$$\begin{aligned} \int_0^\infty \exp\left(-\frac{3}{2\sigma^2}\eta^2\right) I_0\left(\frac{\sqrt{2}x}{\sigma}\eta\right) d\eta &= \frac{\sqrt{\pi}}{2\sqrt{\frac{3}{2\sigma^2}}} \exp\left(\frac{\frac{2x^2}{\sigma^2}}{8\frac{3}{2\sigma^2}}\right) I_0\left(\frac{\frac{2x^2}{\sigma^2}}{8\frac{3}{2\sigma^2}}\right) \\ &= \frac{\sqrt{\pi}\sigma}{\sqrt{6}} \exp\left(\frac{x^2}{6}\right) I_0\left(\frac{x^2}{6}\right) \end{aligned}$$

Plugging this in, we get

$$\begin{aligned}
p &= \frac{2}{\sqrt{2\pi\sigma^2}} \int_0^{\frac{t}{\sigma}} x \exp\left(-\frac{x^2}{2}\right) \frac{\sqrt{\pi}\sigma}{\sqrt{6}} \exp\left(\frac{x^2}{6}\right) I_0\left(\frac{x^2}{6}\right) dx \\
&= \frac{1}{\sqrt{3}} \int_0^{\frac{t}{\sigma}} x \exp\left(-\frac{x^2}{2}\right) \exp\left(\frac{x^2}{6}\right) I_0\left(\frac{x^2}{6}\right) dx \\
&= \frac{1}{\sqrt{3}} \int_0^{\frac{t}{\sigma}} x \exp\left(-\frac{x^2}{3}\right) I_0\left(\frac{x^2}{6}\right) dx
\end{aligned}$$

To proceed, we need to integrate by parts to replace the modified Bessel function I_0 with order zero by a modified Bessel function I_1 with order one.

$$\begin{aligned}
p &= \frac{1}{\sqrt{3}} \int_0^{\frac{t}{\sigma}} x \exp\left(-\frac{x^2}{3}\right) I_0\left(\frac{x^2}{6}\right) dx \\
&= \frac{1}{\sqrt{3}} \left[-\frac{3}{2} \exp\left(-\frac{x^2}{3}\right) I_0\left(\frac{x^2}{6}\right)\right]_0^{\frac{t}{\sigma}} + \frac{1}{2\sqrt{3}} \int_0^{\frac{t}{\sigma}} \exp\left(-\frac{x^2}{3}\right) x I_1\left(\frac{x^2}{6}\right) dx \\
&= \frac{1}{\sqrt{3}} \left(\frac{3}{2} - \frac{3}{2} \exp\left(-\frac{t^2}{3\sigma^2}\right) I_0\left(\frac{t^2}{6\sigma^2}\right)\right) + \frac{1}{2\sqrt{3}} \int_0^{\frac{t}{\sigma}} \exp\left(-\frac{x^2}{3}\right) x I_1\left(\frac{x^2}{6}\right) dx
\end{aligned}$$

In the next step we substitute $y = x^2$ in the remaining integral, which leaves us with

$$p = \frac{1}{\sqrt{3}} \left(\frac{3}{2} - \frac{3}{2} \exp\left(-\frac{t^2}{3\sigma^2}\right) I_0\left(\frac{t^2}{6\sigma^2}\right)\right) + \frac{1}{4\sqrt{3}} \int_0^{\frac{t}{\sigma}} \exp\left(-\frac{y}{3}\right) I_1\left(\frac{y}{6}\right) dy$$

We want to solve the remaining integral. Let $p \neq b$ and $s = \sqrt{p^2 - b^2}$, $u = \sqrt{a(p-s)}$ and $v = \sqrt{a(p+s)}$. Then CITE!!!

$$\int_0^a \exp(-px) I_M(bx) dx = \frac{1}{sb^M} \Big((p-s)^M (1 - Q_M(u,v)) - (p+s)^M (1 - Q_M(v,u))\Big)$$

where Q_M denotes the Marcum Q-function. The Marcum Q-function is only defined for $M \geq 1$, which made the integration by parts necessary. Applying this equation with $M = 1$ to the integral yields

$$\int_0^{\frac{t}{\sigma}} \exp\left(-\frac{y}{3}\right) I_1\left(\frac{y}{6}\right) dy = \frac{1}{\frac{1}{2\sqrt{3}}\frac{1}{6}} \frac{2-\sqrt{3}}{6} \left(1 - Q_1\left(\frac{2-\sqrt{3}}{6}\sqrt{\frac{t}{\sigma}}, \frac{2+\sqrt{3}}{6}\sqrt{\frac{t}{\sigma}}\right)\right)$$

$$- \frac{1}{\frac{1}{2\sqrt{3}}\frac{1}{6}} \frac{2+\sqrt{3}}{6} \left(1 - Q_1\left(\frac{2+\sqrt{3}}{6}\sqrt{\frac{t}{\sigma}}, \frac{2-\sqrt{3}}{6}\sqrt{\frac{t}{\sigma}}\right)\right)$$

$$= 2\sqrt{3}(2-\sqrt{3}) \left(1 - Q_1\left(\frac{2-\sqrt{3}}{6}\sqrt{\frac{t}{\sigma}}, \frac{2+\sqrt{3}}{6}\sqrt{\frac{t}{\sigma}}\right)\right)$$

$$- 2\sqrt{3}(2+\sqrt{3}) \left(1 - Q_1\left(\frac{2+\sqrt{3}}{6}\sqrt{\frac{t}{\sigma}}, \frac{2-\sqrt{3}}{6}\sqrt{\frac{t}{\sigma}}\right)\right)$$

Plugging this in, we obtain the final result

$$\begin{aligned}
p &= \frac{1}{\sqrt{3}}\left(\frac{3}{2} - \frac{3}{2}\exp\left(-\frac{t^2}{3\sigma^2}\right) I_0\left(\frac{t^2}{6\sigma^2}\right)\right) + \frac{1}{4\sqrt{3}}\int_0^{\frac{t}{\sigma}} \exp\left(-\frac{y}{3}\right) I_1\left(\frac{y}{6}\right) dy \\
&= \frac{1}{\sqrt{3}}\left(\frac{3}{2} - \frac{3}{2}\exp\left(-\frac{t^2}{3\sigma^2}\right) I_0\left(\frac{t^2}{6\sigma^2}\right)\right) \\
&+ \frac{2\sqrt{3}}{4\sqrt{3}}(2-\sqrt{3}) \left(1 - Q_1\left(\frac{2-\sqrt{3}}{6}\sqrt{\frac{t}{\sigma}}, \frac{2+\sqrt{3}}{6}\sqrt{\frac{t}{\sigma}}\right)\right) \\
&- \frac{2\sqrt{3}}{4\sqrt{3}}(2+\sqrt{3}) \left(1 - Q_1\left(\frac{2+\sqrt{3}}{6}\sqrt{\frac{t}{\sigma}}, \frac{2-\sqrt{3}}{6}\sqrt{\frac{t}{\sigma}}\right)\right) \\
&= \frac{1}{\sqrt{3}}\left(\frac{3}{2} - \frac{3}{2}\exp\left(-\frac{t^2}{3\sigma^2}\right) I_0\left(\frac{t^2}{6\sigma^2}\right)\right) - \sqrt{3} \\
&- \frac{2-\sqrt{3}}{2} Q_1\left(\frac{2-\sqrt{3}}{6}\sqrt{\frac{t}{\sigma}}, \frac{2+\sqrt{3}}{6}\sqrt{\frac{t}{\sigma}}\right) \\
&+ \frac{2+\sqrt{3}}{2} Q_1\left(\frac{2+\sqrt{3}}{6}\sqrt{\frac{t}{\sigma}}, \frac{2-\sqrt{3}}{6}\sqrt{\frac{t}{\sigma}}\right)
\end{aligned}$$

Thus, we have computed the distribution of $\tilde{D}(i,j)$ given $D^+(i,j) = 0$. It is obvious, that there is no easy way to compute an inverse function of this cumulative distribution function, but for a given α we can find numerical solutions to ensure $\mathbb{P}(\tilde{D}^+(i,j) \leq t \mid D^+(i,j) = 0) \geq 1 - \alpha$. By doing so, we get

$$\mathbb{P}(T(i,j) \geq t \mid H_0(i,j)) \leq \mathbb{P}(\tilde{D}^+(i,j) \geq t \mid D^+(i,j) = 0) \leq \alpha \tag{19}$$

and can thus ensure a statistical significance of α.

In equation 14 we have managed to bound the probability of a type I error in our testing procedure by the probability of a type I error when testing for $D^+(i,j) = 0$ using the test statistic $\tilde{D}^+(i,j)$. We want to do the same for the probability of a type II error.

Theorem 1.5. *Assume the following statistical model:*
Let $M, N \in \mathbb{N}$ and $G = \{1, \ldots, M\} \times \{1, \ldots, N\}$. We are given data

$$F(i, j) = c + V(i, j) + \varepsilon_{i,j} \tag{20}$$

where $(i, j) \in G$, $c \in \mathbb{R}$ is constant, $V(i, j) \in \{0, \pm c\}$ and $\varepsilon_{i,j} \sim \mathcal{N}(0, \sigma^2)$ are i.i.d. normal distributed random variables for some $\sigma > 0$ and for all $(i, j) \in G$.
Assume that V contains a rectangular region of interest R and that R has a checkerboard pattern.
Let $\varphi_\alpha \in \{0, 1\}^{M \times N}$ be the binary matrix, that represents the decision of the testing procedure to a given statistical significance α.
Let $k \in \mathbb{N}$ be odd and B be a square structuring element with side length k.
Then the following inequality holds:

$$\mathbb{P}((\varphi_\alpha \circ B)(i, j) = 1 \mid H_0(i, j)) \leq k\alpha^{\frac{k+1}{2}}$$

Proof. We aim to find an upper bound for the probability

$$\mathbb{P}((\varphi_\alpha \circ B)(i, j) = 1 \mid H_0(i, j))$$

To do this, we first notice that $H_0(i, j)$ is equivalent to $V(i, j) = 0$, but since V contains a rectangular region of interest, this means that $i < i_1$ or $i > i_2$ or $j < j_1$ or $j > j_2$. We need to differentiate cases here. These four cases are not mutually exclusive, but have different implications for the row/column of the index (i, j) and a neighbouring row/column:

case	row/column of (i, j)	neighbouring row/column
$i < i_1$	$V(i, 1) = \cdots = V(i, N) = 0$	$V(i-1, 1) = \cdots = V(i-1, N) = 0$
$i > i_2$	$V(i, 1) = \cdots = V(i, N) = 0$	$V(i+1, 1) = \cdots = V(i+1, N) = 0$
$j < j_1$	$V(1, j) = \cdots = V(M, j) = 0$	$V(1, j-1) = \cdots = V(M, j-1) = 0$
$j > j_2$	$V(1, j) = \cdots = V(M, j) = 0$	$V(1, j+1) = \cdots = V(M, j+1) = 0$

Without loss of generality we assume the first case. This means that the null hypotheses $H(i, 1), \ldots, H(i, N)$ and $H(i-1, 1), \ldots, H(i-1, N)$ are true.

To be even more precise, it implies that $D^-(i,1) = \cdots = D^-(i,N) = 0$.
We have assumed the side length k of the structuring element B to be odd. We define the two index sets $K = \{-\frac{k-1}{2}, -\frac{k-3}{2}, \ldots, \frac{k-3}{2}, \frac{k-1}{2}\}$ and $K_1 = \{-\frac{k-1}{2}, -\frac{k-5}{2}, \ldots, \frac{k-5}{2}, \frac{k-1}{2}\}$. This yields

$$\begin{aligned}
&\mathbb{P}((\varphi_\alpha \circ B)(i,j) = 1 \mid H_0(i,j)) \\
&\quad = \mathbb{P}\left(\bigcup_{\tilde{m},\tilde{n}\in K} \bigcap_{m,n\in K} \{\varphi_\alpha(i+m-\tilde{m}, j+n-\tilde{n}) = 1\} \mid D^-(i,1) = \cdots = D^-(i,N) = 0\right) \\
&\quad = \sum_{\tilde{n}\in K} \mathbb{P}\left(\bigcup_{\tilde{m}\in K} \bigcap_{m,n\in K} \{\varphi_\alpha(i+m-\tilde{m}, j+n-\tilde{n}) = 1\} \mid D^-(i,1) = \cdots = D^-(i,N) = 0\right) \\
&\quad \le \sum_{\tilde{n}\in K} \mathbb{P}\left(\bigcup_{\tilde{m}\in K} \bigcap_{n\in K, m=\tilde{m}} \{\varphi_\alpha(i+m-\tilde{m}, j+n-\tilde{n}) = 1\} \mid D^-(i,1) = \cdots = D^-(i,N) = 0\right) \\
&\quad = \sum_{\tilde{n}\in K} \mathbb{P}\left(\bigcup_{\tilde{m}\in K} \bigcap_{n\in K} \{\varphi_\alpha(i, j+n-\tilde{n}) = 1\} \mid D^-(i,1) = \cdots = D^-(i,N) = 0\right) \\
&\quad = \sum_{\tilde{n}\in K} \mathbb{P}\left(\bigcap_{n\in K} \{\varphi_\alpha(i, j+n-\tilde{n}) = 1\} \mid D^-(i,1) = \cdots = D^-(i,N) = 0\right) \\
&\quad \le \sum_{\tilde{n}\in K} \mathbb{P}\left(\bigcap_{n\in K_1} \{\varphi_\alpha(i, j+n-\tilde{n}) = 1\} \mid D^-(i,1) = \cdots = D^-(i,N) = 0\right) \\
&\quad \le \sum_{\tilde{n}\in K} \prod_{n\in K_1} \mathbb{P}\left(\{\varphi_\alpha(i, j+n-\tilde{n}) = 1\} \mid D^-(i,1) = \cdots = D^-(i,N) = 0\right) \\
&\quad = \sum_{\tilde{n}\in K} \prod_{n\in K_1} \mathbb{P}\left(T(i, j+n-\tilde{n}) \ge t_\alpha \mid D^-(i,1) = \cdots = D^-(i,N) = 0\right) \\
&\quad \le \sum_{\tilde{n}\in K} \prod_{n\in K_1} \mathbb{P}\left(\tilde{D}^-(i, j+n-\tilde{n}) \ge t_\alpha \mid D^-(i,1) = \cdots = D^-(i,N) = 0\right) \\
&\quad = \sum_{\tilde{n}\in K} \prod_{n\in K_1} \mathbb{P}\left(\tilde{D}^-(i, j+n-\tilde{n}) \ge t_\alpha \mid D^-(i, j+n-\tilde{n}) = 0\right) \\
&\quad \le \sum_{\tilde{n}\in K} \prod_{n\in K_1} \alpha \\
&\quad = |K|\alpha^{|K_1|} \\
&\quad = k\alpha^{\frac{k+1}{2}}
\end{aligned}$$

The other three cases can be proven in a similar way by swapping the roles of m or $\tilde{m}$ with n or $\tilde{n}$, respectively and/or by replacing D^- by D^+. $\square$