Multicell Multicast Beamforming
with Delayed SNR Feedback

Markus Jordan, Xitao Gong, and Gerd Ascheid
Institute for Integrated Signal Processing Systems
RWTH Aachen University, Aachen, Germany
E-mail: jordan@iss.rwth-aachen.de

Abstract— We evaluate the performance of beamforming in the
downlink of a Single Frequency Network (SFN), in which com-
mon data is transmitted to all participating users simultaneously.
The beamforming is jointly optimized for all cells in the network
and for all users. The algorithm is based on spatio-temporal
correlation knowledge (long term channel state information at
the transmitter) and regularly updated, but possibly delayed,
SNR information. In a first stage of this algorithm an estimator
for the current SNR is built as a function of the beamforming
weights. In a second stage of the algorithm, the beamforming
weights are optimized in order to maximize the minimum SNR
in the system.

I. INTRODUCTION

The transmission of common information to a multitude
of users is called multicast herein and is opposed to the
transmission to a single user, which is called unicast. Typical
applications of such a multicast scenario include the streaming
of video or audio content to users. Multicast has been consid-
ered under the name MBMS (Multimedia Broadcast/Multicast
Service) [1] as a feature of UMTS and under the name MBS
(Multicast and Broadcast Service) as a feature of WiMAX [2]
to provide the network operators with an additional service.

In order to cover a larger area than a single base station
could, multicast is often transmitted from multiple synchro-
nized base stations. If a suitable modulation as e.g. orthogonal
frequency division multiplex (OFDM) with a sufficient cyclic
prefix length is being used, the contributions from these mul-
tiple base stations add up at the receiver and the symbols can
be detected without intersymbol interference. Such a network
is called a single frequency network.

In previous single frequency networks as e.g. formed for
DVB-T or DAB, no knowledge of the users’ channels exists
in the network as there is no uplink. As a consequence of
this, omnidirectional transmission is the only sensible strategy.
In mobile communication systems such as LTE or WiMAX
however an uplink exists, which opens up the possibility to
obtain knowledge of the downlink channel and to use
this knowledge to ameliorate the multicast transmission via
beamforming.

Previous works have focussed either on the case that either
full knowledge of the instantaneous channel realizations exists
in the transmitter [3] [4] or on the case that the transmitter has
only long-term, i.e. correlation, knowledge of the channel [5].

In the aforementioned communication systems however,
channel quality (i.e. SNR) knowledge is typically available,
since it is required for adaptive modulation and coding and
channel state dependent scheduling. Furthermore, channel
correlation knowledge can be acquired via dedicated low-
rate uplink signalling or via the estimation of the uplink
channel correlation followed by postprocessing to compensate
for different carrier frequencies between up- and downlink in
FDD systems [6] [7].

If both channel correlation knowledge and SNR information
exist, both should be used for beamforming. In this paper, the
performance of a single frequency network is enhanced via
beamforming that takes channel correlation and SNR infor-
mation into account. The objective is to maximize, in every
time slot, the worst link’s SNR based on SNR feedback. The
proposed algorithm consists of two steps: In a first step a linear
minimum mean square error (LMMSE) estimator is built,
using the information about the past beamforming weights, the
SNR that was fed back to the base station that resulted from
the past beamforming weights and the correlation knowledge.
A reformulation of the estimator yields a quadratic form in the
beamformers to be applied in the current time slot. In a second
step, this quadratic form is the objective function for an NP-
hard optimization problem, which allows for an approximation
via convex relaxation. This relaxation can then be solved
efficiently. It is of special importance to consider the effect of
delayed SNR knowledge because of the amount of signalling
over multiple layers of the communication stack involved with
a centralized optimization. The proposed SNR estimator is
effectively a predictor, which counteracts the delay.

After this introduction, we will describe the system setup
and problem statement in Section II and explain the proposed
algorithm in Section III. Results will be shown in Section IV,
before Section V concludes this paper.

II. SYSTEM SETUP AND PROBLEM STATEMENT

We consider a patch of \( N_c \) hexagonal cells (cf. Fig. 1),
which are being formed by base stations with sectorized
antennas, consisting of \( N_T \) elements per cell.

A total of \( N_U \) single-antenna users are being provided with
data by these base stations, and every user receives the same
data (i.e. this is the multicast case). We will only consider
a narrowband scenario, but an extension to a frequency-
selective case is straightforward if OFDM is being used. The
narrowband assumption implies that the symbol duration is
significantly longer than the delay spread of the multicell
channel (consisting of delayed contributions from every base station according to the propagation delay to the user and the time synchronization error between the base stations) and can be disregarded. Furthermore, we assume a perfect frequency synchronization of the base stations.

Let $d_k(l)$ denote the $l$-th transmitted symbol during time slot $k$, then the received signal for a given user can be described by

$$ r_k(l) = \sum_{c=1}^{N_c} {h}_{c,k}^H w_{c,k} d_k(l) + n_k(l) = s_k(l) + n_k(l), $$

where $w_{c,k}$ is the beamforming vector and $h_{c,k}$ is the vector of the MISO-channel of cell $c$ at time slot $k$. Averaging over the transmitted symbols $d$ and assuming unit symbol power, the SNR at time slot $k$ is

$$ \gamma_k = \frac{1}{\sigma_n^2} E_A[|s_k(l)|^2] = \frac{1}{\sigma_n^2} \sum_{c=1}^{N_c} |w_{c,k}^H h_{c,k}|^2, $$

where $\sigma_n^2$ is the noise power.

We assume that the channel is stationary and complex Gaussian distributed and that spatio-temporal channel correlation knowledge $R_{c,n}$ exists in the network for all users as seen from all base stations. This knowledge is long-term knowledge and needs to be reported to the base station only infrequently. Furthermore, if such a feedback does not exist or is assumed to be too bandwidth-consuming, downlink spatio-temporal correlation knowledge can be obtained from uplink measurements [8] [9].

The channels $h_{c,k}$ and $h_{d,k}$ are assumed to be uncorrelated for $c \neq d$ and zero-mean. Then, the mean SNR (averaged across the Rayleigh fading) can be calculated as

$$ \gamma_k = \mathbb{E}_b \left\{ \frac{1}{\sigma_n^2} \sum_{c=1}^{N_c} |w_{c,k}^H h_{c,k}|^2 \right\} = \frac{1}{\sigma_n^2} \sum_{c=1}^{N_c} |w_{c,k}^H R_{c,0} w_{c,k}|, $$

where $R_{c,0}$ is the matrix of auto-correlation.

![Fig. 1. Layout of the cells. Base stations in red, users in blue, lines indicate the base station orientation and the user heading. Number of cells $N_c = 4$, Number of users $N_U = 6$.](image.png)

III. MULTICELL MULTICAST BEAMFORMING WITH DELAYED SNR FEEDBACK

Exploiting spatio-temporal correlation knowledge and feedback SNR from the users, we want to maximize the minimum SNR observed in the system in every time slot, and to transmit common data to all users with a rate given by this minimum SNR.

The proposed algorithm consists of two steps: First, the current SNR is estimated as a function of the beamforming weights to be applied in the current time slot. Second, the beamforming weights are optimized in order to maximize the minimum estimated SNR.

In the first step of the algorithm we use LMMSE estimators $\hat{f}_k = \mathbb{D}_{o,k} d_{o,k}$ with $N_F$ coefficients to estimate the current SNR at time slot $k$ given feedback SNR information. This SNR information is assumed to be delayed by $N_D$ time slots as a consequence of the processing required to estimate the SNR at the receiver, the time for the next feedback channel access, the communication of the SNR from the base station to the centralized optimization entity over the backbone of the network and the transmission of the determined beamforming weights from the optimizer to the respective cells. The total estimated SNR can be written as

$$ \hat{\gamma}_k = f_k^T \gamma_{o,k} = d_{o,k}^T \mathbb{D}_{o,k}^{-1} \gamma_{o,k}, $$

where the feedback SNR values are contained in

$$ \gamma_{o,k} = (\gamma_{k-N_D}, \ldots, \gamma_{k-N_D-N_F+1})^T. $$

The vector of cross-correlation of the observed SNR measurements $\gamma_{o,k}$ and the current SNR $\hat{\gamma}_k$ is given as

$$ d_{i,j,k} = E\{\gamma_{i,k} \gamma_{j,k}\} = (d_0, N_D, \ldots, d_0, N_D + N_F - 1, k)^T $$

with

$$ d_{i,j,k} = E\{\gamma_{i,k} \hat{\gamma}_{j,k}\}. $$

The coefficients $d_{i,j,k}$ are shown in the appendix to yield

$$ d_{i,j,k} = \gamma_{i,k} \hat{\gamma}_{j,k} + \frac{1}{\sigma_n^2} \sum_{c=1}^{N_c} |w_{c,k}^H R_{c,i,j} w_{c,k}|^2. $$

The matrix of auto-correlation $\mathbb{D}_{o,k} = E\{\gamma_{o,k} \gamma_{o,k}^T\}$ contains the coefficients

$$ [\mathbb{D}_{o,k}]_{m,n} = d_{m+N_D-n+N_D-1,k}, \quad 1 \leq m, n \leq N_F. $$

Introducing new filters $p_k = \mathbb{D}_{o,k}^{-1} \gamma_{o,k}$, the estimated SNR can be written as

$$ \hat{\gamma}_k = d_{o,k}^T \mathbb{D}_{o,k}^{-1} \gamma_{o,k} = d_{o,k}^T p_k \gamma_{o,k} = \sum_{f=1}^{N_F} d_0 d_{N_D+f-1,k} p_{f,k}. $$

Please note that the filters $p_k$ depend only on quantities from previous time slots and not on the beamforming weights to be applied in time slot $k$.

Introducing stacked vectors and matrices $w_k = (w_{1,k}^T, \ldots, w_{N_F,k}^T)^T$ and $\mathbb{R}_c = \text{diag}(R_{1,c}, \ldots, R_{N_U,c})$, where $\text{diag}(A_1, \ldots, A_N)$ denotes a block-diagonal matrix containing
the matrices $A_i$ along its main diagonal, the coefficients $d_{i,j,k}$ may be written as

$$d_{i,j,k} = \gamma_k - \bar{\gamma}_{k-j} + \frac{1}{\sigma^2} w^H_{i,j} R_{i,j} w_{k-j}$$

(12)

and the mean SNR as $\bar{\gamma}_k = \frac{1}{\sigma^2} w^H_{i,j} R_{i,j} w_{k}$. With this new notation, the estimated SNR can be expressed as

$$\hat{\gamma}_k = \sum_{f=1}^{N_f} P_{f,k} \left( \frac{1}{\sigma^2} w^H_{i,j} R_{i,j} w_{k-N_f-f+1} + \frac{1}{\sigma^2} w^H_{k,N_f-f+1} R_{N_f+f-1} \right) \theta_{k}.$$  

(13)

This can be abbreviated as

$$\hat{\gamma}_k = w^H_{i,j} A_k w_{k}$$

(14)

with

$$A_k = \sum_{f=1}^{N_f} P_{f,k} \left( \frac{1}{\sigma^2} R_{i,j} w_{k-N_f-f+1} + \frac{1}{\sigma^2} R_{N_f+f-1}. \right.$$  

$$w_{k-N_f-f+1} w^H_{k-N_f-f+1} R_{N_f+f-1} \right).$$

Here, matrix $A_k$ contains information about the channel correlation and the observed SNR and the past beamforming weights but does not depend on the beamforming weights $w_1, \ldots, w_{N_k}$ to be applied in time slot $k$. In this sense, matrix $A_k$ can be understood as an estimator for the SNR of a given user if a set of beamforming weights $(w_{i,k}^T, \ldots, w_{N_k,k})^T$ is being applied.

Since $f_k$ is the LMMSE estimator, the resulting mean squared estimation error $e_k^2 = E_k(\{\gamma_k - \hat{\gamma}_k\}^2)$ can be expressed as [10]

$$e_k^2 = E(\hat{\gamma}_k^2) - E(\hat{\gamma}^2_k) = d_{0,0,k} - E(\hat{\gamma}^2_{i,k} \gamma_{i,k}^T f_k)$$

$$= 2 \frac{1}{\sigma^2} \sum_{i,c=1}^{N} w^H_{i,c} R_{i,c} w_{i,k}^2 - f_k^T D_{0,0,k} f_k.$$  

(15)

In the second step of our proposed algorithm, we use the expression for the estimated SNR to allow for a maximization of the SNR over the applied beamforming weights for time slot $k$. Using the notation introduced in (13), the problem of maximizing the minimum estimated SNR $\bar{\gamma}_{\text{min},k}$ in the system in time slot $k$, using both long-term correlation knowledge and short-term but outdated SNR information, under a total power constraint $P$ may be stated as

$$\bar{\gamma}_{\text{min},k} = \max \min_{i=1}^{N_k} w^H_i A_i w_k$$

subject to $w^H_i w_k = P$,

(16)

where superscript $i$ indicates the user and * indicates the optimum value. Optimization problems of this kind have been the subject of several publications (e.g. [4] [11]), which showed that it is NP-hard in general, but proposed solution strategies based on either sequential quadratic programming (SQP) or on convex relaxation, which yields an approximation of the problem that is convex and therefore easily solvable. In short, after having dropped a rank-1 constraint, the above problem can be expressed via another problem

$$\min \ Tr\{CX\}$$

s.t. $Tr\{E_i X\} = b_i \ \forall i = 1, \ldots, N_U + 1,$

$$X \geq 0,$$

with suitably defined $X, E_i, C$ and $b_i$. This problem is a semidefinite program (SDP) in standard form [12] which can then be solved numerically using software such as SeDuMi [13]. Because having dropped the rank-1 constraint, a solution to this problem has to be post-processed via a so-called randomization procedure to yield a (usually suboptimum) solution to (16). The derivation of (17) is included in the appendix for convenience. For more details, please refer to e.g. [4] or [5].

IV. RESULTS

Table I shows the simulation parameters that we used throughout the evaluation of the proposed algorithm unless otherwise noted. The utilized link-level system-level interface is shown in Figure 2 and is based on the WiMAX technical specifications [2].

Figure 3 shows exemplary a resulting SNR plot over time for the user drop shown in Figure 1. The actually measured SNR at the user side $\gamma_k^i$ for user $i$ according to (2) is plotted.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intersite distance</td>
<td>1732 m</td>
</tr>
<tr>
<td>Sectorized</td>
<td>Yes</td>
</tr>
<tr>
<td>Sectorization pattern</td>
<td>-10 dB at 60°</td>
</tr>
<tr>
<td>Path loss exponent</td>
<td>3.76</td>
</tr>
<tr>
<td>Number of beamforming cells</td>
<td>4</td>
</tr>
<tr>
<td>User placement</td>
<td>uniform (cartesian)</td>
</tr>
<tr>
<td>User velocity</td>
<td>4 km/h</td>
</tr>
<tr>
<td>User heading</td>
<td>$I(0, 2\pi)$</td>
</tr>
<tr>
<td>Transmit antennas</td>
<td>4</td>
</tr>
<tr>
<td>Antenna spacing</td>
<td>half-wavelength</td>
</tr>
<tr>
<td>Channel Model</td>
<td>Rayleigh fading, stationary</td>
</tr>
<tr>
<td>Power angular density</td>
<td>$\pi \cdot$ std. dev.</td>
</tr>
<tr>
<td>Bandwidth of Multicast Chan.</td>
<td>5 MHz</td>
</tr>
<tr>
<td>Time slot duration</td>
<td>2 ms</td>
</tr>
<tr>
<td>SNR feedback delay</td>
<td>2 ms</td>
</tr>
</tbody>
</table>

TABLE I

DEFAULT SIMULATION PARAMETERS
coloured, while the estimated minimum SNR $\hat{\gamma}_{\text{min,k}}$ according to (16) is shown as a black dash-dotted curve. A mean minimum SNR of about 12 dB is achieved, with shallow deep fades and peaks. This allows for a transmission with a relatively constant rate. The SNR of user 1 (the blue line) experiences abrupt and strong changes. This is because the optimization problem only considers the weakest link, and stronger links do not contribute to the objective function. The changes in the SNR for the stronger users are not causing problems in the reception of the data since the SNR requirement for reliable decoding is already oversatisfied. If the fast changes of the beamformers cause problems for the interference handling of surrounding unicasting cells, additional constraints could be added to the optimization problem to counteract this behavior. Furthermore, the SNR is subject to small changes within one or two dB, which are caused by imperfect SNR estimation.

We evaluated the SNR estimation error as a function of the number of filter coefficients $N_F$ as shown in Figure 4. As the weakest link is the most important link in this problem, only the user with the lowest mean SNR (averaged across Rayleigh fading) under omnidirectional transmission of the base stations was considered here. Three different user numbers ($N_U = 4, 6, 8$) were assumed and a SNR feedback delay of $N_D = 1$ and $N_D = 3$ were applied. The solid and dash-dotted lines show the theoretical error obtained via (15) with $N_D = 1$ and $N_D = 3$, respectively, while the markers show simulated results obtained via time averaging across many time slots. A larger SNR feedback delay can be seen to lead to larger estimation errors, as well as a reduction in the error is visible for increasing numbers of filter coefficients. The largest relative reduction of the estimation error can be observed for $N_D = 3$ and $N_U = 4$, where the error was reduced from $\bar{\epsilon} = 8$ to $\bar{\epsilon} = 4$ if the number of filter coefficients was increased from $N_F = 4$ to $N_F = 12$.

The resulting system-level throughput is depicted in Figure 5. The throughput per user increases monotonously with both an increasing number of filter coefficients and a decreasing number of users in the system. Increasing the number of filter coefficients from $N_F = 2$ to $N_F = 12$ leads to an increase in throughput of about 20%. A high absolute throughput is achieved for a low number of users in the system, since the total transmit power is then divided among fewer users.

A higher number of filter coefficients implies a higher numerical complexity, as e.g. the matrix inversion required in (5) becomes more complex. However, the numerical complexity of the algorithm is found to be dominated, in the range of the number of filter coefficients considered herein, by the
solution to (17), which does not depend on the number of filter coefficients. However, a higher number of filter coefficients implies that more correlation knowledge is required in the base station, as the evaluation of $D_{n o, k}$ and $d_{o, k}$ requires the knowledge of $R_{c, n}$ for all $c$ and $-N_D - N_F + 1 \leq n \leq N_F - 1$. Feeding this information back to the base station increases the load of the uplink correspondingly with an increasing number of filter coefficients.

This results in

$$d_{i, j, k} = \frac{1}{\sigma_n^2} E\{\sum_{c=1}^{N_c} w_h^c h_{c, k-i} h_{c, k-j}^* + \sum_{c=1}^{N_c} w_h^c h_{c, k-i} h_{c, k-j}^*\}$$

$$= \frac{1}{\sigma_n^2} E\{\sum_{a=1}^{N_c} \sum_{b=1}^{N_c} \sum_{c=1}^{N_c} \sum_{d=1}^{N_c} w_h^c h_{c, k-i} h_{c, k-j}^* w_h^d h_{d, k-j}^*\}$$

and with the definition of the effective channel $h_{c, k}^e = \frac{1}{\sigma_n} w_h^c h_{c, k}$ it yields

$$d_{i, j, k} = E\{\sum_{a=1}^{N_c} \sum_{b=1}^{N_c} \sum_{c=1}^{N_c} \sum_{d=1}^{N_c} h_{a, k-i}^e h_{a, k-j}^e h_{c, k-i}^e h_{c, k-j}^e + \sum_{a=1}^{N_c} \sum_{b=1}^{N_c} \sum_{c=1}^{N_c} \sum_{d=1}^{N_c} h_{b, k-i}^e h_{b, k-j}^e h_{c, k-i}^e h_{c, k-j}^e\}$$

This results in

$$d_{i, j, k} = \sum_{a=1}^{N_c} \sum_{b=1}^{N_c} \sum_{c=1}^{N_c} \sum_{d=1}^{N_c} E\{h_{a, k-i}^e h_{a, k-j}^e h_{c, k-i}^e h_{c, k-j}^e\} + \sum_{a=1}^{N_c} \sum_{b=1}^{N_c} \sum_{c=1}^{N_c} \sum_{d=1}^{N_c} E\{h_{b, k-i}^e h_{b, k-j}^e h_{c, k-i}^e h_{c, k-j}^e\}$$

V. CONCLUSION

Beamforming in the context of multicast transmission was considered in a multicell environment. We have shown a way to incorporate SNR information, which is typically available in mobile communication standards, with channel correlation information, which can be obtained at low cost, to maximize the SNR of the weakest link on a time-slot basis. In a first step, a LMMSE estimator is determined for the given state, comprising past beamforming vectors, fed back SNR information and spatio-temporal channel correlation knowledge. This allows for rewriting the estimated SNR to obtain a formulation of the problem of maximizing the minimum SNR solvable by convex relaxation. The algorithm does not require complex channel vector feedback but only real valued SNR information.

VI. APPENDIX

A. Derivation of Eq. (9)

The definition

$$d_{i, j, k} = E\{\gamma_{k-i} \gamma_{k-j}\}$$

leads to

$$d_{i, j, k} = \frac{1}{\sigma_n^2} E\{|\sum_{c=1}^{N_c} w_h^c h_{c, k-i} h_{c, k-j}^*| \sum_{c=1}^{N_c} w_h^c h_{c, k-i} h_{c, k-j}^*\}$$

$$= \frac{1}{\sigma_n^2} E\{|\sum_{a=1}^{N_c} w_h^a h_{a, k-i} h_{a, k-j}^*| \sum_{a=1}^{N_c} w_h^a h_{a, k-i} h_{a, k-j}^*\}$$

$$= \frac{1}{\sigma_n^2} E\{|\sum_{a=1}^{N_c} w_h^a h_{a, k-i} h_{a, k-j}^*| \sum_{a=1}^{N_c} w_h^a h_{a, k-i} h_{a, k-j}^*\}$$

where $R_{a, i-j} = R_{a, i-j}$ was used. Eq. (9) follows directly.
B. Derivation of Eq. (17)

The optimization problem of interest is

$$
\begin{align*}
\max_{w_k} & \quad \min_{s=1..N_U} w_k^H A_k^T A_k w_k \\
\text{s.t.} & \quad w_k^H w_k = P.
\end{align*}
$$

(18)

Introducing a new optimization variable $\zeta$ allows a rewriting as

$$
\begin{align*}
\max_{w_k, \zeta} & \quad \zeta \\
\text{s.t.} & \quad \zeta \leq w_k^H A_k^T A_k w_k \quad \forall i = 1, \ldots, N_U, \\
& \quad w_k^H w_k = P, \\
& \quad \zeta \geq 0.
\end{align*}
$$

(19)

Introducing slack variables $s_i$ to turn the inequality constraints into equality and nonnegativity constraints yields

$$
\begin{align*}
\max_{w_k, \zeta, s_1, \ldots, s_{N_U}} & \quad \zeta + s_i = w_k^H A_k^T A_k w_k \quad \forall i = 1, \ldots, N_U, \\
& \quad w_k^H w_k = P, \\
& \quad s_i \geq 0 \quad \forall i = 1, \ldots, N_U, \\
& \quad \zeta \geq 0.
\end{align*}
$$

(20)

Using the identity $w_k^H A w = \text{Tr}(A w w^H)$ and defining a new matrix $W_k = w_k^H w_k$ leads to

$$
\begin{align*}
\max_{w_k, \zeta, s_1, \ldots, s_{N_U}} & \quad \zeta + s_i = \text{Tr}(A_k^T W_k) \quad \forall i = 1, \ldots, N_U, \\
& \quad \text{Tr}(W_k) = P, \\
& \quad s_i \geq 0 \quad \forall i = 1, \ldots, N_U, \\
& \quad \zeta \geq 0, \\
& \quad W_k \succeq 0, \\
& \quad \text{rank}(W_k) = 1.
\end{align*}
$$

(21)

All optimization variables can now be written into one block-diagonal matrix $X = \text{diag}(\zeta, s_1, \ldots, s_{N_U}, W_k)$. Note that this matrix is positive semidefinite if $W_k$ is positive semidefinite and all $\zeta$ and $s_i$ are nonnegative.

Dropping the rank-1 constraint leads to

$$
\begin{align*}
\min & \quad \text{Tr}(CX) \\
\text{s.t.} & \quad \text{Tr}(E_i X) = b_i \quad \forall i = 1, \ldots, N_U + 1, \\
& \quad X \succeq 0,
\end{align*}
$$

(22)

with $C$ being a matrix of all zeros except for $c_{1,1} = -1$, $b_1 = \ldots = b_{N_U} = 0$, $b_{N_U+1} = P$. For $1 \leq i \leq N_U$ is $E_i$ a block-diagonal matrix with entries $c_{i,1} = c_{i+1,i+1} = -1$ and the lower right block containing $A_k^T$ and zeros otherwise. Matrix $E_{N_U+1}$ is contains all zeros except for the last $N_T$ elements of the main diagonal, which are ones.

REFERENCES


