

POLSKA AKADEMIA NAUK
KOMITET ELEKTRONIKI I TELEKOMUNIKACJI

ych,
KWARTALNIK
ELEKTRONIKI I TELEKOMUNIKACJI

ELECTRONICS AND
TELECOMMUNICATIONS
QUARTERLY

TOM 50 — ZESZYT 3

WARSZAWA 2004

RADA REDAKCYJNA

Przewodniczący

Prof. dr hab. inż. STEFAN HAHN
czł. rzecz. PAN

Członkowie

prof. dr hab. inż. DANIEL JÓZEF BEM — czł. koresp. PAN, prof. dr hab. inż. MICHAŁ BIAŁKO — czł. rzecz. PAN, prof. dr hab. inż. MAREK DOMAŃSKI, prof. dr hab. inż. ANDRZEJ HAŁAŚ, prof. dr hab. inż. JÓZEF MODELSKI, prof. dr inż. JERZY OSIOWSKI, prof. dr hab. inż. EDWARD SĘDEK, prof. dr hab. inż. MICHAŁ TADEUSIEWICZ, prof. dr hab. inż. WIESŁAW WOLIŃSKI — czł. koresp. PAN, prof. dr inż. MARIAN ZIENTALSKI

REDAKCJA

Redaktor Naczelny

prof. dr hab. inż. WIESŁAW WOLIŃSKI

Zastępca Redaktora Naczelnego

doc. dr inż. KRYSTYN PLEWKO

Sekretarz Odpowiedzialny

mgr ELŻBIETA SZCZEPANIAK

ADRES REDAKCJI

00-665 Warszawa, ul. Nowowiejska 15/19 Politechnika, pok. 470
Instytut Telekomunikacji, Gmach im. prof. JANUSZA GROSZKOWSKIEGO

Dyżury Redakcji: poniedziałki i piątki, godz. 14–16
tel/fax (022) 660 77 37

Telefony domowe: Redaktora Naczelnego: 812 17 65
Zast. Red. Naczelnego: 826 83 41
Sekretarza Odpowiedzialnego: 633 92 52

Ark. wyd. 13,75	Ark. druk. 11,0	Podpisano do druku wrzesień 2004 r.
Papier offset. kl. III 80 g. B-1		Druk ukończono wrzesień 2004 r.

Skład, druk i oprawa: Warszawska Drukarnia Naukowa PAN
00-656 Warszawa, ul. Śniadeckich 8
Tel./fax: 628-87-77

„Kwart
Quartern
Elektrot
Kwa
demii N
jest czas
prezentu
gładowy
ktroniki,
medyczn
Auto
czeniu, a
Arty
mi bada
danej g
matemat
Commis
Wsz
krajowy
naukowy
w ramac
stawiany
Czas
krajowy
nych. Je
Kažo
ułatwia p
lub za g
w języku
Nade
sprawnj
są zamie
Arty
stronie r

Szanowni Autorzy

„Kwartalnik Elektroniki i Telekomunikacji” — Electronics and Telecommunications Quarterly jest kontynuatorem tradycji powstałego 50 lat temu kwartalnika p.t. „Rozprawy Elektrotechniczne”.

Kwartalnik jest czasopismem Komitetu Elektroniki i Telekomunikacji Polskiej Akademii Nauk. Wydawany jest przez Warszawską Drukarnię Naukową PAN. Kwartalnik jest czasopismem naukowym, na którego łamach są publikowane artykuły i komunikaty prezentujące wyniki oryginalnych prac teoretycznych i doświadczalnych, a także przeglądowych. Związane są one z szeroko rozumianymi dziedzinami współczesnej elektroniki, telekomunikacji, mikroelektroniki, oproelektroniki, radiotechniki i elektroniki medycznej.

Autorami publikacji są wybitni naukowcy, znani specjaliści o wieloletnim doświadczeniu, a także młodzi badacze — głównie doktoranci.

Artykuły charakteryzują się oryginalnym ujęciem zagadnienia, interesującymi wynikami badań, krytyczną oceną teorii lub metod, omówieniem aktualnego stanu, lub postępu danej gałęzi techniki oraz omówieniem perspektyw rozwojowych. Sposób pisania matematycznej części artykułów zgodny jest z wytycznymi IEC (International Electronics Commision) oraz ISO (International Organization of Standarization).

Wszystkie publikowane w Kwartalniku artykuły są recenzowane przez znanych krajowych specjalistów, co zapewnia że publikacje te są uznawane jako autorski dorobek naukowy. Opublikowane w kwartalniku wyniki prac naukowych zrealizowanych w ramach „GRANTów” Komitetu Badań Naukowych spełnia więc jeden z wymogów stawianych tym pracom.

Czasopismo dociera do wszystkich zajmujących się elektroniką i telekomunikacją krajowych ośrodków naukowych oraz technicznych, a także szeregu instytucji zagranicznych. Jest ponadto prenumerowane przez liczne grono specjalistów i biblioteki.

Każdy Autor otrzymuje bezpłatnie 20 egzemplarzy nadbitek swojego artykułu, co ułatwia przesłanie go do indywidualnych wybranych przez Autora osób i instytucji w kraju lub za granicą. Ułatwia to dodatkowo fakt, że w Kwartalniku są publikowane artykuły w języku angielskim.

Nadesłane do redakcji artykuły są publikowane w terminie około pół roku, w przypadku sprawnej współpracy Autora z Redakcją. Wytyczne dla Autorów dotycząc formy publikacji są zamieszczone w zeszytach Kwartalnika, można je także otrzymać w siedzibie Redakcji.

Artykuły można dostarczać osobiście, lub pocztą pod adresem zamieszczanym na stronie redakcyjnej w każdym zeszycie.

Redakcja

KO — czł.
r hab. inż.
r hab. inż.
of. dr inż.

E. Hryn
R. Gola
pre
K. Góre
ZY
A. Boro
—
A. Boro
—
J. Stanc
op
W. Adan
dw
P. Grybo
prz
D. Mako
fot
J. Świde
Informac

E. Hryn
R. Golan
K. Góre
for
A. Boro
—
A. Boro
—
J. Stancel
of c
W. Adan
of I
P. Grybo
D. Mako
mo
J. Świde
Informat

SPIS TREŚCI

E. Hryniewicz, D. Kania: Metody syntezy dedykowane dla struktur FPGA typu tablicowego	325
R. Golański: Badania symulacyjne modulacji różnicowych ze zmienną częstotliwością próbkowania	343
K. Górecki, J. Zarębski: Elektrotermiczny makromodel sterownika przetwornic rezonansowych ZVS dla programu SPICE	365
A. Borowicz, A. A. Petrovsky: Metody motywowane perceptualnie w uzdatnianiu sygnału mowy — część I: eliminacja zjawiska tonów muzycznych	379
A. Borowicz, A.A. Petrovsky: metody motywowane perceptualnie w uzdatnianiu sygnału mowy — część II: psychoakustyczna optymalizacja metod wag widmowych	395
J. Stanlik: Modelowanie zależności maksymalnego napięcia wyjściowego wzmacniaczy operacyjnych od prądu wyjściowego	411
W. Adamski: Czterostanowa metoda bezpośredniego pomiaru współczynnika szumu dwuwrotników o małym wzmacnieniu	429
P. Gryboś: Optymalizacja szumowa wielokanałowych mieszanych układów scalonych na przykładzie układu RX64	441
D. Makowski, W. Marańda, R. Pijarski: Niedrogi miernik charakterystyk modułów fotowoltaicznych	471
J. Świdorski, A. Zając, M. Skórczakowski, P. Konieczny: Iterbowy laser włóknowy dużej mocy	481
Informacje dla Autorów (w jęz. polskim)	489

CONTENTS

E. Hryniewicz, D. Kania: A method of synthesis for LUT-based FPGA's	325
R. Golański: Computer simulation of the delta modulations with sampling frequency adaptation	343
K. Górecki, J. Zarębski: Electrothermal macromodel of the controller of ZVS resonant converter for spice.....	365
A. Borowicz, A. A. Petrovsky: Perceptually motivated approaches to speech enhancement — Part I: Elimination of the musical noise phenomenon	379
A. Borowicz, A.A. Petrovsky: Erceptually motivated approaches to speech enhancement — Part II: Psychoacoustic optimization of spectral weighting rules.....	395
J. Stanlik: Modelling the relation between maximum output voltage and output current of operational amplifiers.....	411
W. Adamski: Four-state measurement method for direct-reading noise figure measurement of low-gain two-ports.....	429
P. Gryboś: Design for a low noise of multichannel integrated circuits on example of RX64 chip	441
D. Makowski, W. Marańda, R. Pijarski: Inexpensive characteristics scanner of photovoltaic modules	471
J. Świdorski, A. Zając, M. Skórczakowski, P. Konieczny: High-power Yb-doped silica fiber laser ...	481
Information for the Authors (w jęz. ang.)	492

N

tab
na,
w
cję
Wy
po
po
syt

Sła

Pod
typu tab
konfigur
Stru
da się z
ścia/wyj
kanały z
do tego
rozmiesz
połączeni

Metody syntezy dedykowane dla struktur FPGA typu tablicowego

EDWARD HRYNKIEWICZ, DARIUSZ KANIA

*Instytut Elektroniki, Politechnika Śląska
ul. Akademicka 16, 44-100 Gliwice
email: eh@boss.iele.polsl.gliwice.pl; kdariusz@zeus.polsl.gliwice.pl*

*Otrzymano 2003.03.25
Autoryzowano 2003.12.05*

W artykule przedstawiono metody syntezy przeznaczone dla układów FPGA typu tablicowego. Głównym elementem zaproponowanych metod jest dekompozycja funkcjonalna, oparta na teorii Curtis'a. Różnica pomiędzy zaproponowanymi metodami syntezy tkwi w kierunku prowadzenia dekompozycji. W pierwszej metodzie stosuje się tzw. dekompozycję „od wejść do wyjść”, w odróżnieniu od drugiej, w której stosuje się „od wyjść do wejść”. Wyniki eksperymentów jednoznacznie pokazały, że lepsze rezultaty uzyskuje się wykonując podział projektowanego układu metodą „od wejść do wyjść”. W artykule zawarto również porównanie przedstawionych metod dekompozycji z innymi, znanymi z literatury metodami syntezy, bazującymi na dekompozycji.

Słowa kluczowe: synteza logiczna, dekompozycja, podział, FPGA

1. WPROWADZENIE

Podstawowym problemem syntezy układów cyfrowych na bazie struktur FPGA typu tablicowego jest sposób podziału projektu na części, realizowane w odrębnych konfigurowalnych blokach logicznych [2, 20].

Struktura wewnętrzna układów FPGA typu tablicowego (ang. Look-up Table) składa się z matrycy konfigurowalnych komórek otoczonych na obrzeżach blokami wejścia/wyjścia. Pomiędzy rzędami i kolumnami konfigurowalnych komórek znajdują się kanały z liniami połączeń. Możliwość prowadzenia połączeń jedynie w wydzielonych do tego celu kanałach sprawia, że równie ważnym etapem syntezy staje się sposób rozmieszczenia bloków wewnątrz struktury (ang. placement) oraz sposób prowadzenia połączeń (ang. routing). Ponieważ zasoby logiczne konfigurowalnych bloków logicz-

nych umożliwiają realizację dowolnej funkcji o określonej liczbie argumentów stąd strategicznym etapem całej syntezy jest dekompozycja, stanowiąca podstawę podziału całego projektu na poszczególne części o zadanej liczbie wejść.

Pojawienie się na rynku układów FPGA typu tablicowego doprowadziło do ogromnego wzrostu zainteresowania elementami dekompozycji, pierwotnie (lata 60-te) wykorzystywanymi w procesie minimalizacji, stanowiąc uzupełnienie metody Quine'a-McCluskey'a. Podstawowe elementy klasycznego modelu dekompozycji funkcjonalnej wprowadzone przez Ashenhurst'a w drugiej połowie lat pięćdziesiątych [3] i rozszerzone w latach sześćdziesiątych przez Curtis'a [8, 9] stanowią do dnia dzisiejszego podwaliny większości opracowanych algorytmów dekompozycji. Klasyczny model dekompozycji doskonale wpasowuje się w potrzeby syntezy logicznej dedykowanej dla struktur FPGA typu tablicowego.

Celem niniejszego artykułu jest przedstawienie dwóch metod syntezy różniących się przede wszystkim kierunkiem prowadzenia dekompozycji („od wejść do wyjść”; „od wyjść do wejść”), ale również sposobem korzystania z dekompozycji złożonych.

2. PODSTAWY TEORETYCZNE

Niech f będzie n -wejściową i m -wyjściową funkcją logiczną odwzorowującą zbiór B^n w zbiór B^m tzn. $f : B^n \rightarrow B^m$, gdzie $B = \{0, 1\}$. Niech $Y = f(I_{n-1}, \dots, I_1, I_0)$, gdzie $Y = \{y_{m-1}, \dots, y_1, y_0\}$. Podstawę rozłącznej dekompozycji funkcjonalnej stanowi poniższe twierdzenie.

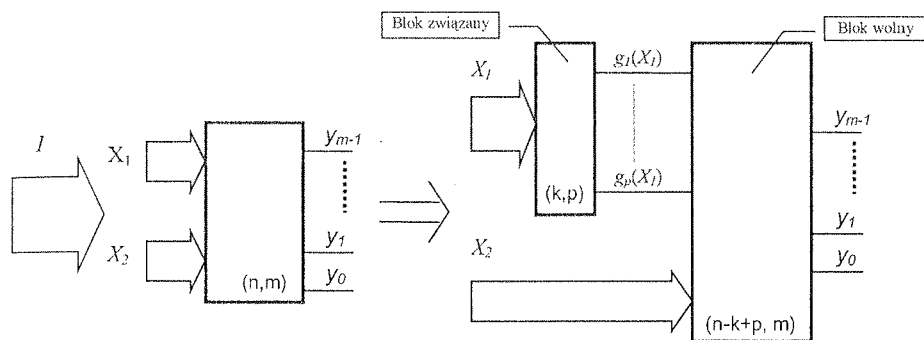
Funkcja $f : B^n \rightarrow B^m$ podlega dekompozycji tzn.

$$\begin{aligned} f(X_2, X_1) &= F[g_1(X_1), g_2(X_1), \dots, g_p(X_1), X_2] \\ X_1 &= \{I_{k-1}, \dots, I_1, I_0\}; X_2 = \{I_{n-1}, \dots, I_{k+1}, I_k\} \end{aligned} \quad (1)$$

wtedy i tylko wtedy, gdy złożoność kolumnowa macierzy podziałów $\nu(X_2|X_1) \leq 2^p$.

Powyższy model dekompozycji, obok rzadziej wykorzystywanego modelu dekompozycji nierozłącznej, stanowi podstawę szeregu algorytmów syntezy dedykowanej dla struktur FPGA typu tablicowego. Popularność dekompozycji rozłącznej wynika z faktu, iż w większości przypadków prowadzi ona do porównywalnych rozwiązań co dekompozycja nierozłączna. Dodatkowo nieznaczne korzyści uzyskiwane po zastosowaniu modelu dekompozycji nierozłącznej okupione są znaczną komplikacją algorytmów poszukiwania tego typu rozwiązań.

Wykorzystanie przedstawionego modelu dekompozycji rozłącznej odpowiada rozbiciu wielowyjściowego bloku scharakteryzowanego uporządkowaną parą liczb (liczba wejść, liczba wyjść) wynoszącą (n, m) na dwa inne o parametrach (k, p) i $(n-k+p, m)$. Rozbicie układu odpowiadające dekompozycji wielowyjściowej funkcji przedstawiono na rysunku 1.



Rys. 1. Rozłączna dekompozycja wielowyjściowej funkcji i odpowiadające jej rozbitcie układu

Fig. 1. Disjoint decomposition of multi-output function along with a circuit partitioning

Przykład 1

Rozważmy przykładową funkcję $f : B^5 \rightarrow B^2$ opisaną za pomocą matrycy podziałów przedstawionej na rysunku 2.

	<i>cde</i>							
<i>ab</i>	000	001	010	011	100	101	110	111
00	00	00	00	00	-1	00	01	00
01	0-	01	01	11	11	01	1-	1-
10	11	11	11	0-	00	11	--	-0
11	10	1-	00	11	-1	10	1-	11
	A	A	D	C	B	A	B	C

Rys. 2. Matryca podziałów przykładowej funkcji $f : B^5 \rightarrow B^2$ Fig. 2. Partition matrix of exemplary function $f : B^5 \rightarrow B^2$ 2^p.

dekom-
anej dla
z faktu,
dekom-
owaniu
ów po-

da roz-
(liczba
+ p, m).
awiono

Zbiór argumentów funkcji $I = \{a, b, c, d, e\}$ można podzielić na dwa rozłączne podzbiory, z których $X_1 = \{c, d, e\}$ zawiera zmienne skojarzone a kolumnami matrycy podziałów, natomiast $X_2 = \{a, b\}$, zmienne skojarzone z wierszami matrycy podziałów. Zbiory X_1, X_2 spełniają warunki twierdzenia o dekompozycji tzn. $X_1 \cup X_2 = I$ oraz $X_1 \cap X_2 = \emptyset$. Zbiory X_1, X_2 nazywane są odpowiednio zbiorem związanym i wolnym. Kolumny matrycy podziałów tworzą cztery różne wzorce kolumn A, B, C, D. Mówimy, że złożoność kolumnowa matrycy podziału $\nu(X_2|X_1) = 4$. Wzorce kolumn można rozróżnić za pomocą dwóch bitów. Prowadzi to ostatecznie do przykładowego przedstawienia analizowanej funkcji w postaci przedstawionej na rys. 3. Blok związany posiada dwa wyjścia odpowiadające funkcjom $g_1(X_1)$ oraz $g_2(X_1)$.

Celem dekompozycji funkcji jest odpowiedni podział projektowanego układu logicznego na podukłady o zadanej liczbie wejść i wyjść (realizacja układu cyfrowego w strukturach FPGA typu tablicowego zawierającego konfigurowalne bloki logiczne).

			$g_l(i) \quad g_z(i)$				
			ab	00	01	11	10
A B C D	$g_l(X_l)$	$g_z(X_l)$	00	00	01	00	00
			01	01	11	11	01
			10	11	00	00	11
			11	10	11	11	00
				A	B	C	D
			$yDy?$				

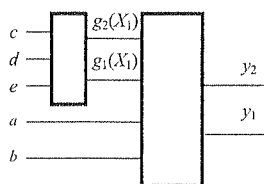
			de				
			c	00	01	11	10
0 1	$g_l(X_l)$	$g_z(X_l)$	0	A	A	C	D
			1	B	A	C	B

$$g_1(X_1) = de + \bar{c}d$$

$$g_2(X_1) = \bar{c}e + de$$

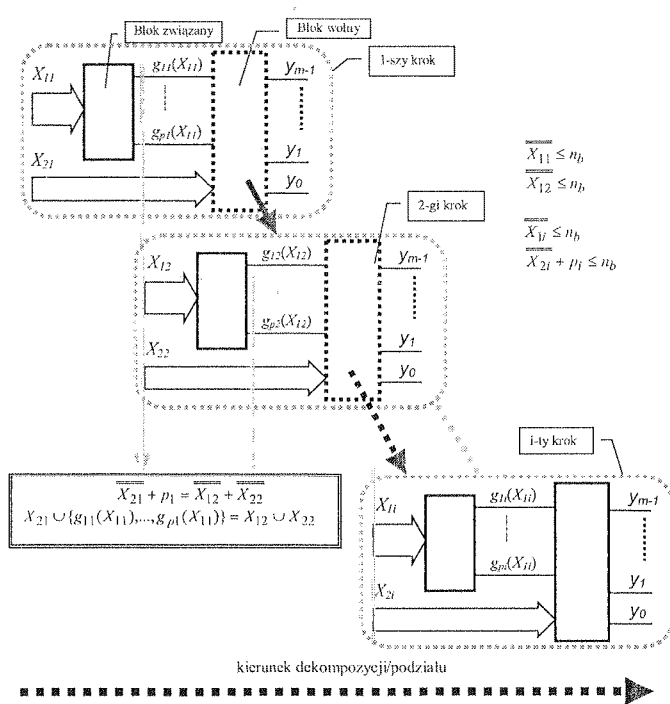
$$y_1 = a \bar{g}_1(X_1) \bar{g}_2(X_1) + a \bar{b} \bar{g}_2(X_1) + b g_2(X_1)$$

$$y_2 = \bar{a} b \bar{g}_2(X_1) + a \bar{b} \bar{g}_2(X_1) + b g_2(X_1) + a \bar{g}_1(X_1) g_2(X_1)$$



Rys. 3. Wynik dekompozycji oraz implementacja przykładowej dwuwyjściowej funkcji

Fig. 3. Result of decomposition and implementation of exemplary two-output function



Rys. 4. Istota dekompozycji metodą „od wejść do wyjść”

Fig. 4. The essence of the „input-to-output” decomposition

Ze względu na kierunek prowadzenia podziału układu w poszczególnych krokach, można wyróżnić metodę dekompozycji „od wejść do wyjść” lub „od wyjść do wejść”.

Pierwsza metoda polega na zapewnieniu realizowalności funkcji, najpierw w bloku związanym, a następnie poszukiwaniu rozbicia bloku wolnego. W procesie poszukiwania odpowiedniej dekompozycji należy wybierać podziały, które maksymalnie ograniczają liczbę wejść bloku wolnego. Jeżeli liczba wejść tego bloku jest większa od zadanej liczby wejść wykorzystywanego bloku logicznego (n_b) to należy ponownie poszukiwać dekompozycji zapewniającej rozbicie bloku wolnego poprzez wprowadzenie kolejnego poziomu logicznego. Istota metody dekompozycji „od wejść do wyjść” przedstawiona jest na rysunku 4.

Drugim sposobem podziału układu jest metoda polegająca na poszukiwaniu rozbicia układu metodą „od wyjść do wejść”. Istota tej metody polega na poszukiwaniu podziałów zapewniających realizowalność funkcji, najpierw w bloku wolnym, a następnie poszukiwaniu rozbicia bloku związanego. W procesie poszukiwania odpowiedniej dekompozycji należy wybierać podziały, które zapewniając realizowalność bloku wolnego w oparciu o bloki logiczne o zadanej liczbie wejść (n_b) minimalizują liczbę wyjść bloku związanego. Jeżeli liczba wyjść tego bloku jest większa od zadanej liczby wyjść wykorzystywanego bloku logicznego (n_b), to należy ponownie poszukiwać dekompozycji zapewniającej rozbicie bloku związanego poprzez wprowadzenie kolejnego poziomu logicznego. Istota metody dekompozycji „od wyjść do wejść” przedstawiona jest na rysunku 5.

Jeżeli wielowyjściowa funkcja podlega kilku dekompozycjom prostym zgodnym z zależnością (1) to możliwe jest poszukiwanie dekompozycji złożonych: wielokrotnych lub iteracyjnych. W pracy [9] można znaleźć szereg twierdzeń dotyczących dekompozycji złożonych jednowyjściowych funkcji. Należą do nich między innymi podstawowe twierdzenia o dekompozycji wielokrotnej i iteracyjnej przedstawione poniżej.

Niech $f(A, B, C) = B^n \rightarrow B^1$, gdzie A, B, C są zbiorami parami rozłącznymi takimi, że $\overline{A \cup B \cup C} = n$.

Twierdzenie 1 (o dekompozycji iteracyjnej):

Jeżeli $f(A, B, C) = F[\xi(A, B), C]$

i $f(A, B, C) = G[\phi(A), B, C]$

to

$$f(A, B, C) = F[\eta[\phi(A), B], C].$$

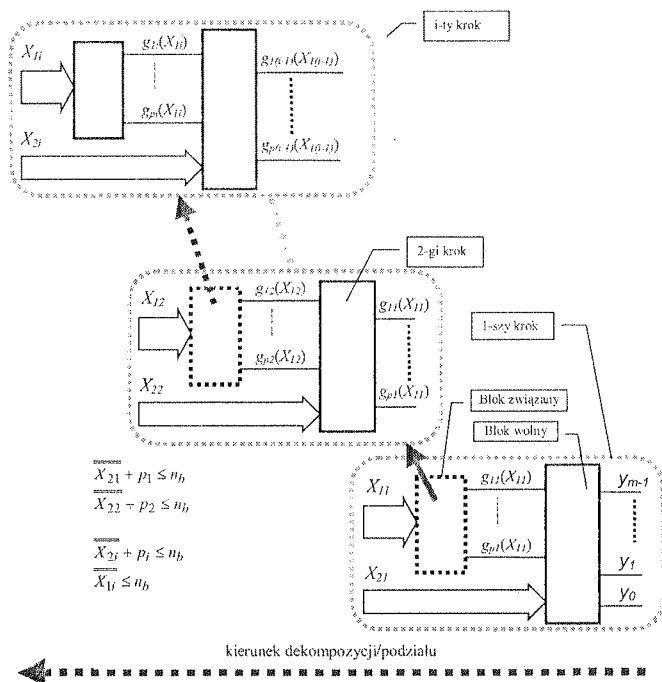
Twierdzenie 2: (o dekompozycji wielokrotnej):

Jeżeli $f(A, B, C) = F[\phi(A), B, C]$

i $f(A, B, C) = G[\chi(B), A, C]$

to

$$f(A, B, C) = H[\phi(A), \chi(B), C].$$



Rys. 5. Istota dekompozycji metodą „od wyjść do wejść”

Fig. 5. The essence of the "output-to-input" decomposition

Powyższe twierdzenia można uogólnić i przedstawić dla wielowyjściowej funkcji w następującej formie.

Niech $Y = f(I_{n-1}, \dots, I_1, I_0) = f(X_{11}, X_{12}, \dots, X_{1l}, X_2) : B^n \rightarrow B^m$ $f(I_{n-1}, \dots, I_1, I_0)$, gdzie $Y = y_{m-1}, \dots, y_1, y_0$. Niech $X_{11}, X_{12}, \dots, X_{1l}, X_2$ są zbiorami parami rozłącznymi, dla których $\overline{X_{11} \cup X_{12} \cup \dots \cup X_{1l} \cup X_2} = n$

Twierdzenie 3 (o dekompozycji iteracyjnej)

Jeżeli

$$\begin{aligned}
 Y &= F_1[G_1(X_{11}), X_{12}, X_{13}, \dots, X_{1l}, X_2] \\
 Y &= F_2[G_2(X_{11}), X_{12}, X_{13}, \dots, X_{1l}, X_2] \\
 Y &= F_3[G_3(X_{11}), X_{12}, X_{13}, \dots, X_{1l}, X_2] \\
 &\vdots \\
 Y &= F_l[G_l(X_{11}), X_{12}, X_{13}, \dots, X_{1l}, X_2]
 \end{aligned}$$

gdzie

$$G_1(X_{11}) = [g_{11}(X_{11}), g_{12}(X_{11}), \dots, g_{1p_1}(X_{11})]$$

gdzie

$$G_2(X_{11}, X_{12}) = [g_{21}(X_{11}, X_{12}), g_{22}(X_{11}, X_{12}), \dots, g_{2p_2}(X_{11}, X_{12})]$$

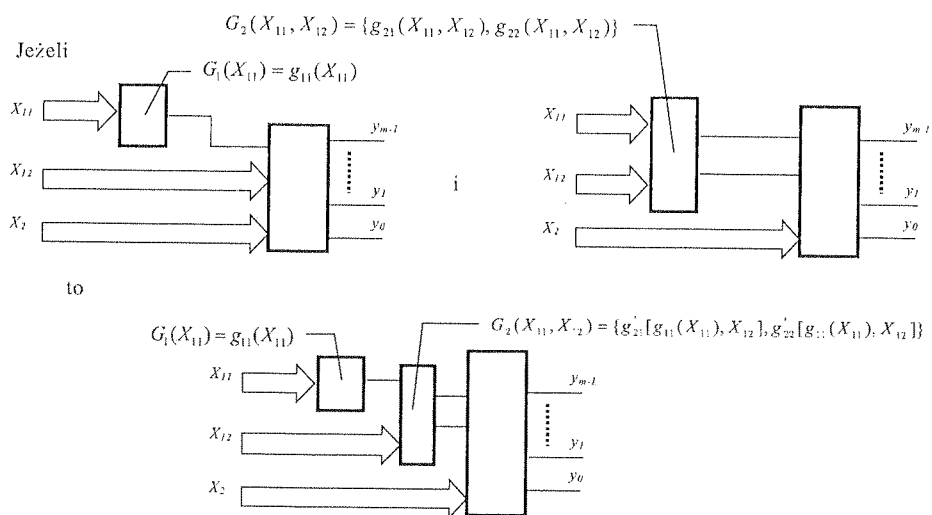
$$\dots$$

$$G_l(X_{11}, \dots, X_{1l}) = [g_{l1}(X_{11}, \dots, X_{1l}), g_{l2}(X_{11}, \dots, X_{1l}), \dots, g_{lp_l}(X_{11}, \dots, X_{1l})]$$

to

$$Y = F_l[G_l[\dots G_2[G_1(X_{11}), X_{12}], \dots, X_{1l}], X_2]$$

Istota twierdzenia wraz ze strukturą układu po podziale, będącym wynikiem dekompozycji iteracyjnej, przedstawiona jest na rys. 6.



Rys. 6. Podział układu będący wynikiem dekompozycji iteracyjnej wielowyjściowej funkcji

Fig. 6. Partitioning of circuits as a result of iterative decomposition of multi-output function

W podobny sposób można przedstawić twierdzenie o dekompozycji wielokrotnej.

Twierdzenie 4 (o dekompozycji wielokrotnej)

Jeżeli

$$Y = F_1[G_1(X_{11}), X_{12}, X_{13}, \dots, X_{1l}, X_2]$$

$$Y = F_2[X_{11}, G_2(X_{12}), X_{13}, \dots, X_{1l}, X_2]$$

$$\dots$$

$$Y = F_l[X_{11}, X_{12}, X_{13}, \dots, G_l(X_{1l}), X_2]$$

gdzie

$$G_1(X_{11}) = [g_{11}(X_{11}), g_{12}(X_{11}), \dots, g_{1p_1}(X_{11})]$$

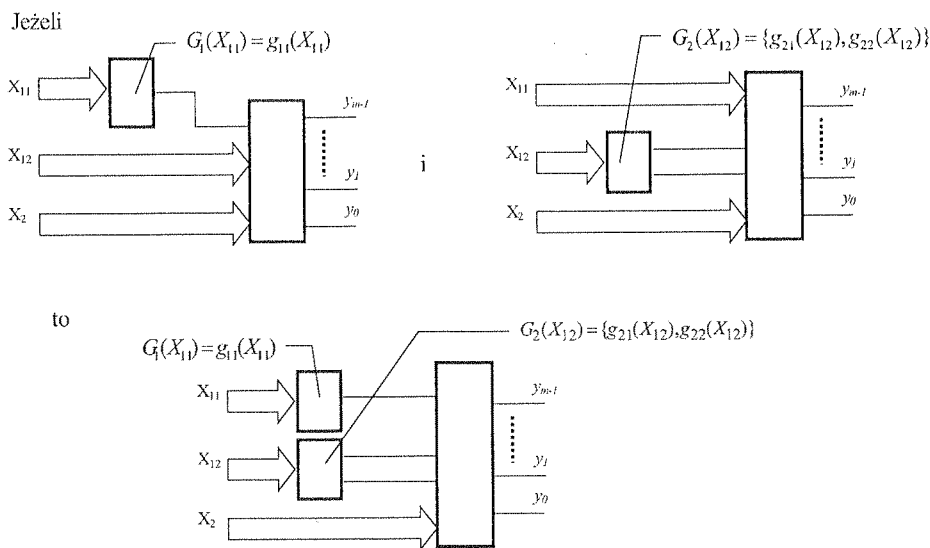
$$G_2(X_{12}) = [g_{21}(X_{12}), g_{22}(X_{12}), \dots, g_{2p_2}(X_{12})]$$

$$G_l(X_{1l}) = [g_{l1}(X_{1l}), g_{l2}(X_{1l}), \dots, g_{lp_l}(X_{1l})]$$

to

$$Y = H[G_1(X_{11}), G_2(X_{12}), \dots, G_l(X_{1l}), X_2]$$

Istota twierdzenia wraz z strukturą układu po podziale, będącym wynikiem dekompozycji wielokrotnej, przedstawiona jest na rys. 7.



Rys. 7. Podział układu będący wynikiem dekompozycji wielokrotnej wielowyjściowej funkcji

Fig. 7. Partitioning of circuits as a result of multiple decomposition of multi-output function

W celu doboru odpowiedniej dekompozycji prostej lub złożonej, stanowiącej teoretyczną podstawę podziału układu i dopasowanie go do struktury bloków logicznych zawartych w układzie programowalnym, zachodzi potrzeba scharakteryzowania poszczególnych dekompozycji.

Niech β będzie maksymalną liczbę wejść podukładów powstających po podziale będącym wynikiem odpowiedniej dekompozycji wielowyjściowej funkcji. Dla podstawowego modelu dekompozycji opisanego zależnością (1) wartość tego współczynnika wynosi $\beta = \max(\overline{X1}, \overline{X2} + p)$.

Większe ograniczenie liczby wejść bloków powstających po podziale uzyskuje się po wyszukaniu odpowiedniej dekompozycji złożonej. W celu przedstawienia istoty takiego poszukiwania konieczne jest wprowadzenie kilku pojęć.

Nieco
wej funk
Podział

przed po

Defin

Tablic

miarach

nom — a

nach odp

podziałów

Defin

Wier

powiadaj

w różnyc

Defin

Wier

gdy X_{1j}

Okaz

dzic do a

Twie

Funk

r wierszy

nad wier

pierwszy

Dow

cji 1,3.

Twie

Funk

k wierszy

Dow

cji 1,2.

Przy

Nieco

do podzi

podziałów

wiersze

Anal

jących da

dekompo

czynnikó

będącym

współczy

Niech P_i będzie podziałem zbioru argumentów $I = \{I_{n-1}, \dots, I_1, I_0\}$ wielowyjściowej funkcji na dwa rozłączne podzbiory $X_1 = \{I_{k-1}, \dots, I_1, I_0\}$ i $X_2 = \{I_{n-1}, \dots, I_{k+1}, I_k\}$. Podział P_i prowadzi do ograniczenia liczby wejść podukładów (w stosunku do układu przed podziałem) wtedy, gdy $\bar{I} = n > \beta = \max(\bar{X}_1, \bar{X}_2 + p)$.

Definicja 1:

Tablicą podziałów wielowyjściowej funkcji $f : B^n \rightarrow B^m$ nazywamy tablicę o wymiarach $r \times n$, której wierszom przyporządkowane są podziały $P_1, P_2, \dots, P_r$, kolumnom — argumenty funkcji, natomiast elementami tablicy są „*”, umieszczone w kolumnach odpowiadających zmiennym, należącym do zbiorów związanych poszczególnych podziałów.

Definicja 2:

Wiersze i, j tablicy podziałów nazywamy rozłącznymi wtedy i tylko wtedy, gdy odpowiadające im zbiory związane spełniają zależność $X_{1i} \cap X_{1j} = \emptyset$. (tzn.: „*” występują w różnych kolumnach tablicy podziałów).

Definicja 3:

Wiersz i -ty tablicy podziałów dominuje nad wierszem j -tym wtedy i tylko wtedy, gdy $X_{1j} \supset X_{1i}$. (tzn.: „*” w wierszu j są podzbiorem „*” w wierszu i)

Okazuje się, że poszukiwanie odpowiedniej dekompozycji funkcji można sprowadzić do analizy wierszy tablicy podziałów.

Twierdzenie 5:

Funkcja $f : B^n \rightarrow B^m$ podlega dekompozycji iteracyjnej wtedy, gdy spośród r wierszy tablicy podziałów można wybrać s wierszy takich, że wiersz s dominuje nad wierszem $(s-1)$, $(s-1)$ nad $(s-2)$, ..., wiersz drugi dominuje nad wierszem pierwszym.

Dowód twierdzenia wynika bezpośrednio z twierdzenia 3 oraz przyjętych definicji 1,3.

Twierdzenie 6:

Funkcja $f : B^n \rightarrow B^m$ podlega dekompozycji wielokrotnej wtedy, gdy spośród k wierszy tablicy podziałów można wybrać l wierszy wzajemnie rozłącznych.

Dowód twierdzenia wynika bezpośrednio z twierdzenia 4 oraz przyjętych definicji 1,2.

Przykład 2:

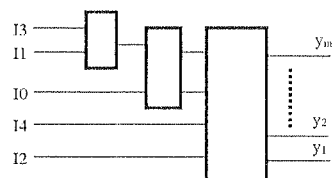
Niech przykładowa funkcja $f : B^n \rightarrow B^m$ podlega dekompozycjom, które prowadzą do podziałów argumentów funkcji, przedstawionych w wierszach P_1, P_2 i P_3 tablicy podziałów. Ponieważ wiersz P_1 tablicy podziałów dominuje nad wierszem P_3 natomiast wiersze P_2 i P_3 są rozłączne stąd możliwe są realizacje przedstawione na rysunku 8.

Analiza wierszy tablicy podziałów oraz współczynników p_i oraz β charakteryzujących dany podział, pozwala po uwzględnieniu twierdzeń 5, 6 wyszukiwać wszystkie dekompozycje iteracyjne oraz wielokrotne. Możliwe jest dla nich wyznaczenie współczynników β_{DI} , β_{DW} charakteryzujących układ uzyskany po podziale, odpowiednio będącym wynikiem dekompozycji iteracyjnej oraz wielokrotnej. Na podstawie wartości współczynników β_{DI} oraz β_{DW} można wybrać dekompozycje złożone, umożliwiające

	I4	I3	I2	I1	I0	p_i	β_i		
P_1		*		*	*	2	4	$X_{11} = \{I3, I1, I0\}$	$X_{21} = \{I4, I2\}$
P_2	*		*		*	2	4	$X_{12} = \{I4, I2, I0\}$	$X_{22} = \{I3, I1\}$
P_3		*		*		1	4	$X_{13} = \{I3, I1\}$	$X_{23} = \{I3, I2, I0\}$

a)

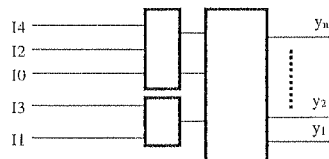
	I4	I3	I2	I1	I0	p_i	β_i
P_1		*		*	*	2	4
P_3		*		*		1	4



$$\beta_{DI} = \max(\overline{X_{13}}; \overline{X_{11}} - \overline{X_{13}} + p_3; \overline{X_{21}} + p_1) = \max(2, 2, 4) = 4$$

b)

	I4	I3	I2	I1	I0	p_i	β_i
P_2	*		*		*	2	4
P_5		*		*		1	4



$$\beta_{DW} = \max(\overline{X_{12}}; \overline{X_{13}}; p_2 + p_3) = \max(3, 2, 3) = 3$$

Rys. 8. Podział układu będący wynikiem dekompozycji iteracyjnej (a) i wielokrotnej (b) przykładowej funkcji

Fig. 8. Partitioning of circuits as a result of iterative decomposition (a) and multiple decomposition (b) of exemplary function

realizację układu po podziale, na blokach logicznych o zadanej liczbie wejść. Metody wyboru odpowiednich dekompozycji prostych lub złożonych, prowadzące do realizacji układów na blokach logicznych o zadanej liczbie wejść zostaną przedstawione w dalszej części pracy.

3. METODY POSZUKIWANIA PODZIAŁU PROJEKTOWANEGO UKŁADU

Pierwszoplanowe znaczenie w procesie syntezy, opartym na elementach dekompozycji, ma dobór w kolejnych krokach odpowiednich podziałów argumentów funkcji. Istotne znaczenie ma również odpowiednie wykorzystanie dekompozycji złożonych, spełniających warunki realizowalności w założonym bloku związanym lub wolnym w zależności od kierunku poszukiwania dekompozycji.

Rozpatrzmy w pierwszej kolejności przypadek dekompozycji „od wejść do wyjść”. Załóżmy, że chcemy zrealizować układ cyfrowy w oparciu o bloki logiczne typu tablicowego zawierające n_b –wejść i m_b –wyjść.

Po to, żeby zapewnić realizowalność bloku związanego w oparciu o blok o n_b –wejściach (patrz rys. 4) należy rozpatrzyć podziały prowadzące do ograniczenia liczby

wejść p
wybiera
bloku w
realizac
warunk

W
mentów
to znac
działów
dla któr
warunek
do mak

Kol
złożony
6) wysz
kroku a
by wejś
fakt, że
dekomp
kompoz
najlepie

Roz
todzie t
wejściac
których
kompoz
ogranicz
rzystany
dekomp
łów, wy
których
($\overline{X_{li}} \leq n$
wia wyb
by wejś
potrzebn
że ze w
maksym
rozwiąz
liczbę b

{4,12}
{3,11}
{3,12,10}

wejść podukładów po podziale (tzn. $\beta < \bar{I}$), dla których $\bar{X}_1 \leq n_b$. Najkorzystniejsze jest wybieranie podziałów, które prowadzą do maksymalnego ograniczenia liczby wejść bloku wolnego równocześnie minimalizując liczbę bloków logicznych potrzebnych do realizacji bloku związanego. Powyższa zasada sprowadza się do spełnienia dwóch warunków: $\bar{X}_2 + p = \min$ oraz $\left\lceil \frac{p}{m_b} \right\rceil = \min$, gdzie p jest liczbą wyjść bloku związanego.

W przypadku dekompozycji „od wyjść do wejść” poszukiwany jest podział argumentów zapewniający realizowalność bloku wolnego w oparciu o blok o n_b – wejściach, to znaczy spełniający zależność $\bar{X}_2 + p \leq n_b$. W tej sytuacji należy poszukiwać podziałów prowadzących do ograniczenia liczby wejść podukładów po podziale ($\beta < \bar{I}$), dla których $\bar{X}_2 < n_b$. Najkorzystniejsze jest wybieranie podziałów, które zapewniając warunek realizowalności funkcji w bloku wolnym ($\bar{X}_2 + p \leq n_b$) prowadzą równocześnie do maksymalnego ograniczenia liczby wejść bloku związanego ($\bar{X}_1 = \min$).

position

Metody
realizacji
w dalszej

ADU

dekom-
funkcji.
łożonych,
wolnym

o wyjść”.
pu tabli-

ok o n_b –
ia liczby

Kolejnym niezwykle istotnym problemem jest sposób wykorzystania dekompozycji złożonych. Mając tablicę podziałów można na podstawie jej wierszy (twierdzenie 5, 6) wyszukiwać dekompozycje złożone. Poszukiwanie może odbywać się w jednym kroku analizy wielowyjściowej funkcji. Prowadzi to do znacznego ograniczenia liczby wejść poszczególnych bloków poddawanych dalszym rozbiciom. Należy podkreślić fakt, że ograniczenie to uzyskuje się w jednym kroku rozbicia układu. W przypadku dekompozycji metodą „od wejść do wyjść” w prosty sposób można wyszukiwać dekompozycje wielokrotne, w odróżnieniu od metody „od wyjść do wejść”, w której najlepiej poszukiwać dekompozycje iteracyjne.

Rozpatrzmy w pierwszej kolejności metodę syntezy „od wejść do wyjść”. W metodzie tej, w celu zapewnienia realizowalności bloku związanego w komórce o n_b – wejściach, można znacznie ograniczyć zbiór poszukiwanych podziałów, do takich, dla których $\bar{X}_1 \leq n_b$. Analiza wierszy tablicy podziałów umożliwia wybór najlepszej dekompozycji prostej lub wielokrotnej tzn. takiej, dla której uzyskuje się maksymalne ograniczenie liczby wejść bloku wolnego, minimalizując równocześnie liczbę wykorzystanych bloków potrzebnych do realizacji bloku związanego. Łatwość poszukiwania dekompozycji wielokrotnych wynika z faktu, że wystarczy analizować tablicę podziałów, wyszukując grupy rozłącznych wierszy opisujących podziały (twierdzenie 6), dla których moc zbiorów związanych jest mniejsza od liczby wejść bloków logicznych ($\bar{X}_{1i} \leq n_b$). Analiza dekompozycji prostych oraz dekompozycji wielokrotnych umożliwia wybór takiej dekompozycji, dla której uzyskuje się maksymalne ograniczenie liczby wejść bloku wolnego, minimalizując równocześnie liczbę wykorzystanych bloków, potrzebnych do realizacji bloku związanego. Przeprowadzone eksperymenty wskazały, że ze względu na szybkość syntezy najlepiej wybierać rozwiązania prowadzące do maksymalnego ograniczenia liczby wejść bloku wolnego. Jeżeli występuje kilka takich rozwiązań, to powinno zostać wybrane takie, dla którego wykorzystano minimalną liczbę bloków logicznych do realizacji bloku związanego.

W metodzie „od wyjść do wejść” analizowane są tylko te podziały, które zapewniają realizowalność bloku wolnego w oparciu o komórkę posiadającą n_b -wejść. Ponieważ blok związany ma co najmniej jedno wyjście ($p = 1$) stąd analizę podziałów można ograniczyć tylko do takich podziałów, dla których $\overline{\overline{X}}_2 < n_b$. Tym razem na podstawie analizy wierszy tablicy podziałów można wybrać podziały spełniające warunki dekompozycji iteracyjnej (twierdzenie 5) zapewniające realizowalność bloku wolnego w komórce o n_b -wejściach. Powyższe rozważania pozwalają na przedstawienie istoty algorytmów obydwu metod syntezy funkcji wielowyjściowych w strukturach FPGA typu tablicowego. W opisie nie ujęto prostych podziałów wyjść [12] wykonywanych zawsze w przypadku, gdy znaleziono podział o zadawalającej liczbie wejść i liczbie wyjść większej od m_b .

Metoda syntezy „od wejść do wyjść”

- A. Wyznaczanie złożoności kolumnowej maczy podziału wielowyjściowych funkcji, dla których $\overline{\overline{X}}_1 \leq n_b$; *(tylko takie podziały zapewniają realizowalność bloku związanego w oparciu o komórki zawierające n_b -wejść)*
- B. Utworzenie tablicy podziałów wielowyjściowej funkcji $f : B^n \rightarrow B^m$, zawierającej podziały, dla których $p < \overline{\overline{X}}_1$, gdzie p jest liczbą wyjść bloku związanego; *(takie podziały prowadzą do ograniczenia liczby wejść bloku wolnego w stosunku do układu przed podziałem tzn. $\overline{\overline{X}}_2 + p < \overline{\overline{X}}_2 + \overline{\overline{X}}_1$)*
- C. Na podstawie wierszy tablicy podziałów poszukiwana jest dekompozycja prosta lub dekompozycja wielokrotna prowadząca do maksymalnego ograniczenia liczby wejść bloku wolnego tzn. $\overline{\overline{X}}_2 + p = \min$ lub $\overline{\overline{X}}_2 + p_{11} + p_{12} + \dots + p_{1l} = \min$, gdzie p oznacza liczbę wyjść bloku związanego dla dekompozycji prostej, natomiast $p_{11}, p_{12}, \dots, p_{1l}$ oznaczają liczby wyjść bloków związanych, skojarzonych z wybranymi, rozłącznymi wierszami tablicy podziałów, prowadzącymi do dekompozycji wielokrotnej (twierdzenie 4, 6);
- D. Jeżeli istnieje kilka rozwiązań spełniających powyższy warunek prowadzący do maksymalnego ograniczenia liczby wejść bloku wolnego, to wybierany jest taki podział, który prowadzi do wykorzystania mniejszej liczby komórek do realizacji bloku związanego
- E. W przypadku, gdy nie istnieje dekompozycja ograniczająca liczbę wejść bloku wolnego tzn. nie istnieje podział, dla którego $p < \overline{\overline{X}}_1$ to wykonywany jest prosty podział wyjść [12] na bloki o liczbie wyjść $m_i \leq m_b$ i ponownie poszukiwany jest podział „mniejszych” wielowyjściowych funkcji p. A

Metoda syntezy „od wyjść do wejść”

- A. Wyznaczanie złożoności kolumnowej maczy podziału wielowyjściowych funkcji, dla których $\overline{\overline{X}}_2 < n_b$; *(tylko takie podziały mogą zapewnić realizowalność bloku wolnego w oparciu o komórki zawierające n_b -wejść)*

B. Utworzenie tablicy podziałów wielowyjściowej funkcji $f : B^n \rightarrow B^m$, zawierającej podziały, dla których $p < \overline{\overline{X}}_1$, gdzie p jest liczbą wyjść bloku związanego; *(takie podziały prowadzą do ograniczenia liczby wejść bloku wolnego w stosunku do układu przed podziałem tzn. $\overline{\overline{X}}_2 + p < \overline{\overline{X}}_2 + \overline{\overline{X}}_1$)*

C. Na podstawie wierszy tablicy podziałów poszukiwana jest dekompozycja prosta lub dekompozycja wielokrotna prowadząca do maksymalnego ograniczenia liczby wejść bloku wolnego tzn. $\overline{\overline{X}}_2 + p = \min$ lub $\overline{\overline{X}}_2 + p_{11} + p_{12} + \dots + p_{1l} = \min$, gdzie p oznacza liczbę wyjść bloku związanego dla dekompozycji prostej, natomiast $p_{11}, p_{12}, \dots, p_{1l}$ oznaczają liczby wyjść bloków związanych, skojarzonych z wybranymi, rozłącznymi wierszami tablicy podziałów, prowadzącymi do dekompozycji wielokrotnej (twierdzenie 4, 6);

D. Jeżeli istnieje kilka rozwiązań spełniających powyższy warunek prowadzący do maksymalnego ograniczenia liczby wejść bloku wolnego, to wybierany jest taki podział, który prowadzi do wykorzystania mniejszej liczby komórek do realizacji bloku związanego

E. W przypadku, gdy nie istnieje dekompozycja ograniczająca liczbę wejść bloku wolnego tzn. nie istnieje podział, dla którego $p < \overline{\overline{X}}_1$ to wykonywany jest prosty podział wyjść [12] na bloki o liczbie wyjść $m_i \leq m_b$ i ponownie poszukiwany jest podział „mniejszych” wielowyjściowych funkcji p. A

Przy
Rozp
wać w s
i jednym
rych $\overline{\overline{X}}_1$
których
odpowia
na rysun
Spec
zycje pro
w komór
P₁ i P₄ p
wejść blo
i P₄ opis
nego zaw
ta prowa
W p
których
dla który
uzyskana
Tylko
zowalnos
w tym p
P₂, w kt
cyjną, kt
zapewnia

- B. Utworzenie tablicy podziałów wielowyjściowej funkcji $f : B^n \rightarrow B^m$, zawierającej tylko podziały, dla których $p < \overline{X}_1$
- C. Na podstawie wierszy tablicy podziałów wyszukiwane są dekompozycje proste, zapewniające realizowalność bloku wolnego w komórce o n_b -wejściach, tzn. podziały, dla których $\overline{X}_2 + p \leq n_b$
- D. Na podstawie wierszy tablicy podziałów wyszukiwane są dekompozycje iteracyjne zapewniające maksymalne ograniczenia liczby wejść bloku związanego
- E. Jeżeli istnieje kilka dekompozycji prostych lub iteracyjnych spełniających warunek minimalizacji liczby wejść bloku związanego to wybierane jest to rozwiązanie, które prowadzi do wykorzystania mniejszej liczby bloków do realizacji bloku wolnego
- F. W przypadku, gdy nie istnieje dekompozycja, dla której $p < \overline{X}_1$ to wykonywany jest prosty podział wyjść na bloki o liczbie wyjść $m_i \leq m_b$ i ponownie poszukiwany jest podział „mniejszych” wielowyjściowych funkcji p. A

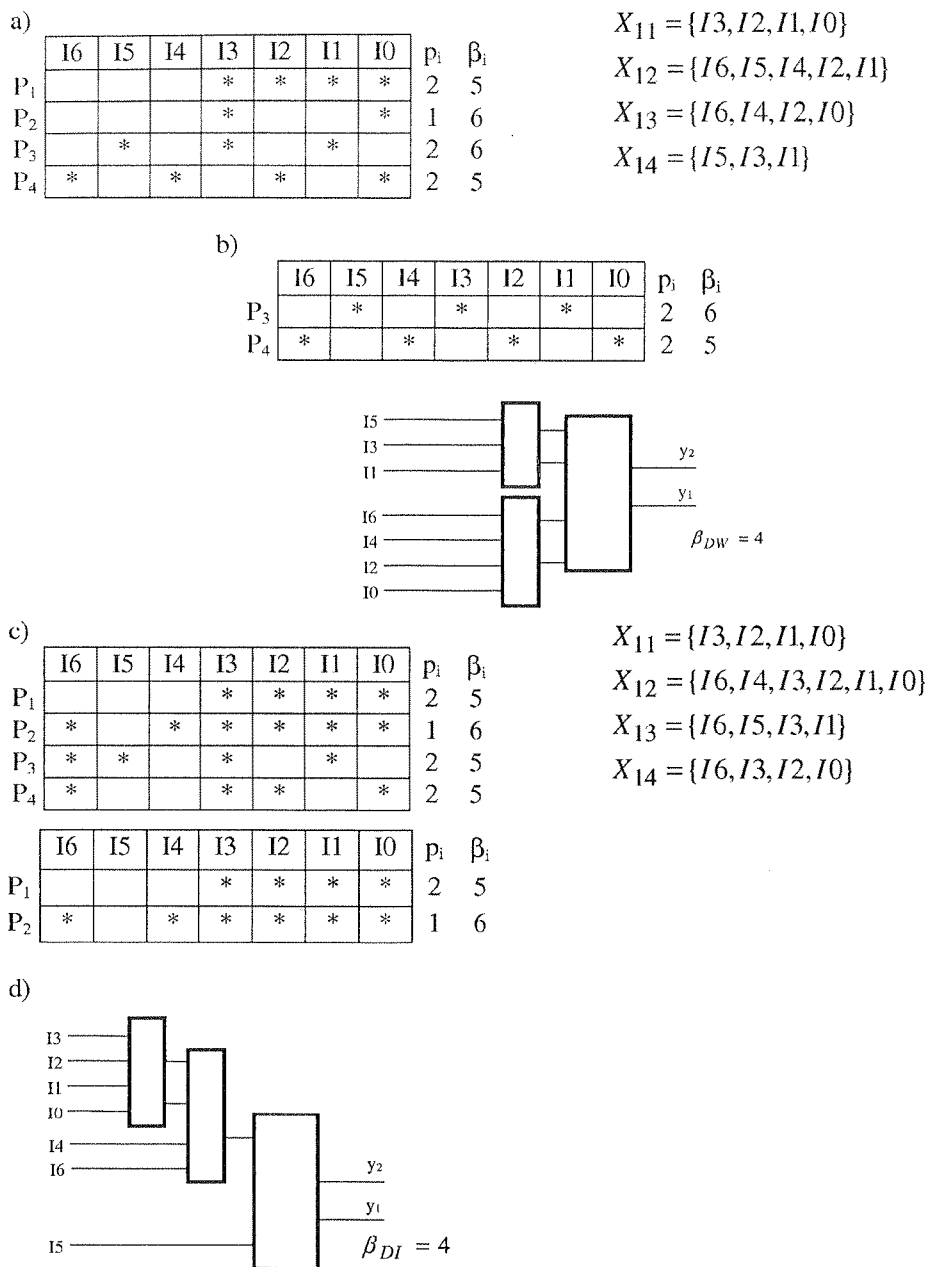
Przykład 3:

Rozpatrzmy przykładową funkcję $f : B^7 \rightarrow B^2$. Załóżmy, że chcemy ją zrealizować w strukturze FPGA zawierającej bloki logiczne typu tablicowego o 4-wejściach i jednym wyjściu. W metodzie „od wejść do wyjść” analizujemy podziały, dla których $\overline{X}_1 \leq n_b = 4$ i tworzymy tablicę podziałów, zawierającą wszystkie podziały, dla których $p < 4$, gdzie p jest liczbą wyjść bloku związanego. Tablica podziałów wraz odpowiadającymi poszczególnym wierszom zbiorami związanymi przedstawiona jest na rysunku 9a.

Specyfika syntezy metodą „od wejść do wyjść” powoduje, że wszystkie dekompozycje proste zawarte w tablicy podziałów, zapewniają realizowalność bloku związanego w komórce o 4-wejściach. Dwie z nich przedstawione w tablicy podziałów w wierszu P_1 i P_4 prowadzą do podziałów, wynikiem których jest minimalna liczba ($\overline{X}_{2i} + p_i = 5$) wejść bloku wolnego. W tablicy podziałów istnieją jednak dwa wiersze rozłączne P_3 i P_4 opisujące dekompozycję wielokrotną (twierdzenie 4, 6), prowadzącą do bloku wolnego zawierającego mniejszą liczbę wejść równą cztery ($p_3 + p_4 = 4$). Dekompozycja ta prowadzi do końcowego rozwiązania przedstawionego na rysunku 9b.

W przypadku syntezy metodą „od wyjść do wejść” analizujemy podziały, dla których $\overline{X}_2 < n_b = 4$ i tworzymy tablicę podziałów, zawierającą wszystkie podziały, dla których $p < \overline{X}_1$, gdzie p jest liczbą wyjść bloku związanego. Tablica podziałów uzyskana w tym przypadku przedstawiona jest na rysunku 9c.

Tylko dekompozycja opisana przez wiersz P_2 tablicy podziałów zapewnia realizowalność bloku wolnego w komórce o 4-wejściach. Liczba wejść bloku związanego w tym przypadku wynosi 6. W tablicy podziałów występuje jednak para wierszy P_1 , P_2 , w której wiersz P_2 dominuje nad wierszem P_1 . Opisują one dekompozycję iteracyjną, która prowadzi do znaczniejszego ograniczenia liczby wejść bloku związanego, zapewniając końcową realizację przedstawioną na rysunku 9d.



Rys. 9. Podział układu będący wynikiem dekompozycji „od wejść do wyjść” a.) b.); podział układu będący wynikiem dekompozycji „od wyjść do wejść” c.) d.); (opis w tekście)

Fig. 9. Partitioning of circuits as a result of decomposition „from input to output” a.) b.) Partitioning of circuits as a result of decomposition „from output to input” c.) d.); (description is in the text)

Przed
przeznacz
W celu zw
zycji wiel
zawierając
dla strukt
(CLB —
Wyniki po

Metod
stałych pr
Warto zw
lepsze roz
sumaryczn
układy tes
„od wejść
do wyjść”
Uzyskuje
ręczną lic

4. WYNIKI EKSPERYMENTÓW

Przedstawione metody syntezy zostały zaimplementowane w systemie Decomp, przeznaczonym do syntezy układów cyfrowych w strukturach FPGA typu tablicowego. W celu zweryfikowania efektywności przedstawionych metod poszukiwania dekompozycji wielowyjściowej funkcji, umożliwiającej jego realizację na blokach logicznych zawierających n_b –wejść i m_b –wyjść, przeprowadzono syntezę układów testowych [22] dla struktur FPGA serii 3000 składających się z konfigurowalnych bloków logicznych (CLB — Configurable Logic Block), zawierających 5–we, 1–wy lub 4–we, 2–wy. Wyniki porównania zawarte są w tabeli 1.

Tabela 1

Porównanie zaproponowanych metod syntezy (B — liczba bloków logicznych;
P — liczba poziomów logicznych)

Performance comparison of proposed methods (B — number of logic blocks;
P — number of logic levels)

	metoda „od wejść do wyjść”		metoda „od wyjść do wejść”	
	B	P	B	P
5xp1	11	2	11	2
9sym	5	3	7	3
Bw	27	1	27	1
Clip	15	3	16	3
F51m	10	3	11	3
Misex1	10	3	13	3
Rd73	5	2	6	3
Rd84	7	3	8	3
Root	19	3	21	4
Sao2	20	5	22	6
Z4m1	4	2	4	2
Σ	133	30	146	33

Metodą „od wejść do wyjść” uzyskano 8 (72%) lepszych rozwiązań. Dla pozostałych przypadków uzyskano równą liczbę bloków logicznych w obydwu metodach. Warto zwrócić uwagę na fakt, że metoda „od wejść do wyjść” czasami daje znacznie lepsze rozwiązania od metody „od wyjść do wejść” np.: 9sym, misex1. Porównując sumaryczną liczbę bloków logicznych wykorzystywanych przez wszystkie, analizowane układy testowe, metoda „od wejść do wyjść” potrzebuje 10% więcej bloków niż metoda „od wyjść do wejść”. Podobnie jak w przypadku bloków logicznych metoda „od wejść do wyjść” jest również skuteczniejsza pod względem liczby poziomów logicznych. Uzyskuje się nią lepsze rozwiązania jednak tylko w 3 przypadkach. Porównując sumaryczną liczbę poziomów logicznych, metodą „od wyjść do wejść” uzyskano rozwiązania

zawierające 10% więcej poziomów logicznych. Wyniki eksperymentów wskazują na większą efektywność metody dekompozycji „od wejść do wyjść” zarówno po kątem liczby bloków jak i poziomów logicznych. Złożoność zaprezentowanych algorytmów sprawia, że mogą być one wykorzystywane dla funkcji z niewielką liczbą argumentów (praktycznie kilkunastu). Złożoność dwóch przedstawionych metod syntezy jest porównywalna.

Dla większości analizowanych układów testowych (np. *rd73*, *rd84*, *9sym*) czas syntezy był krótszy niż 1 s (PC Pentium Celeron 300MHz). Maksymalny czas syntezy dla układu testowego *sao2* wyniósł około 3 s.

Uzyskane wyniki porównano (tabela 2) z wynikami otrzymanymi za pomocą innych metod syntezy, opublikowanymi w pracach [1,4,7,9,11,19,21,22,24,25,26,29,34] (spośród różnych publikacji wybrano najlepsze rozwiązania).

Tabela 2

Wyniki eksperymentów (liczba bloków / liczba warstw)

Experimental results (number of blocks / number of levels)

	ALTO	ASYL	Chortle	Demain	FlowMap	FGSyn	Mispga	TOS (ISODec)	TRADE	Decomp „od we do wy”
5xpl	19/2	13/4	20/3	9/2	22/3	9	17/3	-	11/2	11/2
9sym	7/3	8/3	41/5	5/3	60/5	7	7/3	8	6/3	5/3
Bw	-	27/6	-	-	-	27	27/1	-	27/1	27/1
Clip	31/5	33/6	-	16/2	-	18	23/4	22	29/4	15/3
F51m	15/3	14/6	-	10/2	-	8	11/4	-	9/3	10/3
Misex	14/3	13/2	14/2	8/2	15/2	8	9/3	-	14/2	10/3
Rd73	8/2	8/4	-	5/2	-	5	7/3	7	5/2	5/2
Rd84	13/3	14/4	41/4	7/2	38/6	8	12/3	10	8/3	7/3
Root	-	-	-	16/3	-	-	-	-	21/3	19/3
Sao2	37/5	30/3	-	18/3	-	25	28/5	22	27/3	20/5
z4m1	5/2	4/2	3	4/2	-	4	4/2	4	4/2	4/2

5. PODSUMOWANIE

Podstawowym problemem syntezy logicznej układów cyfrowych realizowanych w oparciu o struktury FPGA typu tablicowego jest problem podziału projektowanego układu na poszczególne konfigurowalne bloki logiczne. Sposób podziału jest oparty na teoretycznych podstawach związanych z dekompozycją funkcji. Optymalny sposób podziału układu na obecnym poziomie wiedzy nie jest znany. Niniejszy artykuł zawiera porównanie dwóch metod syntezy bazujących na dekompozycji. Wyniki eksperymen-

tów je
do wyj
dzi do
metoda
nie sy
tworzą

1. P. thoc 199
2. Alt
3. R. Sym
4. B. Eur
5. S. Gat
6. S. Bas sign
7. J. zati Janu
8. H. thea
9. H. Prin
10. J - I gori (VL
11. D. kom leko
12. D. PLL
13. Y. Imp
14. C h Fun
15. T. ss. 5
16. R. Imp Nov

azują na
o kątem
orytmów
rgumen-
tezy jest

m) czas
syntezy

nością in-
6,29,34]

Tabela 2

comp
do wy"
1/2
3/3
7/1
5/3
0/3
0/3
5/2
7/3
9/3
0/5
4/2

owanych
owanego
st oparty
y sposób
d zawiera
berymen-

tów jednoznacznie potwierdzają, że podział układu (dekompozycja) metodą „od wejść do wyjść” pod każdym względem (liczba bloków/liczba poziomów logicznych) prowadzi do lepszych rozwiązań. Ponieważ czas syntezy jak i złożoność algorytmów w obu metodach są porównywalne dalsze prace powinny być ukierunkowane na udoskonalenie syntezy metodą „od wejść do wyjść” poprzez np. heurystyczny dobór zmiennych tworzących zbiory związane.

6. BIBLIOGRAFIA

1. P. Abouzeid, B. Babba, M. Crastes, G. Saucier: *Input-Driven Partitioning Methods and Application to Synthesis on Table-Lookup-based FPGAs*. IEEE Trans on CAD, 12, No. 7, 1993, pp. 913-925
2. Altera, AMD, Lattice, Xilinx Data Book.
3. R. L. Ashenurst: *The decomposition of switching functions*. Proceedings of an International Symposium on the Theory of Switching, April 1957
4. B. Babba, M. Crastes, G. Saucier: *Input driven synthesis on PLDs and PGAs*. The European Conference on Design Automation, Brussels (Belgium), March 1992
5. S. D. Brown, R. J. Francis, J. Rose, Z. G. Vranesic: *Field Programmable Gate Arrays*. Boston, Kluwer Academic Publishers, 1993, pp. 45-86
6. S. Chang, M. Marek-Sadowska, T. Hwang: *Technology Mapping for TLU FPGA's Based on Decomposition of Binary Decision Diagrams*. IEEE Transactions on Computer-Aided Design, 15(10), October 1996, pp. 1226-1235
7. J. Cong, Y. Ding: *FlowMap: An Optimal Technology Mapping Algorithm for Delay Optimization in Lookup-Table Based FPGA Design*. IEEE Transactions on Computer-Aided Design, 13(1), January 1994, pp. 1-12
8. H. A. Curtis: *Generalized tree circuit — the basic building block of an extended decomposition theory*. J.ACM, vol. 10, 1963, pp. 562-581
9. H. A. Curtis: *The Design of switching Circuits* New Jersey, D.van Nostrand Company, Inc., Princeton, 1962
10. J-D. Huang, J-Y Jou, W-Z Shen: *ALTO: An Iterative Area/Performance Tradeoff Algorithm for LUT-Based FPGA Technology Mapping*. IEEE Transactions on Very Large Integration (VLSI) Systems Vol. 8, NO. 4, August 2000, pp. 392-400
11. D. Kania: *Heurystyczna metoda dekompozycji zespołu funkcji boolowskich wykorzystująca dekompozycje złożone, przeznaczona dla układów FPGA typu tablicowego*. Kwartalnik Elektroniki i Telekomunikacji, 2000, 46, z. 2, ss. 191-206
12. D. Kania: *Algorytmy podziału wyjść umożliwiające realizację układów cyfrowych w strukturach PLD*. Kwartalnik Elektroniki i Telekomunikacji, 1999, 45, z. 2, pp. 189-202
13. Y. Lai, K. R. Pan, M. Pedram: *OBDD-Based Function Decomposition: Algorithms and Implementation*. IEEE Transactions on Computer-Aided Design, 15(8), August 1996, pp. 977-990
14. Ch. Legl, B. Wurth, K. Eckl: *An Implicit Algorithm for Support Minimization during Functional Decomposition*. ED&TC, Paris, 1995, pp. 412-417
15. T. Łuba: *Rola i znaczenie syntezy logicznej w projektowaniu układów cyfrowych*. RUCC'99, 1999, ss. 5-84
16. R. Murgai, N. Shenoy, R. K. Brayton, A. Sangiovanni-Vincentelli: *Improved Logic Synthesis Algorithms for Table Look up Architectures*. ICCAD-91, Santa Clara, CA, November 1991, pp. 564-567

17. M. Nowicka, T. Luba, H. Selvaraj: *Multilevel Decomposition Strategies in Decomposition-Based Algorithms and Tools*. International Workshop on Logic and Architecture Synthesis, Grenoble, 1997, pp. 129-136
18. R. Murgai, Y. Nishizaki, N. Shenay, R. K. Brayton, A. Sangiovanni-Vincentelli: *Logic Synthesis for Programmable Gate Array*. Proc. 27th DAC, June 1990, pp. 620-625
19. M. Rawski, M. Nowicka, P. Tomaszewicz, T. Luba: *Decomposition-based synthesis and its application in FPGA-oriented technology mapping*. PDS'96, 1996, pp. 47-54
20. K. Sharma: *Programmable Logic Handbook, PLDs, CPLDs, & FPGAs*. McGraw-Hill, 1998
21. W. Wan, M. A. Perkowski: *A new Approach to the Decomposition of Incompletely Specified Multi-Output Functions Based on Graph Coloring and Local Transformations and Its Applications to FPGA Mapping*. Proc. of EDAC'92, 1992, pp. 230-235
22. www.cbl.ncsu.edu/ Collaborative Benchmarking Laboratory, Department of Computer Science at North Carolina State University

E. HRYNKIEWICZ, D. KANIA

A METHOD OF SYNTHESIS FOR LUT-BASED FPGAs

S u m m a r y

The main problem of synthesis for look-up table FPGAs is how to split the whole project into sub-circuits that can be realised in separate configurable logic blocks. The algorithm of optimal partitioning has not been discovered so far.

Since recourses of configurable logic blocks enable to implement any function yet with limited arguments only therefore the decomposition, arises as an essential stage of the synthesis. In this paper, two synthesis techniques based on decompositions have been compared. Curtis theory is the background of these methods.

The main difference of the proposed methods is the direction of the decomposition i.e. „from input to output” or „from output to input”. The impact of decomposition direction on synthesis effectiveness is showed. Developed algorithms, implemented within a Decomp system, have been used for partitioning benchmarks due to realisation by means of the Configurable Logic Blocks included in FPGAs.

Experimental results show clearly that the „input-to-output” decomposition method yields better results as far as block number and logic level number are concerned. Further development should be aimed at the improvements of this method only, using elements of non-disjoint decomposition for instance, as the other method yields worse results even though algorithm complexity and time of synthesis are comparable. The obtained results have been compared to the ones published in bibliography.

Keywords: logic synthesis, decomposition, partitioning, FPGA

Badania symulacyjne modulacji różnicowych ze zmienną częstotliwością próbkowania

RYSZARD GOLAŃSKI

*Katedra Elektroniki, Akademia Górniczo-Hutnicza w Krakowie
30-059 Kraków, al. Mickiewicza 30.
golanski@agh.edu.pl*

*Otrzymano 2003.07.26
Autoryzowano 2003.12.31*

Zaprezentowano wyniki symulacyjnych badań parametrów wyjściowych dwóch rodzajów modulacji różnicowych; z których NS-DM posiada możliwość kompowania jedynie częstotliwości próbkowania natomiast ANS-DM zapewnia również kompowanie kroku kwantyzacji. Badaniami objęto cztery podstawowe parametry wyjściowe: współczynnik SNR , (SNR_{max}), średnią częstotliwość próbkowania f_{avg} , współczynnik kompresji CF i zakres dynamiki DR . Maksymalizacja SNR , CF , DR i minimalizacja f_{avg} stanowiła podstawowe kryteria poszukiwania optymalnych zestawów parametrów wewnętrznych. Ustalono, że dla modulacji NS-DM i ANS-DM maksymalna wartość współczynnika SNR zależy od wartości średniej częstotliwości próbkowania f_{avg} (krok kwantyzacji decyduje jedynie o tym, przy jakim poziomie mocy to maksimum wystąpi). Wyniki badań wskazują na logarytmiczny charakter zależności $SNR_{max}[dB] = g(f_{avg})$. Wykazano symulacyjnie wyższą efektywność kodowania sygnałów niestacjonarnych przez modulatory różnicowe ze zmienną szybkością próbkowania niż przez modulatory ze stałą częstotliwością. Badania symulacyjne potwierdzają wyniki analiz matematycznych, że całkowity zakres dynamiki modulacji ANS-DM jest w [dB] równy sumie zysków kompowacji od każdego z adaptowanych parametrów. W efekcie modulacja ANS-DM wykazuje w porównaniu z modulacjami z kompowacją tylko jednego parametru (CFDM i NS-DM) nie tylko wzrost zakresu dynamiki DR , ale także współczynnika SNR_{max} . Wyznaczony na podstawie badań symulacyjnych współczynnik kompresji wynosi około 2 dla modulacji CFDM i NS-DM oraz około 3 dla modulacji ANS-DM.

Słowa kluczowe: adaptacyjne modulacje różnicowe, zmienna częstotliwość próbkowania, kompowacja, zakres dynamiki, badania symulacyjne, stosunek sygnał /szum

1. WSTĘP

Technika Digital Waveform Coding jest szeroko stosowana w inżynierii akustycznej ze względu na pracę w czasie rzeczywistym, duży zakres dynamiki przetwarzania, wysoką odporność na zakłócenia i dobrą dokładność aproksymacji.

Wyniki badań analitycznych, pozwalające na wyznaczanie parametrów modulacji różnicowych z adaptacją częstotliwości próbkowania zostały przedstawione w pierwszej części niniejszej pracy [2].

Wielu właściwości koderów delta nie da się opisać przejrzystym modelem analitycznym. Modele matematyczne z natury ich funkcji idealizują działanie urządzeń. Z powodu dużej liczby zewnętrznych i wewnętrznych parametrów modulatorów NS-DM i ANS-DM, badania analityczne wymagają użycia bardzo rozbudowanego modelu i sprowadzają się do iteracyjnego rozwiązywania, za pomocą metod komputerowych, skomplikowanych układów równań. Wobec wielu założeń upraszczających, które muszą być przyjmowane przy poszukiwaniu rozwiązań metoda analityczna nie daje w tym przypadku wyników dokładniejszych niż symulacje [3]. Metoda symulacyjna pozwala łatwo ustalić optymalne parametry wewnętrzne, a sposób i kolejność ich doboru najczęściej wynika z rozważań analitycznych [4, 5, 6]. Obie metody symulacyjna i analityczna uzupełniają się i wraz z badaniami prototypów pozwalają nie tylko oceniać właściwości modulatorów, ale także je projektować.

Przeprowadzone badania miały na celu wyznaczenie parametrów wyjściowych dwóch rodzajów modulacji różnicowych, z których NS-DM [2, 3] posiada możliwość kompowania tylko częstotliwości próbkowania natomiast ANS-DM [2, 4] zapewnia również kompowanie kroku kwantyzacji.

Oznaczenia parametrów wewnętrznych modulacji różnicowych

- $f_{smax}, (f_{smin}), (f_{start}), f_s$ — maksymalna (minimalna), (startowa) częstotliwość próbkowania
- $T_{smax}, (T_{smin}), (T_{start}), T_s$ — maksymalny (minimalny), (startowy) okres próbkowania
- $K1$ — krotność wzrostu częstotliwości próbkowania
- $K2$ — krotność redukcji częstotliwości próbkowania
- Δ — krok kwantyzacji
- $\Delta_{min}, (\Delta_{max})$ minimalny, (maksymalny) krok kwantyzacji
- P — krotność wzrostu kroku kwantyzacji
- Q — krotność redukcji kroku kwantyzacji
- f_{3dB} — częstotliwość 3 decybelowego spadku amplitudy sygnału wejściowego
- f_o — częstotliwość odcięcia sygnału wejściowego
- Δ_{opt} — wartość kroku kwantyzacji pozwalająca uzyskać $SNR = SNR_{max}$ przy zadanej f_s i dla zadanej poziomu mocy wejściowej S_{wej}
- K_f — stosunek maksymalnej do minimalnej częstotliwości próbkowania
- K_Δ — stosunek maksymalnego do minimalnego kroku kwantyzacji

• S_{wej}

Ozna

• SNR • SNR • $DR =$

który

nie m

• BR_{av}

(DM)

• f_{savg}

żany

• $CR =$ • $CF =$

nych

niezł

ADM

LDM

NS-DM

ANS-DM

CFDM

CVSD

ADM

ASR-AL

USZ

DA-DM

Bad

ze zmier

badane a

¹ Mod

alizu

param

- S_{wej} — średnia moc sygnału wejściowego

Oznaczenia najważniejszych parametrów wyjściowych modulacji różnicowych

- SNR — stosunek sygnał-szum
- SNR_{max} — maksymalna wartość SNR
- $DR = \frac{S_2}{S_1}$ — stosunek granicznych wartości mocy sygnału wejściowego, pomiędzy którymi współczynnik SNR modulatora ADM przyjmuje wartość maksymalną lub nie mniejszą niż zadana.
- BR_{avg} — średnia przepływność bitowa (definiowana dla modulacji NS-DM i ANS-DM)
- f_{savg} — średnia częstotliwość próbkowania, parametr odpowiadający BR_{avg} , wyrażony w jednostkach częstotliwości,
- CR — zysk kompresji,
- CF — współczynnik kompresji modulatora ADM: średnia liczba bitów niezbędnych do zakodowania sygnału przez modulator LDM odniesiona do liczby bitów niezbędnych do zakodowania tego sygnału z tą samą jakością przez modulator ADM.

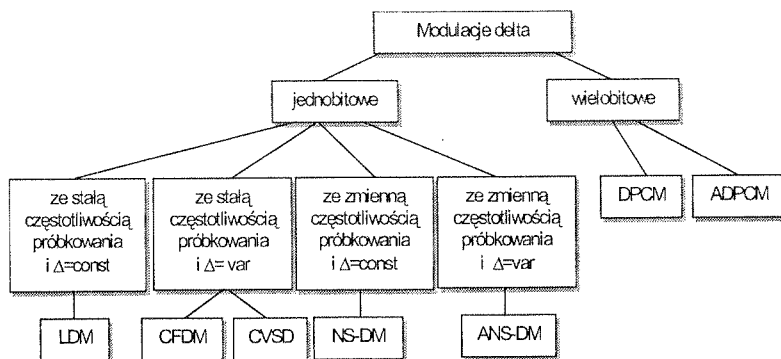
Nazewnictwo

LDM	— Linear Delta Modulation
NS-DM	— Non-uniform Sampling Delta Modulation
ANS-DM	— Adaptive Non-uniform Sampling Delta Modulation
CFDM	— Constant Factor Delta Modulation
CVSD	— Continuously Variable Slope Delta
ADM	— Adaptive Delta Modulation
ASR-ADM	— Adaptive Sampling Rate- ADM
USZ	— Ujemne Sprzężenie Zwrotne
DA-DM	— Doubly Adapted Delta Modulation

2. CEL BADAŃ EKSPERYMENTALNYCH

Badania miały za zadanie wyznaczenie tych parametrów modulatorów różnicowych ze zmienną częstotliwością próbkowania (rys. 1), które albo nie były do tej pory w ogóle badane albo są bardzo trudne do zbadania metodami analitycznymi¹

¹ Modele analityczne nie uwzględniają niektórych parametrów wewnętrznych modulatorów lub idealizują ich praktyczne działanie (w przypadku CFDM, ANS-DM i NS-DM dotyczy to np. takich parametrów jak f_{start} , $K1$, $K2$, P , Q).



Rys. 1. Ogólny podział współczesnych modulacji delta

Fig. 1. The classification of delta modulations on the ground of adapted parameters

Potrzeba przeprowadzenia wielu z prezentowanych dalej badań wynika także z konieczności weryfikacji rozważań analitycznych. Taka weryfikacja pozwoli uznać, czy zależności wyprowadzone z modeli matematycznych mogą być podstawą do obliczania parametrów wyjściowych modulatorów NS-DM i ANS-DM.

3. KONCEPCJA BADAŃ

Przygotowana koncepcja służy realizacji postawionych celów badań. Ustalono następujące etapy przeprowadzenia badań:

- A. Ustalenie najważniejszych parametrów wyjściowych (w wielu przypadkach jest równoznaczne z ustaleniem kryteriów, według których będą realizowane poszczególne pomiary. Wyselekcjonowano cztery podstawowe parametry wyjściowe powszechnie używane [7, 8, 9] do scharakteryzowania i porównywania właściwości modulatorów różnicowych ADM: SNR_{max} , DR , CF i f_{savg} lub BR_{avg} ².
- B. Badania wstępne. Wynikają z konieczności ustalenia relacji pomiędzy najważniejszymi parametrami wewnętrznymi badanych modulatorów.
- C. Poszukiwania zestawów parametrów wewnętrznych zapewniających uzyskanie ekstremum dla poszczególnych parametrów wyjściowych.
- D. Analiza wyników pomiarów, ilustracje graficzne, ustalenie zależności i wnioski końcowe.

² f_{savg} [kHz] i BR_{avg} [kb/s] w 1-bitowych modulacjach z nierównomiernym próbkowaniem są parametrami wyjściowymi (wynikowymi) i są sobie równe co do wartości bezwzględnej.

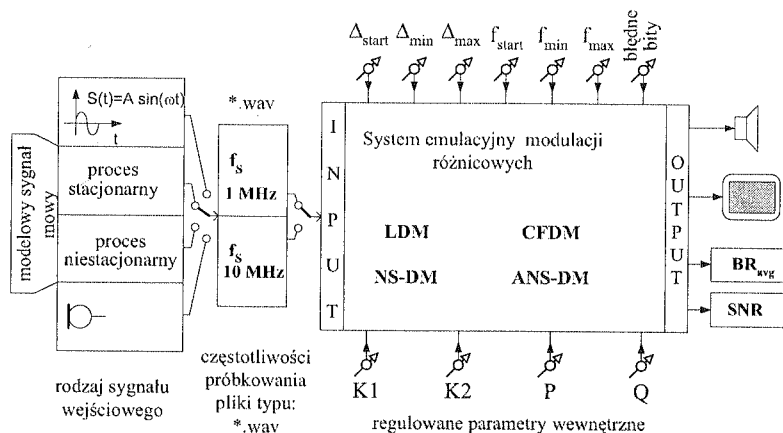
3.1. ZADANIA PROCESU BADAWCZEGO

Zadania badawcze zostały ustalone zgodnie z wytyczonymi celami badań na podstawie przygotowanej koncepcji i wykonanych badań analitycznych. Mają one pozwolić na ustalenie właściwości modulacji z nierównomiernym próbkowaniem i dokonanie ich oceny na tle właściwości modulacji różnicowych z równomiernym próbkowaniem. Podstawowe zadania eksperymentalnych badań modulacji NS-DM i ANS-DM to:

- Poszukiwanie optymalnych zestawów parametrów wewnętrznych ($K1$, $K2$, Δ , T_{start} oraz T_{smin} , T_{smax}). Optymalne zestawy parametrów to te, które pozwalają maksymalizować wartości wybranego parametru wyjściowego przy zadanych kryteriach np. maksymalizacja SNR przy zadanej wartości średniej przepływności bitowej BR_{avg} [8, 9, 10].
- ✓ opracowanie metody doboru współczynników adaptacji okresu próbkowania ($K1$, $K2$) i wartości częstotliwości startowej f_{start} (przy zadanych f_{smax} i f_{smin}).
- Zbadanie efektywności zmian zakresu dynamiki DR .
- Wyznaczenie współczynników kompresji CF .

3.2. RODZAJE POMIARÓW

Rys. 2 przedstawia schemat funkcjonalny systemu pomiarowego używanego podczas badań.



Rys. 2. System badawczy modulacji różnicowych

Fig. 2. Diagram of computing delta modulations performances in simulation

Pomiary, których zdecydowano się dokonać można podzielić na 3 grupy:

- wstępne
- indywidualne
- porównawcze

Pomiary wstępne. Przydatne do zapewnienia właściwych warunków pomiarów i umożliwiające poznanie zarówno podstawowych różnic jak i cech wspólnych modulacji ze stałą jak i zmienną częstotliwością próbkowania.

Pomiary indywidualne. Do tej grupy pomiarów należały, te których wyniki miały pozwolić na ustalenie, które parametry wewnętrzne mają najbardziej istotne znaczenie dla optymalnej realizacji procesu modulacji ze zmienną szybkością próbkowania. Pomiary indywidualne pozwoliły ocenić właściwości ściśle charakteryzujące dany typ modulacji delta.

Badania porównawcze. Wybrano następujące parametry wyjściowe w oparciu, o które dokonywano porównań właściwości wszystkich rozważanych rodzajów modulacji różnicowych: SNR , SNR_{max} , DR , CF oraz f_{savg} lub BR_{avg} .

3.3. KRYTERIA UZYSKIWANIA OPTYMALNYCH PARAMETRÓW WEWNĘTRZNYCH MODULATORÓW RÓŻNICOWYCH ZE ZMIENNĄ SZYBKością PRÓBKOWANIA

Wstępnie dokonywano optymalizacji parametrów wewnętrznych przy dwóch zasadniczych kryteriach:

- poszukiwanie największego możliwego do osiągnięcia współczynnika SNR , przy zadanej wartości średniej przepływności bitowej BR_{avg} i ustalonym zakresie dynamiki DR .
- poszukiwanie minimalnej wartości średniej przepływności bitowej BR_{avg} przy ustalonych a priori stosunku SNR_{max} i zakresie dynamiki DR .

Stwierdzono, że oba kryteria dają podobne wartości parametrów wewnętrznych. Ustalono, że w dalszych badaniach będzie stosowane kryterium stałości BR_{avg} . Pozostałe kryteria:

- zakres dynamiki DR . Poszukiwano takich zestawów parametrów wewnętrznych, które pozwalają na maksymalizację zakresu dynamiki przy zadanej wartości SNR i/lub BR_{avg} .
- współczynnik kompresji CF . Wartość CF pośrednio wynika z relacji $SNR_{max} = f(BR_{avg})$.

3.4. SPOSÓB OPTYMALIZACJI PARAMETRÓW WEWNĘTRZNYCH MODULATORÓW NS-DM I ANS-DM

Dobór optymalnych zestawów parametrów wewnętrznych modulatorów NS-DM i ANS-DM ze względu na ilość tych parametrów a także na liczbę kryteriów jest

procesem
na kilka
Skor
metodą C
rametrów
i po prze
Szc
zentowan
przyjmo
wstępnych

Bada
nych sym
zbudowan
różnicowy
przeznacz
Dalej
gramów s

4.1.

”Mo
plików d
metrów w
metrów z
wprowadz
wartości
bitowej B
wych w z
cych. Pro
wyjściu p
w linii tr
wy po de
*.wave, sy
nego. Pos
i demodu
2-kanalów

procesem żmudnym i skomplikowanym (ilość możliwych przekrojów można oszacować na kilkadziesiąt).

Skorzystano z jednej z najstarszych metod szukania ekstremum funkcji zwanej metodą Gaussa-Seidla lub metodą relaksacyjną [10, 11]. Kolejność optymalizacji parametrów wewnętrznych ustalono na podstawie wcześniejszych badań analitycznych i po przeprowadzeniu badań wstępnych.

Szczegóły tych badań były kilkakrotnie publikowane [4, 5, 6, 10, 12, 13]. W prezentowanych w dalszym ciągu niniejszego artykułu wynikach eksperymentów należy przyjmować, że parametry wewnętrzne zostały dobrane zgodnie z wynikami badań wstępnych prezentowanych a [4, 5, 6, 12, 13].

4. NARZĘDZIA BADAŃ EKSPERYMENTALNYCH

Badań eksperymentalnych dokonano przede wszystkim za pomocą algorytmicznych symulatorów modulacji różnicowych [14, 15, 16]. Celem przeprowadzenia badań, zbudowano kilka programów symulacyjnych dla 1-bitowych modulacji i demodulacji różnicowych. Programy posiadają różne możliwości funkcjonalne, bo różne jest ich przeznaczenie. Zostały one szczegółowo opisane a [14, 15, 16].

Dalej zostaną scharakteryzowane cechy funkcjonalne dwóch podstawowych programów symulacyjnych, które pozwoliły zrealizować większość założonych badań.

4.1. CHARAKTERYSTYKA PODSTAWOWYCH PROGRAMÓW SYMULACYJNYCH

4.1.1. Emulator „modulacje”

„Modulacje” — to program emulacyjny, umożliwiający kodowanie różnicowe plików dźwiękowych typu *.wave [17]. Emulator umożliwia regulację wszystkich parametrów wewnętrznych (np. f_s , Δ , P , Q , $K1$, $K2$, T_{smax} , T_{smin}) i najważniejszych parametrów zewnętrznych (S_{wej} , rodzaj sygnału wejściowego). W programie możliwe jest wprowadzanie błędnych bitów w torze transmisyjnym. Symulator dokonuje obliczeń wartości SNR a dla modulatorów NS-DM i ANS-DM również średniej przepływności bitowej BR_{avg} . Umożliwia automatyczne wykonywanie obliczeń parametrów wyjściowych w zadanym zakresie zmian poszczególnych parametrów wewnętrznych i sterujących. Program wizualizuje sygnały: wejściowe pliki *.wave, sygnał aproksymujący na wyjściu predyktora w nadajniku, ciąg binarny wyemitowany z nadajnika, ciąg binarny w linii transmisyjnej, ciąg binarny zdetektowany przez odbiornik i sygnał wyjściowy po demodulacji (odtworzony plik *.wave). Umożliwia odsłuch wejściowego pliku *.wave, sygnału aproksymującego na wyjściu predyktora w nadajniku i zdemodulowanego. Posiada możliwość synchronizacji przebiegów czasowych w oknie modulatora i demodulatora, ustalania podstawy czasu i większość regulacji jak w oscyloskopie 2-kanalowym.

4.1.2. Symulator „nsdm-sym”

„Nsdm-Sym” [18] — Nsdm-Sym jest zaawansowaną aplikacją symulatora algorytmicznego autonomiczną, jeśli chodzi o sygnał wejściowy. Realizuje funkcje symulatora oraz analizatora. W oknie graficznym wyświetlany jest, zależnie od funkcji, sygnał modulowany, jego aproksymacja lub wykres $SNR = f(S_{wej})$ będący wynikiem analizy. Funkcje symulatora Nsdm-Sym:

- o Generuje próbki kilku rodzajów sygnałów (sinusoida, prostokąt, trójkąt, szum gaussowski skalkowany), które mogą być modulowane zgodnie z algorytmem każdej z badanych modulacji. Dla każdego z sygnałów okresowych istnieje możliwość zadania amplitudy, fazy początkowej oraz częstotliwości. W przypadku szumu gaussowskiego skalkowanego można zadawać poziom mocy sygnału wejściowego a [dB],

- o posiada możliwość ustawienia wszystkich parametrów wewnętrznych dla każdej z modulacji, szczegółowo wizualizuje przebieg modulowany, jego aproksymację i wyjściowy sygnał cyfrowy lub sygnał błędu,

- o wyznacza stosunek SNR (równocześnie dla kilku modulatorów) w zadanym zakresie poziomów mocy sygnału wejściowego i graficznie wizualizuje uzyskane wyniki.

Program Nsdm-Sym jest szczególnie przydatny w tych badaniach symulacyjnych, w których konieczne jest śledzenie pracy algorytmu adaptacji lub szybka weryfikacja takich parametrów wyjściowych jak współczynnik SNR oraz średnia przepływność bitowa sygnału BR_{avg} .

4.2. RODZAJE BADANYCH SYGNAŁÓW

Podczas badań symulacyjnych używano następujących rodzajów sygnałów (w postaci plików typu *.wave):

- Próbkę sygnału mowy — o regulowanej do 20 [kHz] częstotliwości odcięcia f_o i czasie trwania do 5 [s],
- Sinusoida 800[Hz] — czas trwania około 1 [s].
- „analityczny” stacjonarny sygnał mowy — szumu biały skalkowany. Przyjęto $f_{3dB}/f_o = 0,23$ i $f_o = 3,4$ kHz [1]. Charakteryzuje się gaussowskim lub wykładniczym rozkładem gęstości prawdopodobieństwa amplitud. Ten stacjonarny sygnał aproksymujący mowę (bez okresów ciszy) jest w badaniach analitycznych przyjmowany najczęściej jako wejściowy przy wyznaczaniu szumów kwantyzacji [1, 7, 19]. Czas trwania około 1 [s].
- Sygnał niestacjonarny (syntetyczny sygnał mowy) — w pracy przyjęto tak nazwać sygnał składający się z kilku do kilkunastu fragmentów sygnału stacjonarnego każdy o długości nie mniejszej niż 200 ms [26]. Poszczególne fragmenty różnią się wartością skuteczną. Źródło to odzwierciedla sygnał niestacjonarny, w którym mamy do czynienia ze zmiennym poziomem mocy dźwięku oraz z chwilami „ciszy”. Czasem nazywany jest syntetycznym sygnałem mowy.

Jako że
Edit Pro” [2]
liwia kształt
przeznaczony
mocy i usta

5. PRZ

Prezent
i ANS-DM

Badani
21, 22] poz
wewnętrznych

między f
Badani
na podstaw

- a) maksym
- b) minimal
- z maksy
- c) maksym
- d) zbadanie

Proble
kością prób
szego artyk

W dalsz
cji poziomu
trach wewn
przedstawio

Zostana
nić działani

W celu
pomiarów z
sygnału wej

Charak
pracy modu
wą ujemną

Jako źródło plików dźwiękowych typu *.wave użyto edytora dźwiękowego „Cool Edit Pro” [20]. Edytor ten zapewnia realizację wyżej wymienionych sygnałów. Umożliwia kształtowanie sygnału np. konwersję częstotliwości próbkowania pliku *.wave przeznaczonego do modulacji, przeprowadzenie analizy widmowej, analizy poziomu mocy i ustalenie wielu innych właściwości sygnału zdemodulowanego.

5. PRZEGLĄD NAJWAŻNIEJSZYCH CHARAKTERYSTYK MODULACJI NS-DM I ANS-DM

Prezentowany w niniejszym rozdziale przegląd charakterystyk modulacji NS-DM i ANS-DM jest efektem realizacji zaplanowanych zadań badawczych i pomiarowych.

Badania symulacyjne rozpoczęto od tzw. badań wstępnych [4, 5, 6, 10, 12, 13, 21, 22] pozwalających przede wszystkim na ustalenie zakresu zmian parametrów wewnętrznych oraz relacji wzajemnej pomiędzy tymi parametrami (m.in. ustalenie relacji pomiędzy f_{start} a BR_{avg}).

Badania symulacyjne zgodnie z przyjętą koncepcją [rozdz.3] zostały podzielone na podstawie czterech podstawowych kryteriów:

- a) maksymalizacja stosunku SNR ,
- b) minimalizacja średniej przepływności bitowej BR_{avg} (jest najczęściej równoważne z maksymalizacją współczynnika kompresji CF),
- c) maksymalizacji zakresu dynamiki DR ,
- d) zbadanie odporności na załócenia.

Problematyka dotycząca odporności na zakłócenia modulatorów ze zmienną szybkością próbkowania [12, 13] zostanie zaprezentowane a całości a drugiej części niniejszego artykułu.

W dalszej części pracy przedstawiono charakterystyki SNR , SNR_{max} i BR_{avg} w funkcji poziomu mocy lub częstotliwości próbkowania przy zadanych optymalnych parametrach wewnętrznych. Charakterystyki te uzyskano na podstawie pomiarów wg schematu przedstawionego na rys. 2.

5.1. BADANIA WSTĘPNE

Zostaną zaprezentowane wyniki wybranych badań wstępnych, które pozwoliły ocenić działanie modulatorów ze zmienną częstotliwością próbkowania.

W celu zobrazowania mechanizmu funkcjonowania modulatora NS-DM dokonano pomiarów zależności $SNR = g(f_s)$ modulatora LDM dla różnych poziomów mocy sygnału wejściowego, ale przy stałym kroku kwantyzacji.

Charakterystyki na rys. 3 mogą w sposób poglądowy posłużyć do zobrazowania pracy modulatora NS-DM ($f_s = \text{var}$, $\Delta = \text{const}$, S_{wej} - parametr) jako układu z cyfrową ujemną pętlą SZ. Przy ustalonym kroku kwantyzacji, po wzroście poziomu mocy

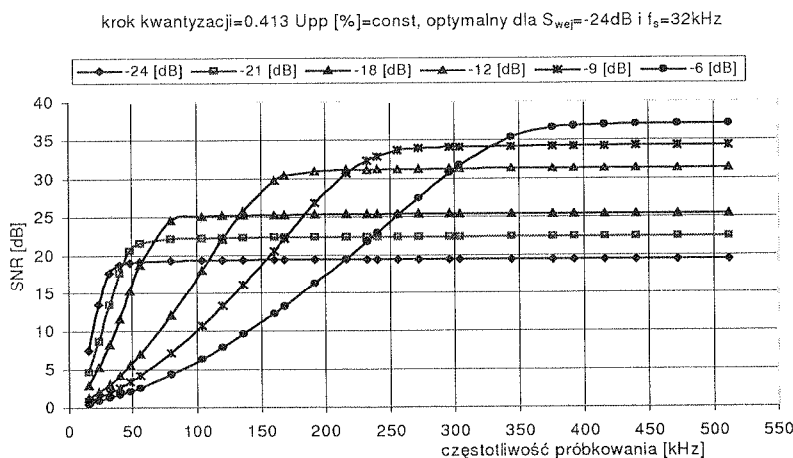
dzięki działaniu pętli USZ częstotliwość próbkowania wzrasta dotąd aż zostaną wyeliminowane szumy przeciążenia stromości (obszar nasycenia krzywych na rys. 3). Kiedy pozostaną już tylko szumy granulacji ustaje proces zwiększania częstotliwości bowiem szumy te nie zależą od częstotliwości a wyłącznie od wartości kroku kwantyzacji a ten w modulacji NS-DM pozostaje stały.

W praktyce szumy przeciążenia stromości nie są całkowicie eliminowane a jedynie minimalizowane ze względu na skończoną wartość kroku kwantyzacji i ograniczoną czułość cyfrowej pętli USZ. Podczas modulacji typu NS-DM (ANS-DM) każda wartość nachylenia sygnału wejściowego mniejsza od maksymalnej powoduje, że pętla USZ zmniejsza częstotliwość próbkowania utrzymując przy tym minimalne szumy przeciążenia stromości, czyli nie pogarsza jakości. Dlatego średnia przepływność bitowa modulacji NS-DM (ANS-DM), przy której uzyskuje się tą samą wartość SNR ulega zmniejszeniu (kompresja danych) w stosunku do modulacji LDM (rys. 6, 7).

Problemem optymalizacyjnym jest ustalenie „miejsca na kolanie” (rys. 3) poprzez wybór wartości kroku kwantyzacji. Na „części płaskiej” kosztem niewielkiego spadku SNR (< 1 dB) można uzyskać zauważalną poprawę współczynnika kompresji. Na części stromej można zwiększyć SNR (< 1 dB) kosztem niewielkiego pogorszenia współczynnika kompresji. Wybór „miejsca na kolanie” zależy, więc od konkretnej aplikacji.

Rys. 3 wskazuje ponadto, że każde czterokrotne (o 6 dB) zwiększenie poziomu mocy wejściowej wymaga, co najmniej podwojenia częstotliwości próbkowania dla zminimalizowania szumów przeciążenia stromości.

Przy doborze parametrów jak w przypadku przedstawionym na rys. 3 „dla naj-



Rys. 3. Współczynnik SNR w funkcji częstotliwości próbkowania. Krok kwantyzacji stały, optymalny dla $S_{wej} = -24 \text{ dB}$ i $f_{smin} = 32 \text{ kHz}$

Fig. 3. SNR 's versus sampling frequency. Family of curves for different input signal levels. Step size fixed, optimal for $f_{smin} = 32 \text{ kHz}$ and input level equals -24 dB

mniejsz
kość p
pozio
często
zwięks
szumó

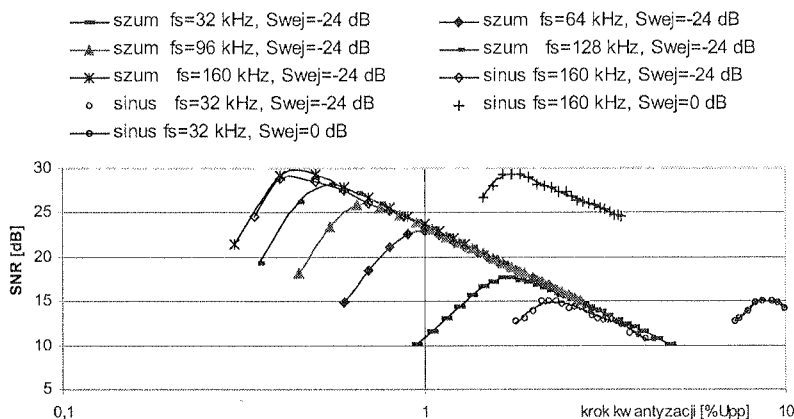
Rys. 4.

Fig. 4.
paramet

Po
ściwoś
√ Prz
o 6
√ op
zał
koś
to
√ As
kro
źró
Ro
wego j

³ Pod
pró
co p

mnijszego poziomu mocy” i „najmniejszej częstotliwości próbkowania” założona jakość przetwarzania (SNR) zostaje zapewniona począwszy od tego poziomu mocy do poziomu największego. Maksymalny poziom ustalany jest wartością górnej granicznej częstotliwości próbkowania $f_{s\max}$. Powyżej, tego poziomu mocy pętla USZ nie może zwiększyć częstotliwości i dalej układ zachowuje się jak modulator LDM w obszarze szumów przeciążenia stromości.



Rys. 4. Rodzina charakterystyk $SNR = f(\Delta)$ modulacji LDM dla różnych częstotliwości próbkowania. Parametry wewnętrzne dobrano oddzielnie dla dwóch poziomów mocy i dwóch rodzajów sygnału wejściowego

Fig. 4. SNR 's of LDM against step size. Family of curves for different sampling frequencies. Inside parameters have been selected for two's kinds of input signals and two's input signal levels, separately

Pomiary, których wynikiem jest rys. 4 pozwoliły ustalić kilka zasadniczych właściwości modulacji adaptacyjnych delta:

- ✓ Przy stałym poziomie mocy każde podwojenie częstotliwości zwiększa $SNR_{\max}$ o 6 dB^3 a krok optymalny maleje zgodnie z zależnościami [1, 7, 19, 24],
- ✓ optymalny, ze względu na maksymalizację współczynnika SNR krok kwantyzacji zależy od poziomu mocy dla którego tą maksymalizację się przeprowadza, szybkości próbkowania i rodzaju kodowanego źródła. Przy danym poziomie mocy jest to krok, przy którym szumy przeciążenia stromości zostają zminimalizowane.
- ✓ Asymptoty nie zmieniają swego nachylenia przy zmianie szybkości próbkowania, kroku kwantyzacji, poziomu mocy. Niewiele różnią się dla wszystkich badanych źródeł sygnału wejściowego.

Rodzaj źródła wpływa na maksymalną wartość SNR , przy czym dla sygnału tonowego jest ona mniejsza niż dla szumu scałkowanego, ale różnice maleją ze wzrostem

³ Podczas badań nie stosowano filtru na wyjściu dekodera. Przyrost $SNR_{\max}$ w funkcji częstotliwości próbkowania jest tylko w przybliżeniu stały i zmniejsza się ze wzrostem częstotliwości próbkowania, co potwierdza wyniki badań analitycznych [19].

szybkości próbkowania. Modulacje różnicowe mają mniejszą efektywność kodowania dla sygnałów o mniejszym współczynniku korelacji i jest to widoczne szczególnie przy małych częstotliwościach próbkowania [19].

5.1.1. KONKLUZJE WYNIKAJĄCE Z BADAŃ WSTĘPNYCH

Poniżej zostaną przedstawione najważniejsze konkluzje wynikające z przeprowadzonych badań wstępnych dotyczących ustalenia wzajemnych zależności pomiędzy parametrami modulatorów z nierównomiernym próbkowaniem [4, 6, 10, 21].

- Maksymalna wartość współczynnika SNR zależy wyłącznie od średniej częstotliwości próbkowania f_{avg} (krok kwantyzacji decyduje jedynie o tym, przy jakim poziomie mocy to maksimum wystąpi).
- Wykazano [10, 12, 21], jedno–jednoznaczny charakter zależności pomiędzy f_{avg} i SNR . Pozwoliło to na ograniczenie liczby kryteriów stosowanych podczas badań eksperymentalnych.
- Należy stosować duże częstotliwości próbkowania plików *.wave używanych jako sygnały wejściowe symulatorów modulacji adaptacyjnych delta. Zapewnia to warunki pomiarów zbliżone do tych, które przeprowadza się w rozwiązaniach sprzętowych z sygnałem analogowym jako wejściowym. Szczegółowe badania z uwzględnieniem zakresów zmian częstotliwości próbkowania i kroku kwantyzacji modulatorów NS-DM i ANS-DM wykazały, że minimalna wartość częstotliwości pliku *.wave powinna wynosić 1 MHz.

5.2. BADANIA WPLYWU RODZAJU SYGNAŁU WEJŚCIOWEGO NA PARAMETRY MODULATORÓW ZE ZMIENNĄ CZĘSTOTLIWOŚCIĄ PRÓBKOWANIA

Prezentowane w kilku pracach [6, 8, 9, 19, 23, 25] wyniki badań parametrów wyjściowych modulacji ze zmienną szybkością próbkowania wskazują, że rodzaj kodowanego sygnału wejściowego zasadniczo nie zmienia kształtu charakterystyk $SNR = G(f_{avg})$ natomiast ma wpływ na wartości bezwzględne tych parametrów.

Na rys. 5 zaprezentowano rodzinę charakterystyk SNR w funkcji poziomu mocy wejściowej dla dwóch rodzajów modulacji i dwóch rodzajów sygnału wejściowego.

Pomiary zobrazowane na rys. 5 wskazują, że przy przyjętych [6, 10, 18, 21, 22] parametrach wewnętrznych badanych modulatorów:

- charakterystyki $SNR = g(f_s)$ i $SNR = h(f_{avg})$ zachowują zbliżony kształt niezależnie od rodzaju przetwarzanego sygnału i rodzaju modulacji delta. Uzasadnieniem zbliżonych wartości nachylenia asymptot w modulacjach z nierównomiernym i równomiernym próbkowaniem jest fakt, że dla poziomów mocy, przy których asymptoty te określają wartość SNR mechanizm kompandacji częstotliwości próbkowania występuje rzadko (zarówno częstotliwość próbkowania jak i krok kwantyzacji przyjmują wartości graniczne). W efekcie modulator NS-DM (ANS-DM)

Rys. 5. W

Fig. 5. SN

zachow
więc z
• przy s
ANS-I
• Zysk k
spełnio
• Zmian
kuje z
kompa

6. PO

6.1. BAD

Ważny
wewnętrzne n
zmienną c
o wartości
średnia prz
ści pomięc

kodowania
gólnie przy

przeprowa-
między pa-

ej częstotli-
przy jakim

między f_{savg}
czas badań

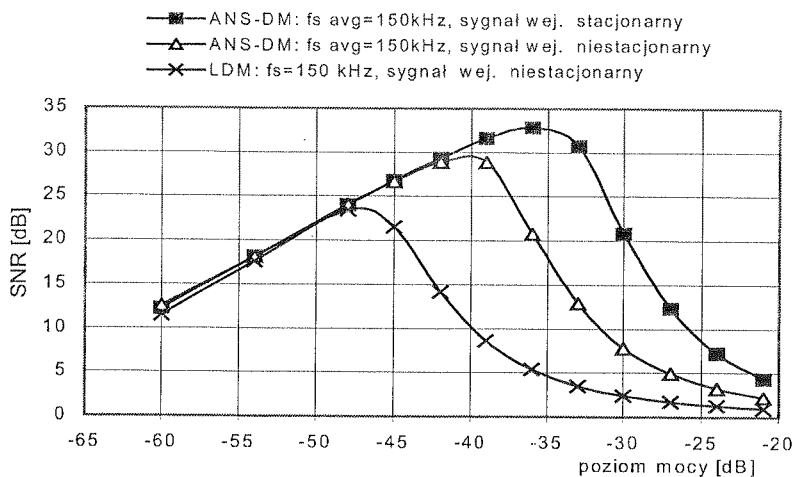
anych jako
nia to wa-
ch sprzeto-
z uwzględ-
acji modu-
wości pliku

ETRY

parametrów
rodzaj ko-
tyk $SNR =$

omu mocy
iowego.
[8, 21, 22]

stałt nieza-
Uzasadnie-
niomniom
zy których
wości prób-
rok kwan-
ANS-DM)



Rys. 5. Współczynnik SNR w funkcji poziomu mocy sygnału wejściowego dla modulacji ANS-DM i LDM. ANS-DM dla sygnału stacjonarnego i niestacjonarnego

Fig. 5. SNR 's of LDM, ANS-DM versus input signal level. Family of ANS-DM curves for stationary and non-stationary signal

zachowuje się podobnie do modulatora LDM ($f_s = \text{const}$, $\Delta = \text{const}$). Obowiązują, więc zależności opisane analitycznie dla tej modulacji [1, 7, 19, 24].

- przy spełnionym warunku $f_{sLDM} = f_{savg\text{ANS-DM}}$ współczynnik SNR_{max} modulacji ANS-DM przewyższa o około 9 dB wartość uzyskiwaną w modulacji LDM
- Zysk komandacji modulacji ANS-DM wynosi około 15 dB dla $SNR = 20$ dB przy spełnionym warunku $f_{sLDM} = f_{savg\text{ANS-DM}}$.
- Zmiana rodzaju przetwarzanego sygnału ze stacjonarnego na niestacjonarny skutkuje zmniejszeniem maksymalnej wartości stosunku SNR i zmniejszeniem zysku komandacji. Wielkość zmniejszenia zależy od stopnia niestacjonarności.

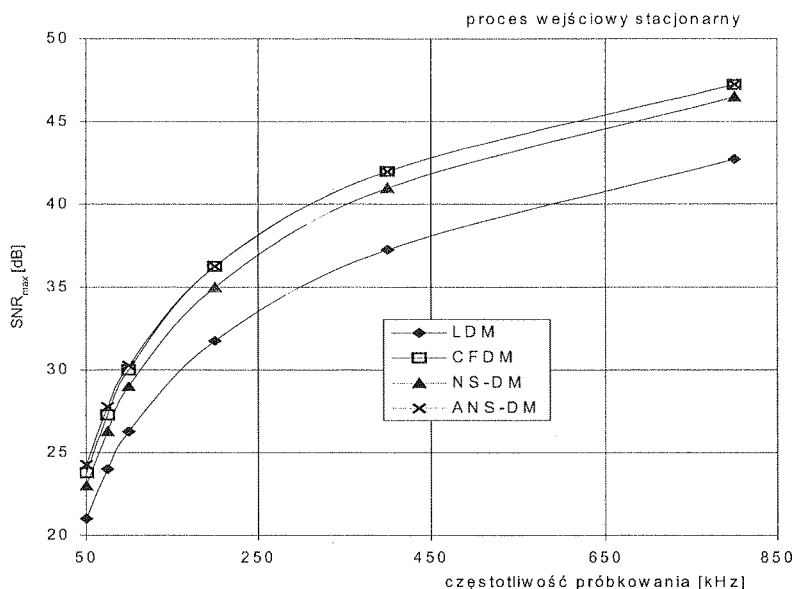
6. PORÓWNANIE PARAMETRÓW MODULATORÓW ADM ZE STAŁĄ I ZMIENNĄ CZĘSTOTLIWOŚCIĄ PRÓBKOWANIA

6.1. BADANIE WPLYWU ZMIENNEJ CZĘSTOTLIWOŚCI PRÓBKOWANIA NA PARAMETRY MODULACJI

Ważnym zagadnieniem badań symulacyjnych było ustalenie, które parametry wewnętrzne mają decydujący wpływ na maksymalną wartość SNR w modulacjach ze zmienną częstotliwością próbkowania. Wykazano, [4, 5, 10, 12] że przy $f_o = \text{const}$ o wartości SNR_{max} decyduje średnia częstotliwość próbkowania f_{savg} a licząc w [kb/s] średnia przepływność bitowa BR_{avg} . Dalsze badania miały ustalić charakter zależności pomiędzy SNR_{max} i średnią przepływnością bitową BR_{avg} a także porównanie tej

zależności z odpowiadającymi jej dla modulacji z równomiernym próbkowaniem. Charakterystyki przedstawione na rys. 6 i 7 pozwalają wyciągnąć kilka ważnych wniosków dotyczących zależności stosunku SNR_{max} i średniej częstotliwości próbkowania f_{savg} (modulacje NS-DM i ANS-DM).

Rys. 6. wskazuje na logarytmiczny charakter zmian SNR_{max} [dB] w funkcji BR_{avg} dla wszystkich badanych rodzajów modulacji delta. Nachylenie poszczególnych krzywych jest zbliżone a różnice dotyczą jedynie wartości SNR_{max} . Ponadto można zauważyć, że przy przetwarzaniu sygnałów stacjonarnych zastosowanie adaptacji częstotliwości próbkowania w modulacji ADM nie zwiększa SNR_{max} . Znajduje to potwierdzenie w badaniach analitycznych [1, 7].



Rys. 6. Rodzina charakterystyk SNR_{max} w funkcji częstotliwości próbkowania. Poziom mocy stacjonarnego sygnału wejściowego równy -6 dB

Fig. 6. Comparison of LDM, CFDM, NS-DM and ANS-DM optimum performance (SNR_{max}) with stationary signal. Input signal level equal -6 dB

Opierając się na powyższych spostrzeżeniach i wynikach dotychczasowych badań modulacji delta z równomiernym próbkowaniem [1, 4, 6, 7, 9, 15, 16, 19, 25], można znaleźć eksperymentalną zależność opisującą zmiany SNR_{max} w funkcji średniej przepływności bitowej dla modulatorów z nierównomiernym próbkowaniem.

Badania modulatorów delta dla sygnałów niestacjonarnych zilustrowane na rys. 7 pozwalają stwierdzić, że charakter zmian SNR_{max} [dB] pozostaje zbliżony do logarytmicznego. Następuje znaczne zmniejszenie współczynnika SNR_{max} dla modulacji bez adaptacji np. LDM. Zmniejszenie współczynnika SNR_{max} obserwuje się najczęściej także dla modulacji delta z adaptacją, a wielkość redukcji zależy od stopnia niesta-

Rys. 7. R

Fig. 7. C

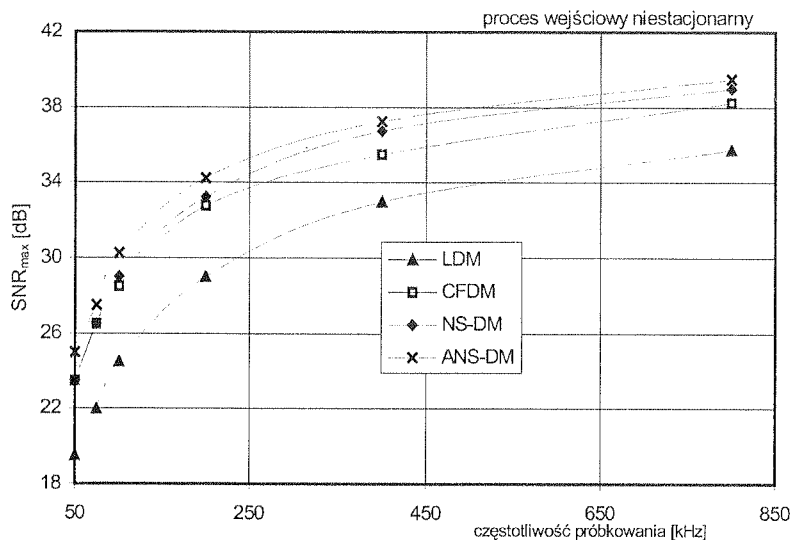
cjonarnoś
modulacji
ADM z p
Przep
teoretyczn

6.2. I

Zgod
efektywno
te same w
modulacji
rys. 8.

Na p
dowania r
uzyskiwan
wielkiego
sygnału w
CFDM je
towej przy

niem. Cha-
wniosków
wania f_{savg}
kcji BR_{avg}
nych krzy-
na zauwa-
częstotliwo-
wierdzenie



Rys. 7. Rodzina charakterystyk SNR_{max} w funkcji częstotliwości próbkowania. Poziom mocy sygnału wej. równy -6 dB

Fig. 7. Comparison of LDM, CFDM, NS-DM and ANS-DM optimum performance (SNR_{max}) with non-stationary signal. Input signal level equal -6 dB

cjonarności oraz rodzaju procesu wejściowego. Jednak redukcja wartości SNR_{max} dla modulacji z kompandacją częstotliwości próbkowania jest mniejsza niż dla modulacji ADM z próbkowaniem równomiernym.

Przeprowadzone pomiary symulacyjne (rys. 7) potwierdzają, więc przewidywania teoretyczne [25, 26].

6.2. PORÓWNANIE EFEKTYWNOŚCI KODOWANIA MODULACJI DELTA ZE STAŁĄ I ZMIENNĄ SZYBKościĄ PRÓBKOWANIA

a mocy

R_{max}) with

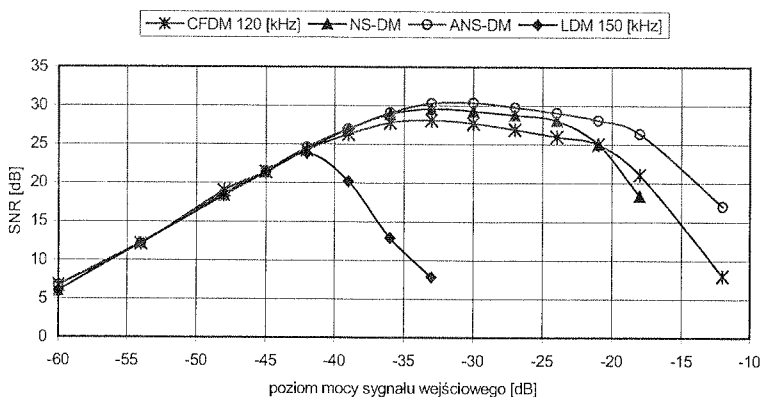
ych badań
[25], moż-
cji średniej
n.

e na rys. 7
do logaryt-
modulacji bez
najczęściej
nia niesta-

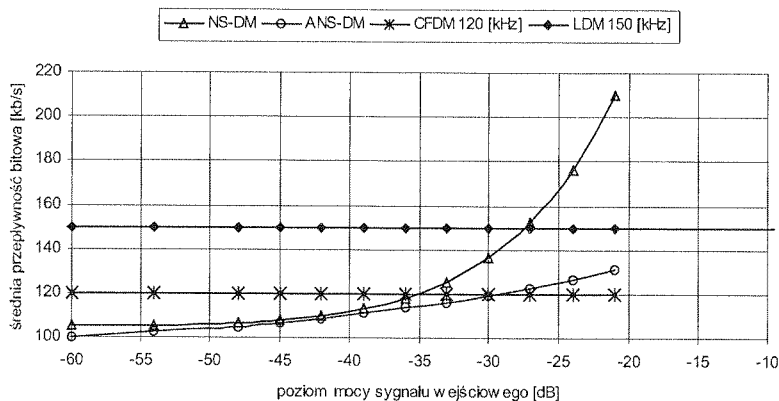
Zgodnie z założonym planem badań dokonano pomiarów pozwalających porównać efektywność modulacji ze stałą i zmienną szybkością próbkowania. W tym celu przyjęto te same wartości współczynników K_f , K_Δ i dobrano pozostałe parametry wewnętrzne modulacji tak aby uzyskać wymagany współczynnik SNR . Rezultaty przedstawiono na rys. 8.

Na podstawie rys. 8a i 8b można stwierdzić znacznie wyższą efektywność kodowania metody ANS-DM niż pozostałych badanych modulacji delta. Widać większe uzyskiwane wartości SNR i szerszy zakres dynamiki osiągany kosztem jedynie niewielkiego wzrostu średniej przepływności bitowej przy największych poziomach mocy sygnału wejściowego. Modulacja NS-DM wprawdzie przewyższa jakością modulację CFDM jednak odbywa się to kosztem wyraźnego wzrostu średniej przepływności bitowej przy największych poziomach mocy sygnału wejściowego.

a)



b)



Rys. 8. Współczynnik SNR (a) oraz przepływność bitowa BR_{avg} (b) a funkcji poziomu mocy. Sygnał wejściowy niestacjonarny

Fig. 8. SNR 's (a) and bit rate BR_{avg} (b) of DM modulations against input signal level. Non-stationary input signal

6.3. KOMPRESYJNE WŁAŚCIWOŚCI METOD ZE ZMIENNĄ SZYBKościĄ PRÓBKOWANIA

Badania właściwości kompresyjnych są podstawowymi wykonywanymi m.in. w celu dokonania oceny porównawczej różnych metod kodowania. Również po pojawieniu się pierwszych modulatorów ze zmienną szybkością kodowania opublikowano wiele wyników dotyczących tego parametru [8, 9, 19, 25, 27]. Wartość współczynnika kompresji CF wynika z zależności $BR_{avg} = f(S_{wej})|_{SNR=const}$. Zależność tą w sposób pośredni reprezentuje rys 8. Należy również podkreślić, że zmiana zestawu parametrów wewnętrznych (np. kroku kwantyzacji) pozwala a pewnych granicach zwiększyć współczynnik kompresji kosztem niewielkiego zmniejszenia współczynnika SNR (rys. 3).

Bada
towa prz
zapisany
własność

6.4. F

Wyk
czterech
dacji. Dw
(-42 dB
zapewnia

Rys. 9. SNR
wewn
Fig. 9

Tabela 1

Współczynniki kompresji CF wybranych modulacji ADM przy $SNR_{max} = \text{const}$.
 Sygnał wejściowy niestacjonarny

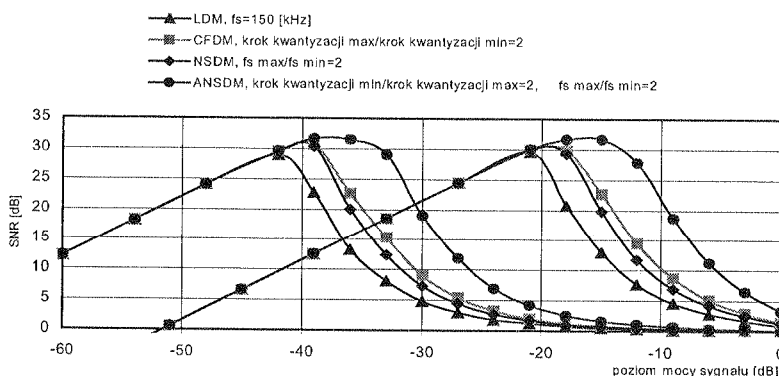
Compression factors CF of the selected ADM modulations. The same SNR 's.
 Non-stationary input signal

Rodzaj modulacji	CFDM	NS-DM	ANS-DM
f_{sLDM} w odniesieniu do f_{sCFDM} , $f_{savgNS-DM}$ i $f_{savgANS-DM}$	~ 2	~ 2	~ 3

Badania symulacyjne efektywności bitowej wyrażanej średnią przepływnością bitową przeprowadzono na syntetycznym sygnale mowy, ale także na próbkach mowy zapisanych w postaci plików *.wave. W tabeli 2 zebrano uogólnione wyniki dotyczące własności kompresyjnych poszczególnych modulacji.

6.4. PORÓWNANIE ZYSKU KOMPANDACJI MODULATORÓW ZE ZMIENNĄ I STAŁĄ CZĘSTOTLIWOŚCIĄ PRÓBKOWANIA

Wykonano pomiary, które pozwoliły na ilustrację zależności $SNR = f(S_{wej})$ dla czterech rodzajów modulatorów objętych porównaniem i oszacowanie zysków kompan-dacji. Dwie rodziny charakterystyk przedstawione na rys. 9 różnią się poziomem mocy (−42 dB i −21 dB) dla którego optymalizowano zestaw parametrów wewnętrznych zapewniający $SNR = SNR_{max}$.



Rys. 9. SNR w funkcji poziomu mocy. Dwie rodziny charakterystyk uzyskano optymalizując parametry wewnętrzne każdej modulacji raz dla poziomu mocy −42 dB a drugi dla poziomu −21 dB

Fig. 9. SNR 's of LDM, CFDM, NS-DM and ANS-DM versus, stationary input signal level.

Inside parameters optimized for −42 and −21 dB input level, separately

Analizując rys. 9 można wyciągnąć następujące wnioski:

- Modulacja ANS-DM, posiadająca możliwość komandowania dwóch parametrów wykazuje wzrost zakresu dynamiki w porównaniu do modulacji z komandacją jednego parametru. Zysk komandacji modulacji ANS-DM jest w [dB] równy sumie zysków komandacji od zakresu zmian każdego z adaptowanych parametrów (NS-DM i CFDM). Zysk komandacji modulacji NS-DM jest mniejszy niż zysk modulacji CFDM przy tym samym zakresie zmian parametru adaptowanego ($K_f = K_\Delta$). Te wyniki symulacyjnych badań własności komandacyjnych stanowią potwierdzenie wniosków wpływających z przeprowadzonych przez autora analiz matematycznych [23, 28].
- Potwierdzają się wyniki badań analitycznych [1, 8, 19, 25] wskazujące niezależność wartości SNR_{max} od poziomu mocy, dla którego optymalizuje się parametry wewnętrzne modulacji delta, jeśli tylko utrzymuje się tą samą wartość częstotliwości próbkowania f_s lub f_{avg} .

7. ANALIZA WYNIKÓW BADAŃ

Wszelkie badania symulacyjne przeprowadzono dla dwóch rodzajów modulacji różnicowych (NS-DM i ANS-DM), w których częstotliwość próbkowania podlega zmianom zgodnie z zastosowanym algorytmem. Równocześnie a celach głównie porównawczych dokonano tych samych badań dla dwóch najbardziej znanych rodzajów modulacji z równomiernym próbkowaniem: LDM i CFDM (adaptacja wg algorytmu Jayant'a).

Badaniami objęto cztery podstawowe parametry wyjściowe SNR , (SNR_{max}), BR_{avg} , (f_{avg}), CF , DR . Większość badań symulacyjnych dotyczyła parametrów nie objętych modelem matematycznym i ich wyniki mają znaczenie poznawcze. Pozwalają odzwierciedlić różnice między parametrami wyjściowymi modelu (modulatora idealnego) a parametrami wyjściowymi modulatora rzeczywistego. Poniżej zostaną wymienione najważniejsze wnioski wpływające z badań symulacyjnych.

- Modulatory ze zmienną częstotliwością próbkowania znacznie lepiej radzą sobie z przetwarzaniem sygnałów niestacjonarnych niż modulatory z próbkowaniem równomiernym. Dla badanych sygnałów pomiary symulacyjne wykazują różnicę SNR_{max} wynoszącą od 2 do 9 dB na korzyść modulatorów, w których stosuje się nierównomierne próbkowanie. Kompandacja częstotliwości próbkowania nie przynosi korzyści względem komandacji kroku kwantyzacji przy przetwarzaniu źródeł stacjonarnych, co przewidują wyniki badań analitycznych [25, 26].
- Kodowanie sygnałów niestacjonarnych w porównaniu do stacjonarnych nie wpływa na kształt charakterystyk SNR a jedynie na zmianę wartości.
- Zysk komandacji modulacji delta z podwójną adaptacją (np. ANS-DM) jest w mierze decybelowej sumą zysków komandacji od zakresu zmian kroku kwantyzacji i szybkości próbkowania [rys. 9]. Zależność ta jest bardzo korzystna pozwala, bo-

wiem
paran
• Wsp
około
wych
sji w
27].
• W m
tości
wejś
6, 7)
(śred
(wyn
wani
sygna
i war
modu
miej
stały
na po
ność
z ada
• Dla r
gnału
tości
tliwo
czynn
wość
dem
i wsp
same
• Bada
prakt
z obl
system
podej
nych,
zawsz
kich
wani
szcze
efeko

wiem poszerzać zakres dynamiki bez znacznej zmiany każdego z kompowanych parametrów.

- Współczynnik kompresji CF wynosi około 2 dla modulacji CFDM i NS-DM oraz około 3 dla modulacji ANS-DM. Są to uśrednione wyniki uzyskane dla modelowych sygnałów mowy. Przeprowadzone badania dotyczące współczynnika kompresji wykazują, że uzyskiwane wartości silnie zależą od rodzaju widma sygnału [19, 27].
- W modulacjach delta ze zmienną częstotliwością próbkowania o maksymalnej wartości SNR nie decyduje ani krok kwantyzacji (rys.4) ani poziom mocy sygnału wejściowego (rys. 8a, 9) ani f_{smax} , f_{smin} czy f_{start} [10, 12]. Badania (rys. 3, 4, 6, 7) wykazały natomiast, że decyduje o niej średnia częstotliwość próbkowania (średnia przepływność bitowa), która w tej modulacji jest parametrem wyjściowym (wynikowym). Dokładniej decyduje nadmiarowość średniej częstotliwości próbkowania w stosunku do maksymalnej częstotliwości f_o sygnału i rodzaj kodowanego sygnału. Badania wykazały, że zarówno warunki uzyskania maksimum SNR jak i wartość tego maksimum mogą być wyrażone tymi samymi funkcjami, co dla modulacji z próbkowaniem równomiernym [1,7,19, 24]. We wzorach Abate'a [1] miejsce f_s zajmuje f_{avg} . Różnice dotyczą wartości współczynników ewentualnie stałych występujących w tych wzorach. Współczynniki te lub stałe można określić na podstawie charakterystyk (rys. 6, 7) i w efekcie uzyskać eksperymentalną zależność pozwalającą wyznaczać SNR_{max} w funkcji BR_{avg} (i odwrotnie) dla modulacji z adaptacją częstości próbkowania.
- Dla modulacji NSDM i ANS-DM, w zależności od charakterystyk widmowych sygnału wejściowego następuje charakterystyczna „wymiana” zysku kompresji i wartości współczynnika SNR_{max} . I tak dla źródeł bogatych w składniki wysokoczęstotliwościowe następuje znaczna poprawa współczynnika SNR_{max} a niewielka współczynnika kompresji, natomiast dla źródeł z przewagą składników niskoczęstotliwościowych w widmie zmiany są odwrotne. Dla źródeł z równomiernym rozkładem widma następuje umiarkowany wzrost zarówno współczynnika kompresji, jaki i współczynnika SNR_{max} [8, 9]. Efekt „wymiany” można również uzyskać dla tego samego sygnału poprzez odpowiedni dobór parametrów wewnętrznych.
- Badania symulacyjne wykazały, że istnieje szereg ograniczeń, które nie pozwalają praktycznym rozwiązaniom systemów ADM na uzyskanie parametrów równych tym z obliczeń analitycznych. Niektóre z tych ograniczeń mają charakter tzw. błędów systemowych i bez zmiany metody nie mogą zostać usunięte. Systemy adaptacyjne podejmują decyzje, a co ważniejsze rozpoczynają zmianę parametrów adaptowanych, z opóźnieniem (na podstawie pewnej liczby bitów poprzednich) a więc nie zawsze adekwatnie do zachowania źródła w aktualnej chwili (jest to cechą wszystkich systemów z predykcją wsteczną). Efektem jest wzrost szumów (przeregulowania), które można nazwać szumami adaptacji lub lokalnymi. Pojawiają się one szczególnie przy zmianie nachylenia sygnału z dużego na małe i odwrotnie. W efekcie są powodem np. zmniejszenia SNR_{max} modulatorów adaptacyjnych [19].

Dlatego jego wartość jest często zbliżona do tej, którą uzyskuje się w modulacji LDM. Niektóre analizy [1] błędnie wskazują tę równość jako prawidłowość w systemach modulacji delta. Szumy adaptacji można minimalizować poprzez umiejętny dobór zarówno metody modulacji (np. ANS-DM a nie CFDM itp.) jak i parametrów adaptacji a w szczególności częstotliwości próbkowania [12, 19]. Drugi rodzaj błędów aproksymacji jest związany z praktycznym funkcjonowaniem algorytmu (błędy te niezależnie od tego, jaki algorytm zostanie wybrany zawsze występują a zmieniając algorytm można je tylko minimalizować w zależności od rodzaju aplikacji, w której modulator ma znaleźć zastosowanie). Należy podkreślić, że optymalizacja algorytmu pod kątem jednego z parametrów często prowadzi do pogorszenia innych, np. poprawa odporności algorytmu na zakłócenia bitów podczas transmisji wiąże się z pogorszeniem SNR_{max} i BR_{avg} [4, 18]. Problemy minimalizacji szumów lokalnych w systemach ADM w świetle badań symulacyjnych i analitycznych zostaną przedstawione w drugiej części niniejszej pracy.

8. BIBLIOGRAFIA

1. J. E. Abate: *Linear and Adaptive Delta Modulation*. Proceedings of the IEEE, Vol. 55, No. 3, March 1967, pp. 298-308.
2. R. Golański: *Adaptacja częstotliwości próbkowania jako metoda komandowania w modulacjach różnicowych*. Kwartalnik Elektrotechniki i Telekomunikacji 2/2001, Warszawa, PWN, 2001.
3. R. Golański, J. Kołodziej: *Próby analitycznego rozwiązania problemu wyznaczania średniej przepływności bitowej w modulatorach NSDM*. Mat. Konf.: KST'2001, Tom B, Bydgoszcz, wrzesień 2001.
4. R. Golański, J. Kołodziej, P. Świrek: *Badanie symulacyjne modulacji różnicowych ze zmiennym krokiem kwantyzacji i nierównomiernym próbkowaniem*. Mat. Konf. KST'99, Bydgoszcz 1999.
5. R. Golański: *Propozycja doboru parametrów adaptacji modulatora różnicowego z nierównomiernym próbkowaniem*. Mat. Konf. KST'98, Tom B, Bydgoszcz 1998.
6. R. Golański, R. Dmoch, J. Kochanowski, K. Szklarski: *Propozycja znajdowania optymalnych parametrów modulacji różnicowej NSDM w oparciu o badania symulacyjne*. Mat. Konf. KST, Tom B, Bydgoszcz 2000.
7. H. Taub, D.L. Schilling: *Principles of Communication Systems*. McGraw-Hill, Ltd. 1986.
8. Y.S. Zhu, W. Leung, C.M. Wong: *Adaptive Non-Uniform Sampling Delta Modulation for Audio/Image Processing*. IEEE Trans. Consumer Electron, vol. 42, No. 4, 1996, pp. 1062-1072.
9. Y.S. Zhu, W. Leung, C.M. Wong: *A digital audio processing system based on non-uniform sampling delta modulation*. IEEE Trans. on Consumer Electronics, vol. 42, No. 1, 1996, pp. 80-85.
10. R. Dmoch, R. Sadkiewicz: *Porównanie właściwości wybranych modulacji różnicowych a oparciu o wyniki symulacji algorytmicznych i rozważań analitycznych.*, Pr. Dyp., Katedra Elektroniki, Kraków, AGH, 2000.
11. J. Seidler, A. Badach, W. Molisz: *Metody rozwiązywania zadań optymalizacji*. Warszawa, WNT, 1980.
12. R. Golański, J. Kołodziej: *Wpływ częstotliwości startowej i współczynników adaptacji na parametry modulacji NSDM*. Mat. Konf.: KST'2001, Tom B, Bydgoszcz, wrzesień 2001.

13. R. C. ...
dulat
14. R. C. ...
Konf.
15. Wybr ...
g o, F
16. R. C. ...
Mat.
17. S. K. ...
mulac
18. L. C. ...
Pr. Dy
19. R. S. ...
20. Syntr
21. T. F. ...
niki A
22. M. F. ...
cowej
23. R. G. ...
Proc.
24. W. F. ...
często
25. C. K. ...
IEEE
26. J. J. ...
vol. 58
27. J. L. ...
sion op
28. R. G. ...
Proc. c

Author
improvement
puter simulation
NS-DM, ANS-DM
in graphical
DM modulation
is discussed
CFDM, NS-DM
performance

13. R. Golański, J. Kołodziej: *Problemy dyskretyzacji wartości szybkości próbkowania a modulatorach ANS-DM*. Mat. Konf. KST, Bydgoszcz 2003.
14. R. Golański, D. Kościelnik: *System badawczy i analiza modulacji różnicowych*. Mat. Konf. KST'93, Bydgoszcz 1993.
15. *Wybrane systemy i układy scalone w telekomunikacji cyfrowej*. Pr. zb. pod red. R. Golańskiego, Kraków, Wyd. AGH, 1995.
16. R. Golański: *Stanowisko do badania właściwości współczesnych modulatorów różnicowych*. Mat. Konf.: PWT'99, str.1.2-1 do 1.2-6, Poznań 1999.
17. S. Kończyk: *Wpływ zakłóceń i przerw w transmisji na pracę demodulatorów delta-metoda symulacyjna*, Pr. Dypl., Katedra Elektroniki AGH, Kraków 1998.
18. Ł. Olszewski: *Symulator modulatorów różnicowych ze zmienną częstotliwością próbkowania*. Pr. Dypl., Katedra Elektroniki AGH, Kraków 1999.
19. R. Steele: *Delta Modulation Systems*. London, Pentech Press, 1986.
20. Syntrillium Software Corporation— *Cool Edit Pro*, www.syntrillium.com.
21. T. Fornalik: *Porównanie właściwości modulacji NS-DM i ANS-DM*, Pr. Dypl., Katedra Elektroniki AGH, Kraków 2002.
22. M. Kilian: *Badania symulacyjne wpływu częstotliwości startowej na parametry modulacji różnicowej ze zmienną szybkością próbkowania*. Pr. Dypl., Katedra Elektroniki AGH, Kraków 2001.
23. R. Golański: *Influence of the Non-uniform Sampling on the Dynamic Range of Delta Modulators*. Proc. of 3rd Conf. on Telecom., Session TM4, Portugal, 2001, pp. 307-314.
24. W. Pogribny, I. Różnakowski, I. Zieliński, Z. Nowakowski: *Graniczna częstotliwość dyskretyzacji eliminująca przeciążenie delta-kodera*. Mat. Konf. KST'98, tom.B, Bydgoszcz, wrzesień 1998, pp. 390-395.
25. C.K. Un, D.H. Cho: *Hybrid Companding Delta Modulation with Variable-Rate Sampling*. IEEE Trans. On Comm., vol. COM-30, No. 4, 1982, pp. 593-599.
26. J.J. Dubnowski, R.E. Crochiere: *Variable Rate Coding of Speech*. Bell Sys. Tech. J., vol. 58, 1979, pp. 577-600.
27. J.L. LoCicero, P.P. Prezias: *Doubly adaptive delta modulation for bandwidth compression of speech signals*. Proc. of Nat. Telecom. Conf., 80 CH 1539-6, vol 2, 1980, pp. 36.5.1-36.5.5.
28. R. Golański: *Analytical Investigations of Delta Modulation with Variable Sampling Frequency*. Proc. of ICSES'2000, A.1.6, Ustroń, Poland, 2000, pp. 81-87.

R. GOLAŃSKI

COMPUTER SIMULATION OF THE DELTA MODULATIONS WITH SAMPLING FREQUENCY ADAPTATION

Summary

Author have been decided to investigate variable rate delta coding systems because of it's theoretical improvements over coding delta systems with fixed rate, especially for non-stationary input signals. Computer simulation results are presented. Characteristics of SNR ratio as a function of input signal level of NS-DM, ANS-DM modulation and the conventional LDM, CFDM have been obtained. The paper presents in graphical form dependences between the most important output parameters of the NS-DM and ANS-DM modulation. The performance of the SNR ratio, bit rate average, dynamic range and compression ratio is discussed. Simulation investigations for input signals in the form of .wave files on the emulator of LDM, CFDM, NSDM and ANS-DM have been carried out. Results of simulation investigations concerning the performance of these modulations have been compared. The following general comments can be made.

It was measured, that BR_{avg} is the parameter, which in decisive manner influences on SNR_{max} value in delta modulation systems with variable-sampling-rate (Fig.3, 4, 7, 9).

The maximum SNR ratio is greater by about 6 dB for variable-sampling delta system than LDM for stationary signals. Using for coding non-stationary signals the delta modulation with variable-sampling-rate one can improve SNR_{max} by 2 to 9 dB over the fixed-rate adaptive schemes (Fig. 2-9).

The shape of the curves of SNR ratio as a function of input signal power is similar for both variable- and fixed-sampling delta systems (Fig. 5, 8, 9).

The change of dynamic range from the fs adaptation is smaller then the change from the step size adaptation, which is constant and equal to 6 decibels per octave of the step size change (Fig. 9).

The total dynamic range of the ANS-DM modulation is a product of the dynamic range derived both from sampling frequency and step size adaptation (Fig. 9). Hence, ANS-DM modulation guarantees the same range as in case of the adaptive modulations where only single parameter is adapted, but at the within range of the each parameter. In practice, it is reflected by the increase of maximum value of SNR (Fig. 9). The doubly adapted delta systems (eg ANS-DM) require a lower BR_{avg} , and have a greater DR than the systems with step size adapted only. The ANS-DM modulation is modest in complexity compared with other medium range speech coders, but yields good performance regardless of channel conditions.

The results of computer simulation of the NS-DM and ANS-DM performances agree well with the theoretical evaluations [23, 28]. As seen from the results of computer simulation variable-rate-sampling improves the performances of delta modulations, at conversion of the highly non-stationary signal (i.e. speech and television).

Keywords: differential modulations, variable sampling frequency, computer simulation, dynamic range, compression ratio, signal-to-noise ratio.

V/R_{max} valuean LDM for
e-sampling-

th variable-

he step size
(9).nge derived
guaranteesd, but at the
lue of SNRgreater DR
y comparedconditions.
well with thee-sampling
signal (i.e.

mic

Elektrotermiczny makromodel sterownika przetwornic rezonansowych ZVS dla programu SPICE

KRZYSZTOF GÓRECKI, JANUSZ ZARĘBSKI

*Katedra Radioelektroniki Morskiej, Akademia Morska w Gdyni
ul.Morska 83, 81-225 Gdynia
E-mail:zarebski@am.gdynia.pl,
gorecki@am.gdynia.pl*

Otrzymano 2003.12.02

Autoryzowano 2004.02.20

Praca dotyczy problematyki modelowania układów sterujących przetwornic rezonansowych ZVS na przykładzie monolitycznego układu scalonego MC34067. W pracy sformułowano elektrotermiczny makromodel rozważanego sterownika dla programu SPICE. Poprawność opracowanego makromodelu zweryfikowano eksperymentalnie w opracowanym przez autorów układzie testowym.

Słowa kluczowe: sterowniki przetwornic rezonansowych, modelowanie, SPICE

1. WPROWADZENIE

Wśród powszechnie wykorzystywanych w układach zasilających impulsowych przetwornic dc-dc coraz większy udział mają przetwornice rezonansowe, charakteryzujące się małymi stratami mocy w elementach kluczujących w czasie przełączania. W zależności od sposobu włączenia elementów obwodu rezonansowego w obwód mocy przetwornicy, wyróżnia się przetwornice pracujące z tranzystorami włączanymi przy zerowej wartości napięcia (ZVS) lub wyłączanymi przy zerowej wartości prądu (ZCS) [1, 2, 3].

W procesie projektowania układów elektronicznych, np. przetwornic dc-dc, kluczową rolę odgrywają wiarygodne modele elementów elektronicznych występujących w tych układach oraz programy komputerowe, akceptujące postać tych modeli. Obecnie bardzo popularnym i wygodnym narzędziem obliczeniowym jest program SPICE, w którym wbudowane są modele podstawowych elementów elektronicznych, a biblio-

teki tego programu zawierają wartości parametrów modeli setek typów elementów elektronicznych lub makromodele półprzewodnikowych elementów mocy i układów scalonych, np. [4, 5]. Dodatkowo, na stronach internetowych producentów przyrządów półprzewodnikowych dostępne są makromodele przyrządów półprzewodnikowych przeznaczone dla programu SPICE. Niestety, jak dotychczas ani w bibliotekach tego programu, ani na stronach internetowych producenta nie ma modeli sterowników przetwornic rezonansowych.

Modele elementów elektronicznych w programie SPICE są modelami izotermicznymi, tzn. są one słuszne dla ustalonej wartości temperatury, która nie może ulegać zmianie w czasie trwania analizy. Tymczasem w rzeczywistych elementach półprzewodnikowych występuje zjawisko samonagrzewania, wynikające z wydzielania w elemencie energii elektrycznej i zamianie jej na ciepło przy nieidealnym chłodzeniu struktury półprzewodnikowej. Skutkiem samonagrzewania jest wzrost temperatury wnętrza elementu ponad temperaturę otoczenia oraz zmiana przebiegu charakterystyk tego elementu, nazywanych charakterystykami nieizotermicznymi. Uwzględnienie zjawiska samonagrzewania w analizach komputerowych wykonywanych za pomocą wymienionego programu wymaga sformułowania specjalnego rodzaju modelu, nazywanego modelem elektrotermicznym w postaci akceptowanej przez program SPICE.

Celem niniejszej pracy jest sformułowanie i weryfikacja poprawności elektrotermicznego makromodelu sterownika przetwornic rezonansowych ZVS typu MC34067 dla programu SPICE przy uwzględnieniu wpływu temperatury, napięcia zasilania oraz samonagrzewania na charakterystyki i parametry robocze tego sterownika. Przy formułowaniu tego makromodelu autorzy wykorzystali doświadczenia zdobyte przy opracowaniu elektrotermicznego makromodelu sterownika PWM [6, 7] oraz izotermicznego makromodelu sterownika rezonansowej przetwornicy ZVS [8], przeznaczonych dla programu SPICE.

W kolejnych rozdziałach przedstawiono strukturę wewnętrzną rozważanego w pracy sterownika, opisano postać makromodelu tego układu oraz przedstawiono wyniki obliczeń uzyskanych za pomocą zaproponowanego w pracy makromodelu, przy czym część z tych wyników zweryfikowano doświadczalnie.

2. STEROWNIK MC34067

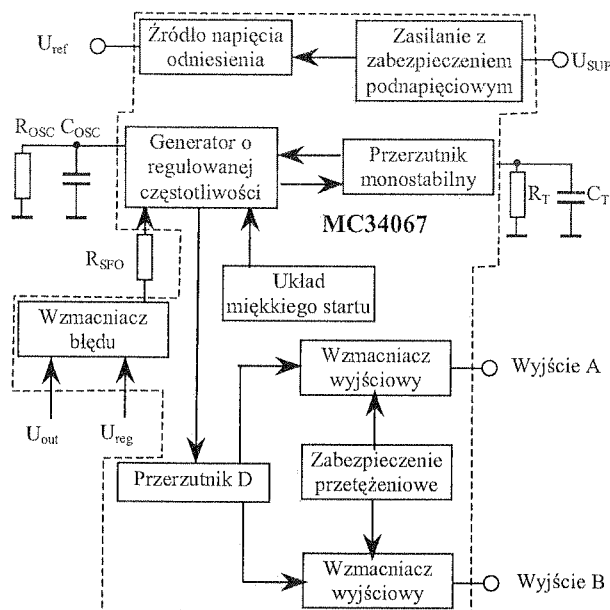
Sterownik MC34067 jest monolitycznym układem scalonym przeznaczonym do pracy w impulsowych stabilizatorach napięcia pracujących w trybie ZVS. Sterownik ten realizuje modulację częstotliwości sygnału sterującego elementy kluczujące przetwornicy przy stałej wartości czasu wyłączenia [9]. Schemat blokowy sterownika przedstawiono na rys. 1.

Układ MC34067 składa się z generatora przebiegu piłozębowego o częstotliwości regulowanej napięciem, wyzwalanego przerzutnika monostabilnego o precyzyjnie regulowanym czasie trwania impulsu, skompensowanego temperaturowo źródła napięcia odniesienia, szerokopasmowego wzmacniacza błędów o dużym wzmocnieniu, przerzut-

nika i p
żej wyd
impulsy
w przet
prądowa
tranzyst
zabezpie
czenie p
impulsó
błędów

Z p
którego
proporc
trwania
wnętrzn
Ze wzgl
gólne w
gnału os
częstotli

Blo
5,1 V, k



Rys. 1. Schemat blokowy sterownika MC34067

Fig. 1. The block diagram of MC34067 controller

nika i podwójnego układu totem-pole, stanowiącego końcowy stopień wyjściowy o dużej wydajności prądowej. Obecność w rozważanym układzie dwóch wyjść, na których impulsy są przesunięte w fazie o 180° umożliwia wykorzystanie tego sterownika m. in. w przetwornicach przeciwsobnych, np. w układzie mostkowym. Z kolei duża wydajność prądowa stopnia wyjściowego umożliwia bezpośrednie sterowanie z układu MC34067 tranzystorów kluczujących MOSFET lub IGBT. Rozważany sterownik zawiera także zabezpieczenie podnapięciowe, programowany układ miękiego startu oraz zabezpieczenie przetężeniowe. Zgodnie z danymi podanymi przez producenta [9], częstotliwość impulsów wyjściowych może zmieniać się nawet 1000-krotnie a pasmo wzmacniacza błędów wynosi 4 MHz.

Z punktu widzenia zasady działania układu kluczowym elementem jest generator, którego częstotliwość jest regulowana przez napięcie wyjściowe wzmacniacza błędów, proporcjonalne do różnicy napięcia odniesienia i części napięcia wyjściowego. Czas trwania stanu niskiego na wyjściu generatora jest precyzyjnie regulowany przez zewnętrzne elementy R_T oraz C_T dołączone do zacisków przerzutnika monostabilnego. Ze względu na konieczność sterowania elementami układu przeciwsobnego, poszczególne wyjścia sterownika są wysterowywane naprzemiennie w kolejnych okresach sygnału oscylatora i stąd częstotliwość impulsów wyjściowych jest dwa razy mniejsza od częstotliwości oscylatora. Podziału tej częstotliwości dokonuje przerzutnik D.

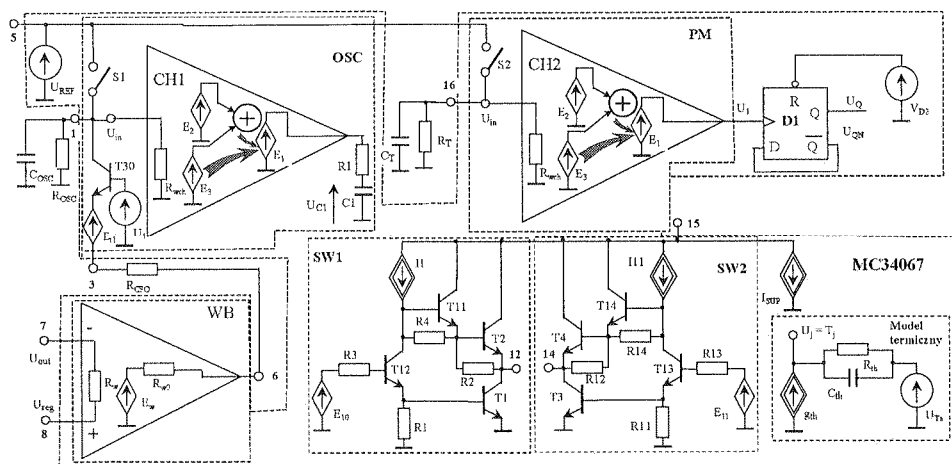
Blok napięcia odniesienia wytwarza na swoim wyjściu stałe napięcie o wartości 5,1 V, które zasilą część wewnętrznych bloków sterownika oraz jest wykorzystywane

jako napięcie wzorcowe dla wzmacniacza błędu. Układ napięcia zasilania z zabezpieczeniem podnapięciowym monitoruje wartość zewnętrznego napięcia zasilającego i blokuje pracę sterownika w przypadku zbyt niskiej wartości napięcia zasilania U_{SUP} .

Układ miękkiego startu zapewnia minimalną częstotliwość oscylatora po włączeniu zasilania i płynny jej wzrost do częstotliwości nominalnej ustalonej przez sygnał sprzężenia zwrotnego U_{out} , podawany na wzmacniacz błędu z wyjścia przetwornicy.

3. STRUKTURA MAKROMODELU

Makromodel rozważanego sterownika sformułowano w postaci obwodowej, akceptowanej przez program SPICE. Struktura makromodelu wynika z jego schematu blokowego, podawanego przez producenta tego układu [9] i zasady działania sterownika przetwornic ZVS, opisanej w rozdziale 2. Opracowany przez autorów makromodel opisuje najważniejsze zjawiska, występujące w rozważanym układzie scalonym, istotne przy jego pracy w stanie ustalonym, tzn. zależność czasu trwania impulsu oraz okresu sygnału wyjściowego od różnicy napięcia odniesienia i napięcia błędu, a także od wartości elementów zewnętrznych RC sterownika. Makromodel ten uwzględnia także pobór prądu ze źródła zasilania oraz zjawisko samonagrzewania. Ze względu na przeznaczenie makromodelu do analiz przetwornic rezonansowych w stanie ustalonym, pominięto w nim podukłady sterownika istotne tylko przy jego uruchomieniu, np. układ miękkiego startu oraz w stanach awaryjnych, np. układy zabezpieczeń przetężeniowego i podnapięciowego. Schemat blokowy opracowanego makromodelu przedstawiono na rys. 2.



Rys. 2. Struktura makromodelu sterownika MC34067

Fig. 2. The network representation of MC34067 macromodel

z zabez-
lającego
ia U_{SUP} .
włącze-
z sygnał
ornicy.

Zgodnie z zasadami podanymi w pracy [10], elektrotermiczny makromodel przyrządu półprzewodnikowego składa się z modelu elektrycznego, modelu termicznego oraz modelu mocy. W najbardziej złożonym, spośród wymienionych modeli — modelu elektrycznym można wyróżnić siedem bloków: źródło napięcia odniesienia (U_{REF}), wzmacniacz błędu (WB), oscylator (OSC), przerzutnik monostabilny (PM), przerzutnik D (D1) oraz dwa identyczne stopnie wyjściowe (SW1 oraz SW2). Sposób modelowania poszczególnych bloków opisano poniżej.

Wzmacniacz błędu (WB) zamodelowano jako quasi-idealny wzmacniacz operacyjny, w którego strukturze uwzględniono rezystancję wejściową R_w , sterowane źródło napięciowe E_w modelujące wzmocnienie wzmacniacza w otwartej pętli, natomiast rezystor R_{w0} reprezentuje jego rezystancję wyjściową. Wydajność źródła E_w dana jest wzorem

$$E_w = K_u \cdot U_{wew} \quad (1)$$

wej, ak-
chematu
terowni-
romodel
m, istot-
lsu oraz
, a także
zględnia
głędu na
stalonym,
p. układ
niowego
wiono na

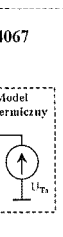
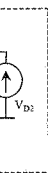
gdzie K_u oznacza wzmocnienie napięciowe wzmacniacza przy otwartej pętli sprzężenia zwrotnego, natomiast U_{wew} jest różnicowym napięciem wejściowym wzmacniacza. W modelu przyjęto typowe wartości katalogowe [9] parametrów wzmacniacza błędu: $R_w = 5 \text{ M}\Omega$, $K_u = 10^5$, $R_{w0} = 75 \text{ }\Omega$.

W prezentowanym makromodelu, wejścia wzmacniacza błędu pobudzane są napięciem zadany U_{reg} oraz napięciem sprzężenia zwrotnego U_{out} , natomiast jego wyjście steruje prądem rozładowania zewnętrznego kondensatora C_{osc} .

Oscylator (OSC) zbudowany jest z komparatora z histerezą (CH1), źródła napięciowego o stałej wydajności U_1 , tranzystora bipolarnego T30, sterowanego źródła napięciowego E_{t1} oraz przełącznika S1, sterowanego napięciem wyjściowym układu całkującego, zawierającego rezystor R1 i kondensator C1. Do zacisków oscylatora dołączone są zewnętrzne elementy R_{osc} , C_{osc} oraz R_{CF0} , umożliwiające regulację maksymalnej i minimalnej częstotliwości sygnału wyjściowego oscylatora. Oscylator pracuje w dwóch cyklach obejmujących ładowanie pojemności C_{osc} oraz jej rozładowanie. Pojemność ta jest ładowana ze źródła napięcia odniesienia U_{REF} w czasie załączenia przełącznika S1, przy czym wartość prądu ładowania ogranicza rezystancja włączenia tego przełącznika.

Przełącznik S1 jest załączany w sytuacji, gdy różnica między napięciem na pojemności C1 (U_{C1}), a wejściem komparatora CH1 wzrośnie powyżej 0,6 V, natomiast jest wyłączany, gdy różnica ta zmaleje poniżej 0,5 V. Regulacja częstotliwości sygnałów wyjściowych jest realizowana przez zmianę prądu rozładowania pojemności C_{osc} , płynącego przez rezystor R_{CF0} , za pomocą zmian napięcia wyjściowego wzmacniacza błędu WB. Tranzystor T30 blokuje prąd rezystora R_{CF0} , gdy napięcie wyjściowe wzmacniacza błędu jest większe od różnicy napięcia polaryzującego bazę tranzystora T30 (równego 3,1 V) oraz napięcia przewodzenia złącza baza-emiter tego tranzystora.

Czas rozładowania pojemności C_{osc} , a zatem i częstotliwość sygnału wyjściowego sterownika zależą m.in. od napięcia baza-emiter tranzystora T30. A zatem istotne znaczenie ma uwzględnienie w modelu zależności tego napięcia od temperatury. W celu



zamodelowania tej zależności zastosowano sterowane źródło napięciowe E_{t1} o wydajności danej wzorem

$$E_{t1} = h \cdot \left[\ln \left(\frac{i_{RCF0}}{I_0} \right) + 2 \cdot \ln \left(\frac{T_j}{T_0} \right) - 2 \right] \cdot (T_j - T_a) \quad (2)$$

gdzie T_j oznacza temperaturę wnętrza sterownika, wyznaczaną w modelu termicznym, omówionym w dalszej części rozdziału, T_a — temperaturę otoczenia, T_0 — temperaturę odniesienia, w której wyznaczono wartości parametrów modelu, h — iloraz stałej Boltzmanna przez ładunek elektronu, i_{RCF0} — prąd rezystora R_{CF0} , natomiast I_0 to parametr modelu tranzystora. Występujący w oscylatorze komparator z histerezą (CH1) zamodelowano za pomocą układu zawierającego rezystor R_{wch} reprezentujący rezystancję wejściową oraz trzy sterowane źródła napięciowe E_1 , E_2 oraz E_3 , modelujące jego charakterystykę przenoszenia, wraz z pętlą histerezy. Źródło napięciowe E_2 modeluje charakterystykę komparatora bez histerezy, natomiast źródło E_3 zapewnia przesunięcie tej charakterystyki w prawo o 1,4 V przy narastaniu napięcia wejściowego. Wydajności tych źródeł dane są wzorami

$$E_1 = 5 \cdot LIMIT(x, 1, 0) \quad (3)$$

gdzie $LIMIT$ jest standardową funkcją programu SPICE, natomiast zmienna pomocnicza x opisana jest zależnością

$$x = \begin{cases} E_2 & \text{dla } U_{in} < U_{min} \text{ oraz } U_{in} > U_{max} \\ E_2 + E_3 & \text{dla } U_{min} < U_{in} < U_{max} \end{cases} \quad (4)$$

$$E_2 = \begin{cases} 1 & \text{dla } U_{in} < U_{min} \\ \frac{1}{U_{mx} - U_{min}} * (U_{in} - U_{min}) & \text{dla } U_{min} < U_{in} < U_{mx} \\ 0 & \text{dla } U_{in} > U_{mx} \end{cases} \quad (5)$$

$$E_3 = \begin{cases} 0 & \text{dla } dU_{in}/dt < 0 \\ dU_{in}/dt & \text{dla } 0 < dU_{in}/dt < 1 \\ 1 & \text{dla } dU_{in}/dt > 1 \end{cases} \quad (6)$$

przy czym U_{in} oznacza napięcie wejściowe komparatora, U_{min} i U_{max} oznaczają wartości dolnego i górnego krańca histerezy. Na podstawie danych katalogowych [9] przyjęto następujące wartości parametrów oscylatora $U_{min} = 3.5$ V, $U_{max} = 4.9$ V, $U_{mx} = 3.51$ V oraz $R_{wch} = 1$ GΩ.

W modelach wzmacniacza błędu i komparatora pominięto pojemności wewnętrzne, ponieważ są one znacznie mniejsze od pojemności filtrującej występującej w obwodzie mocy sterowanej przetwornicy oraz od zewnętrznych pojemności oscylatora C_{osc} i przerzutnika monostabilnego C_T . Pojemności te oraz histereza charakterystyk komparatorów zapewniały stabilną pracę sterownika we wszystkich analizowanych przypadkach.

Źródło
napięciowe
napięcie
napięcia
Prze-

ny identy-
i C_T , dete-
napięcia
jest przez
wyjściow
wyłączon
osiąga sta-
jest mniej
mniejsze
trwania, r

Prze-

razy mniej
okresu, st-
pomocą w
poprawna
za pomoc-

o wysokoś-
Z kole-
źródła nap-
oraz odpo-
tych źróde-

Stopień

Każdy z ni-
prądowe. W
sterownika,
przyspiesze-

W mak-
bipolarnego
rów przyzł-
rezystancję

Model
Wartość ter-
źródła prąd-

Źródło napięcia odniesienia (U_{REF}) zamodelowano za pomocą niezależnego źródła napięciowego U_{REF} . Ponieważ poprawna praca układu wymaga, aby w chwili $t = 0$ napięcie odniesienia było równe zero, dlatego wydajność źródła U_{REF} opisano uskokiem napięcia o wartości 5,1 V.

Przerzutnik monostabilny (PM) zawiera komparator z histerezą CH2 (modelowany identycznie jak komparator CH1), przełącznik S2 oraz zewnętrzne elementy R_T i C_T , determinujące wartość czasu wyłączenia. Pojemność C_T jest ładowana ze źródła napięcia odniesienia U_{REF} przez włączony przełącznik S2, natomiast rozładowywana jest przez rezystor R_T . Przełącznik S2 jest włączony, gdy różnica między napięciem wyjściowym komparatora CH2 a napięciem U_{C1} jest większa niż 0,6 V, natomiast wyłączony, gdy różnica ta jest mniejsza od 0,5 V. Napięcie na wyjściu komparatora osiąga stan wysoki, gdy w czasie ładowania kondensatora napięcie na jego wejściu jest mniejsze od 4,9 V oraz gdy w czasie rozładowania kondensatora napięcie to jest mniejsze od 3,5 V. Komparator CH2 wytwarza na wyjściu ciąg impulsów o czasie trwania, równym różnicy okresu sygnału oscylatora i czasu wyłączenia.

Przerzutnik D (D1) umożliwia uzyskanie dwóch sygnałów o częstotliwości dwa razy mniejszej od częstotliwości oscylatora i przesuniętych względem siebie o połowę okresu, sterujących obydwoma stopniami wyjściowymi sterownika. Przerzutnik D opisano za pomocą wbudowanego w programie SPICE modelu tego układu cyfrowego, przy czym poprawna praca tego układu wymagała wytworzenia sygnału zerującego, realizowanego za pomocą źródła napięciowego V_{D2} , którego wydajność opisano uskokiem napięcia o wysokości równej napięciu zasilania układu cyfrowego (5 V).

Z kolei, stopnie wyjściowe, pracujące w układzie totem-pole, sterowane są przez źródła napięciowe E_{10} i E_{11} , realizujące negację iloczynów sygnału wejściowego u_1 oraz odpowiedniego sygnału wyjściowego u_Q lub u_{QN} przerzutnika D. Wydajności tych źródeł dane są wzorami

$$E_{10} = LIMIT(-2 \cdot LIMIT(u_1 \cdot (u_Q - 1), 0, 15) + 15, 0, 6) \quad (7)$$

$$E_{11} = LIMIT(-2 \cdot LIMIT(u_1 \cdot (u_{QN} - 1), 0, 15) + 15, 0, 6) \quad (8)$$

Stopień wyjściowy składa się z dwóch identycznych bloków (SW1 oraz SW2). Każdy z nich zawiera cztery tranzystory bipolarne, cztery rezystory oraz jedno źródło prądowe. W celu zapewnienia odpowiednio szybkiego przełączania sygnału na wyjściu sterownika, wykorzystano rezystory bocznikujące złącza baza-emiter tranzystorów do przyspieszenia procesu ich wyłączania [1, 11].

W makromodelu wykorzystano wbudowany w programie SPICE model tranzystora bipolarnego, przy czym w celu uwzględnienia spadku napięcia na rezystancji obszarów przyłączowych w czasie włączenia tranzystorów uwzględniono w ich modelach rezystancję szeregową kolektora.

Model termiczny bazuje na koncepcji przejściowej impedancji termicznej [10]. Wartość temperatury wnętrza wyznaczana jest w układzie złożonym ze sterowanego źródła prądowego g_{th} , rezystora R_{th} , kondensatora C_{th} oraz źródła napięciowego U_{Ta} .

Układ ten stanowi odpowiednik obwodowy skupionego modelu termicznego elementu półprzewodnikowego, w tym także układu scalonego [10]. W modelu tym, temperatura wnętrza T_j , odpowiadająca napięciu U_j na zaciskach źródła prądowego g_{th} , jest sumą temperatury otoczenia T_a oraz nadwyżki temperatury wynikającej z samonagrzewania, natomiast R_{th} i C_{th} reprezentują przejściową impedancję termiczną sterownika.

W celu skrócenia czasu trwania obliczeń niezbędnego do uzyskania stanu ustalonego, z uwagi na przeznaczenie rozważanego makromodelu do analiz w stanie ustalonym i znaczne różnice między okresem sygnału sterującego a rzeczywistymi termicznymi stałymi czasowymi rozważanego sterownika, przyjęto w modelu termicznym tylko jedną termiczną stałą czasową. Uproszczenie to jest uzasadnione, gdyż najkrótsze termiczne stałe czasowe są typowo co najmniej o rząd wielkości większe od okresu sygnału wyjściowego sterownika, a zatem w czasie trwania jednego okresu sygnału wyjściowego temperatura wnętrza sterownika praktycznie nie zmienia się (zmiany nie przekraczają 0,1%). Stąd termiczną stałą czasową w modelu dobrano tak, aby była ona znacznie dłuższa od okresu sygnału wyjściowego, a jednocześnie znacznie krótsza od stałych czasowych obwodu mocy przetwornicy, w której pracuje rozważany sterownik.

Moc wydzielana w sterowniku jest sumą dwóch składników: mocy wydzielanej w stopniu wyjściowym, której wartość zależy silnie od obciążenia sterownika oraz mocy wykorzystywanej do zasilania poszczególnych bloków sterownika, reprezentowanej przez moc wydzielaną na źródle prądowym I_{SUP} . A zatem, wydajność źródła prądowego g_{th} , modelującego moc wydzielaną w układzie dana jest wzorem

$$g_{th} = U_{SUP} \cdot I_{SUP} + \sum_{i=1}^4 i_{Ci} \cdot u_{CEi} \quad (9)$$

gdzie U_{SUP} oznacza napięcie zasilania sterownika, natomiast i_{Ci} oraz u_{CEi} oznaczają odpowiednio prądy kolektora oraz napięcia kolektor-emiter tranzystorów końcowych stopnia wyjściowego sterownika.

4. WYNIKI SYMULACJI I POMIARÓW

W celu weryfikacji poprawności opracowanego modelu wykonano pomiary i obliczenia charakterystyk rozważanego sterownika, pracującego w układzie testowym, przedstawionym na rys. 3. Zakres badań obejmował symulacje i pomiary charakterystyk sterownika przy różnych wartościach pojemności C_{OSC} oraz C_T , napięcia zasilania, rezystancji obciążenia oraz temperatury otoczenia.

W obliczeniach wykorzystano następujące wartości parametrów makromodelu: $R_{th} = 100 \Omega$, $C_{th} = 120 \text{ pF}$, $I_{SUP} = 25 \text{ mA}$, $R_1 = 100 \Omega$, $C_1 = 200 \text{ pF}$, $I_1 = I_{11} = 1 \text{ mA}$, $R_4 = R_{14} = 2 \text{ k}\Omega$, $R_{12} = R_2 = 100 \Omega$, $R_1 = R_{11} = 1 \text{ k}\Omega$, $R_3 = R_{13} = 10 \Omega$, $T_0 = 27^\circ\text{C}$, parametry tranzystora T30: $I_0 = 6 \cdot 10^6 \text{ A}$, $I_S = 30 \text{ fA}$, $BF = 250$, $BR = 10$, $TR = 50 \text{ ps}$, $TF = 10 \text{ ps}$, $C_{JE} = 4 \text{ pF}$, $C_{JC} = 2 \text{ pF}$, $RC = 1 \Omega$, parametry pozostałych tranzystorów: $I_S = 1 \text{ fA}$, $BF = 250$, $BR = 10$, $TR = 5 \text{ ns}$, $TF = 0,6 \text{ ns}$, $C_{JE} = 15 \text{ pF}$, $C_{JC} = 5 \text{ pF}$, RC

elementu
temperatura
jest sumą
przewania,
ka.

ustalone-
ustalonym
termiczny-
nym tylko
rótsze ter-
od okresu
u sygnału
miany nie
y była ona
krótsza od
sterownik.
dzielanej
nika oraz
prezentow-
ość źródła
n

(9)

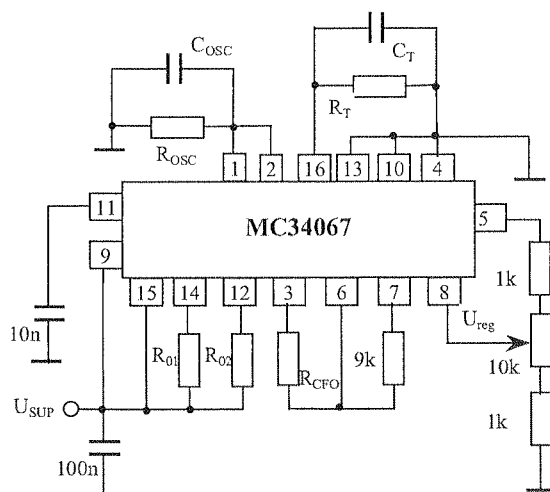
oznaczają
końcowych

niary i ob-
testowym,
charaktery-
a zasilania,

romodelu:
 $I_{11} = 1 \text{ mA}$,
 $T_0 = 27^\circ\text{C}$,
 $R = 50 \text{ ps}$,
nzystorów:
 $= 5 \text{ pF}$, RC

$= 50 \Omega$, $R_B = 200 \Omega$. Wartości parametrów występujących w opisie oscylatora oraz wzmacniacza błęd podano w rozdziale 3.

Na rys. 4 przedstawiono obliczone i zmierzone nieizotermiczne zależności czasu trwania impulsu t_w oraz okresu T od napięcia zadanego U_{reg} dla temperatury otoczenia $T_a = 20^\circ\text{C}$ przy dwóch wartościach pojemności C_{OSC} . Pozostałe, zewnętrzne elementy RC miały następujące wartości $R_{\text{OSC}} = 18 \text{ k}\Omega$, $R_T = 2,7 \text{ k}\Omega$, $C_T = 100 \text{ pF}$, $R_{01} = R_{02} = R_0 = 240 \Omega$.



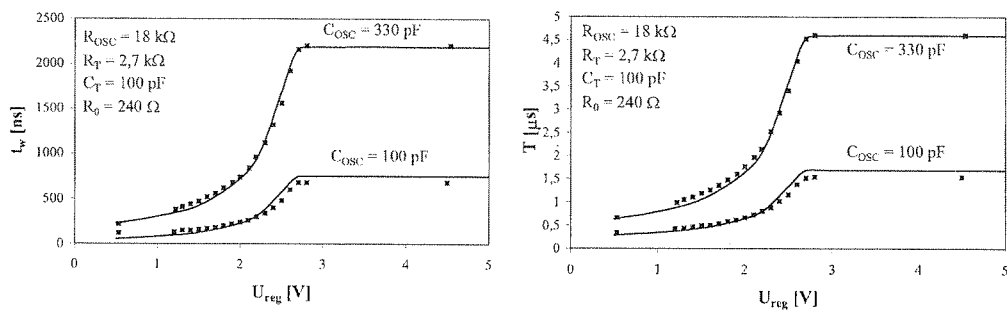
Rys. 3. Układ testowy sterownika MC34067

Fig. 3. The test circuit of MC34067 controller

Jak widać, uzyskano dobrą zgodność wyników obliczeń i pomiarów. Warto zauważyć, że wartości parametrów impulsów wyjściowych (t_w oraz T) zmieniają się prawie dziesięciokrotnie w zakresie regulacji, a zmiana wartości pojemności C_{OSC} z 330 pF na 100 pF powoduje ponad trzykrotne skrócenie czasu trwania impulsu i okresu sygnału wyjściowego.

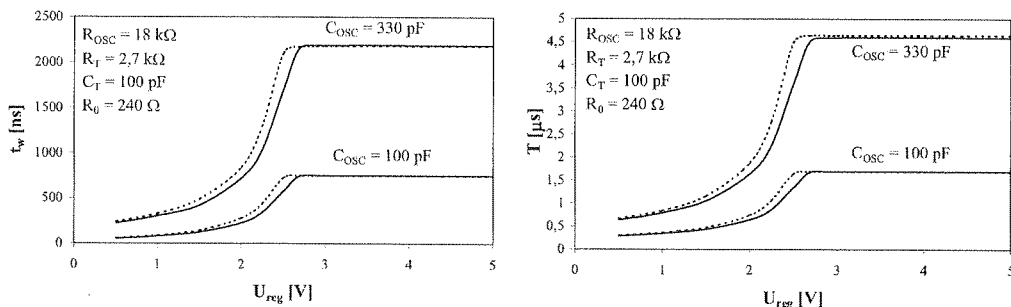
Rys. 5 ilustruje wpływ samonagrzewania na zależność $t_w(U_{\text{reg}})$ oraz $T(U_{\text{reg}})$. Na rysunku tym, linią ciągłą oznaczono wyniki obliczeń nieizotermicznych, natomiast linią kreskową - obliczeń izotermicznych. Jak widać, w rozpatrywanym przypadku różnice w przebiegu charakterystyk są widoczne, zwłaszcza w zakresie małych wartości U_{reg} . Samonagrzewanie powoduje skrócenie czasu trwania impulsu i okresu sygnału wyjściowego oraz rozszerzenie zakresu zmian napięcia U_{reg} , w którym następuje regulacja parametrów impulsów wyjściowych.

Na kolejnych rysunkach przedstawiono obliczone zależności średniego prądu zasilania I_{SUP} (rys. 6) oraz temperatury wnętrza T_j (rys. 7) od napięcia zadanego U_{reg} dla różnych wartości pojemności C_{OSC} i rezystancji obciążenia R_0 . Na rys. 6 zaznaczono dodatkowo wyniki pomiarów (punkty).



Rys. 4. Obliczone i zmierzone nieizotermiczne zależności $t_w(U_{reg})$ oraz $T(U_{reg})$ dla sterownika MC34067

Fig. 4. The calculated and measured nonisothermal $t_w(U_{reg})$ and $T(U_{reg})$ characteristics of MC34067 controller



Rys. 5. Zależność czasu trwania impulsu wyjściowego od napięcia odniesienia w warunkach izotermicznych i nieizotermicznych

Fig. 5. The measurement and calculation results of t_w and T on U_{reg} for the isothermal and nonisothermal case

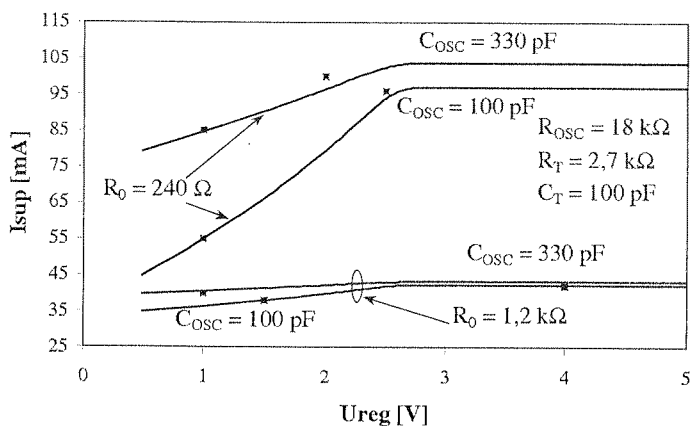
W przypadku zależności $I_{SUP}(U_{reg})$ uzyskano zadawalającą zgodność wyników symulacji i pomiarów. Obserwowany wzrost prądu zasilania I_{SUP} wraz ze wzrostem napięcia U_{reg} wynika ze wzrostu wartości współczynnika wypełnienia impulsów wyjściowych.

Jak widać na rys. 7, temperatura wnętrza maleje wraz ze wzrostem napięcia U_{reg} . Obserwowany wzrost temperatury T_j przy spadku wartości rezystancji obciążenia R_0 wynika ze wzrostu mocy wydzielanej w tranzystorach wyjściowych. Z kolei, dla mniejszych wartości pojemności C_{OSC} występuje większa częstotliwość przełączania i większy jest udział w mocy średniej jej składnika związanego z przełączaniem, a stąd uzyskuje się większe wartości temperatury T_j .

Wzrost temperatury nie powoduje jakościowych zmian w przebiegu charakterystyk sterownika, a zmiany ilościowe dochodzą do około 25% przy wzroście temperatury wnętrza sterownika o około 70°C. Zmiany napięcia zadanego U_{reg} powodują zmiany

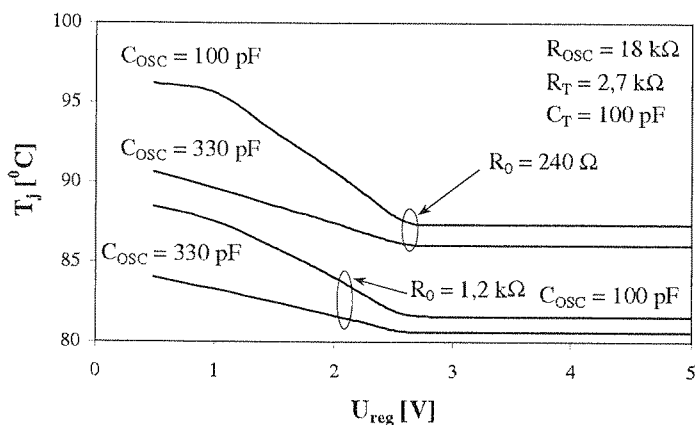
wartości
zasilania

W p
scaloneg
stawie p
makrom



Rys. 6. Zależność średniego prądu zasilania sterownika od napięcia zadanego

Fig. 6. The dependence of the controller supply current on controlling voltage



Rys. 7. Zależność temperatury wnętrza sterownika od napięcia zadanego

Fig. 7. The dependence of the controller inner temperature on controlling voltage

wartości tej temperatury zaledwie o kilka stopni Celsjusza, natomiast zmiany prądu zasilania przy obciążeniu $R_0 = 240 \text{ }\Omega$ dochodzą nawet do 50%.

5. PODSUMOWANIE

W pracy przedstawiono opracowany przez autorów elektrotermiczny makromodel scalonego sterownika przetwornic rezonansowych ZVS dla programu SPICE. Na podstawie przeprowadzonych symulacji i pomiarów wykazano poprawność opracowanego makromodelu na przykładzie sterownika MC34067 w szerokim zakresie zmian napięcia

zadanego U_{reg} oraz wartości zewnętrznych elementów RC. Makromodel ten umożliwia wyznaczenie charakterystyk zaciskowych sterownika oraz jego temperatury wnętrza, ograniczającej zakres bezpiecznej pracy oraz niezawodność sterownika.

Przedstawiony w niniejszej pracy makromodel może być wykorzystany zarówno przy projektowaniu układów zasilających z przetwornicami ZVS, jak i w dydaktyce do ilustracji działania takich przetwornic.

Opracowany makromodel przeznaczony jest do analiz sterownika w stanie ustalonym, natomiast wykorzystanie go do analiz stanów przejściowych, np. po załączeniu zasilania wymaga uzupełnienia jego struktury o układ miękkiego startu oraz układy zabezpieczeń podnapięciowego i przetężeniowego.

6. BIBLIOGRAFIA

1. N. Mohan, T. M. Undeland, W. P. Robbins: *Power Electronics: Converters, Applications, and Design*. New York, John Wiley & Sons, 1995.
2. M. Kazimierzczuk, D. Czarkowski: *Resonant Power Converters*. New York, Wiley & Sons, 1995.
3. R. W. Ericson, D. Maksimovic: *Fundamentals of Power Electronics*. Kluwer Academic Publisher, 2001.
4. Switch-reg.lib, biblioteka programu SPICE, 1998.
5. Ch. P. Basso: *Switch-Mode Power Supply SPICE Cookbook*. McGraw-Hill, 2001.
6. J. Zarębski, K. Górecki: *SPICE Modelling of PWM Controllers*. IEEE International Power Electronics Congress CIEP'2002, Guadalajara, 2002, p. 61.
7. J. Zarębski, K. Górecki: *Temperature Affects the UC3842 Current Mode PWM Controller*. 9th International Workshop on Thermal Investigations of ICs and Systems THERMINIC, Aix-en-Provence, 2003, p. 117.
8. K. Górecki, J. Zarębski: *A new SPICE macromodel of ZVS resonant converter controller*. 6th International Seminar on Power Semiconductors ISPS'2002, Praga 2002, p. 195.
9. MC34067, MC33067 High Performance Resonant Mode Controllers. Catalogue data. ON Semiconductor, 2000.
10. J. Zarębski: *Modelowanie, symulacja i pomiary przebiegów elektrotermicznych w elementach półprzewodnikowych i układach elektronicznych*. Prace Naukowe WSM w Gdyni, Gdynia, 1996.
11. W. J. Stepowicz: *Elementy półprzewodnikowe*. Wydawnictwo Akademii Morskiej, Gdynia, 2002.

K. GÓRECKI, J. ZARĘBSKI

ELECTROTHERMAL MACROMODEL OF THE CONTROLLER OF ZVS RESONANT CONVERTER FOR SPICE

S u m m a r y

For a long time, the switching-mode dc-dc converters are a very important class of power electronic circuits. Among these circuits, the resonant converters switching at zero voltage (ZVS) are particularly interesting.

In
compon
are nec
a produ
Unfortu
users of
Th
for SPIC
block di
converte
phenom
Th
features
(resultin
on the to
Be
steady-st
omitted
In
of the co
voltage s
the sawto
external
signal as
generate
and the
controls
Foster ne
Th
Some no
of supply
terminals
in TC. T
A s
pulse dur

Keyword

In the process of designing of such a class of converters, full credible models describing the converter components, or describing a whole circuit, play the fundamental role and the proper computer tools are necessary as well. Today, SPICE along with the built-in models and macromodels worked out by a producer of some devices, such as: diodes, transistors, inductors, transformers etc. is a comfortable tool. Unfortunately, there are no models or macromodels of resonant converter controllers available for the users of SPICE.

The aim of this paper is to propose a new form of MC34067 resonant converter controller macromodel for SPICE including phenomena of selfheating. The macromodel structure results both from the controller block diagram presented in catalogue data and the principle of operation of the controller of the resonant converters. The formulated electrothermal macromodel is dedicated to describe the fundamental physical phenomena existing in the considered IC, indispensable in its steady-state analysis.

This macromodel, formed as an electrical network, describes the fundamental and most important features of the considered IC, as the influence of the ambient temperature, supply voltage and selfheating (resulting from changing the electrical energy dissipated in ICs to heat at non-ideal cooling conditions) on the terminal characteristics and its working parameters.

Because the presented macromodel is designed for the analysis of the resonant converters in the steady-state, such phenomena, as the soft-start, the overcurrent and undervoltage protection have been omitted in the macromodel construction.

In the macromodel, two signals — the reference voltage and the error voltage arises from the output of the converter, excite the inputs of the error amplifier (WB). The oscillator is composed of the reference voltage source (U_{REF}), the comparator with the hysteresis (CH1) and the controlled switch (S1). To generate the sawtooth pulses of the time duration depending on the value of the error amplifier output voltage, three external elements R_{OSC} , C_{OSC} and R_{CFO} are needed. The switch S2 is controlled by the use of the identical signal as the switch S1, whereas the external elements C_T and R_T determine the off-time. Comparator generates the pulse train of the time, being the difference of the period of the switch controlling signal and the off-time. The pulse train of the required frequency (after passing through the flip-flop circuit) controls two Totem-Pole output circuit. The lumped thermal model of the considered IC in the form of Foster network is used.

The proposed macromodel was verified experimentally in the test circuit (TC) designed by authors. Some nonisothermal (d.c. measurements at the thermal steady-state) measurements for the various values of supply voltage, the load resistance as well as for some values of RC elements connected to the oscillator terminals and determining the period and the pulse duration time of output signals, have been performed in TC. The supply current and the inner temperature of controller are calculated as well.

A satisfactory agreement between the calculations and measurements of the dependencies of the pulse duration time and the period of the output signal on the controlling voltage have been achieved.

Keywords: resonant converter, semiconductor devices, modelling, SPICE

Perc

Fin
we
sys
con
ad
exp
me

Ke

The
vestigate
the prob
we obse
rise of n
conferen
availabil
with inc
speech e

Mos
especial

Perceptually Motivated Approaches to Speech Enhancement — Part I: Elimination of the Musical Noise Phenomenon

ADAM BOROWICZ, ALEXANDER A. PETROVSKY

*Wydział Informatyki, Politechnika Białostocka
ul. Wiejska 45A, 15-351 Białystok
E-mail: borowicz@ii.bialystok.pl*

*Otrzymano 2003.06.11
Autoryzowano 2003.11.14*

This paper addresses the problem of noise reduction in speech communication devices. Firstly, basis of the spectral weighting techniques and related problems are presented. Next, we describe psychoacoustic principles as a tool for improvement of speech enhancement systems. Conventional methods aimed at elimination of residual noise and their psychoacoustic modifications are discussed. The differences between them are emphasized. Finally, advantages of perceptual approaches over conventional ones are presented. Paper shows that exploitation of psychoacoustic models in speech enhancement provides significant improvements especially residual noise reduction and decrease of speech distortion.

Keywords: noise reduction, echo cancellation, psychoacoustic.

1. INTRODUCTION

The problem of noise reduction in speech communication devices has been investigated for many years. Although many sub-optimal solutions have been developed, the problem is not fully solved and it still exists now [1]. Moreover, in recent years we observe an increasing interest in this field of research. This trend is caused by the rise of requirements for modern, military and civilian communication devices (teleconferencing and speech recognition systems, hands-free telephony). Additionally, the availability of new digital signal processors is supporting the design of new algorithms with increased computational complexity. Actually there is a strong need to improve speech enhancement techniques, especially to increase an intelligibility of the speech.

Most of the known suppression systems are based on spectral weighting technique, especially on spectral subtraction [2, 3]. Although these methods are very simple and

easy to implement, their drawback is the residual noise, also known as musical tones, and the distortion of the speech. In effect, intelligibility of the speech is drastically decreased. Moreover, in very noisy conditions, this problem grows rapidly and the enhanced speech may be even more disturbed for a human listener than the original noisy signal.

For these reasons many modifications of standard weighting rules have been developed [4–7]. Unfortunately, most of these approaches solve the problem only partially. However, recent studies [8–13] show that it is possible to obtain significant improvements by using perceptual properties of the human auditory system.

In Part I of the review, we describe conventional methods aimed at elimination of residual noise together with their psychoacoustic modifications. Section 2 covers a basis of spectral weighting technique and related problems. In Section 3 we present conventional spectral weighting methods aimed at suppression of residual noise. In Section 4 we describe the psychoacoustic properties of the human auditory system. Perceptual modifications of the conventional spectral weighting rules are discussed in Section 5. Especially, the comparison of improvements in musical noise cancellation and increasing speech intelligibility is performed. In Section 6 we summarize paper.

In Part II, we will cover recent psychoacoustically motivated approaches [10–13] to speech enhancement and discuss perspectives of joint noise suppression and coding system based on psychoacoustic principles. Both parts present an evolution of psychoacoustically motivated approaches to single channel speech enhancement.

2. INTRODUCTION TO SPEECH ENHANCEMENT

In single channel speech enhancement, it is usually assumed that the time-sampled, disturbed signal $y(k)$ is a sum of a clean speech signal $s(k)$ and a noise $n(k)$:

$$y(k) = s(k) + n(k), \quad (1)$$

where integer k denotes the time index. Moreover, we assume that $s(k)$ and $n(k)$ are statistically independent and stationary in a specified time interval [14]. Most speech enhancement algorithms operate in frequency domain in a frame-by-frame basis. Input signal is divided into overlapped frames. Each frame is multiplied by window function and transformed into frequency domain (in this paper, we often omit frame index for simplicity). Frequency decomposition is usually implemented with Fast Fourier Transform (FFT) or filter bank. This process is often called as Short-Time Spectral Analysis (STSA). A noise reduction is usually achieved by using the spectral weighting technique. In this approach, the noise is suppressed by multiplying spectral magnitude of the noisy speech $|Y(\omega)|$ by gain function $0 \leq G(\omega) \leq 1$, where $\omega \in \langle -\pi; \pi \rangle$. The phase of the disturbed signal is not modified. After magnitude modification, enhanced spectrum is transformed back into time domain and synthesized with overlap-add method. The simplest scheme of this method is given in Fig. 1.

$$y(k)=s(k)$$

Fig. 1. E

Rys. 1.

All
between
and differ
only in sp
use minim
spectral c
where the
coefficient
touch this
popular w

Metho
tion [2, 3]
speech sp

where $\hat{N}(\omega)$
the metho
spectral su
Subtraction
context of
formulated

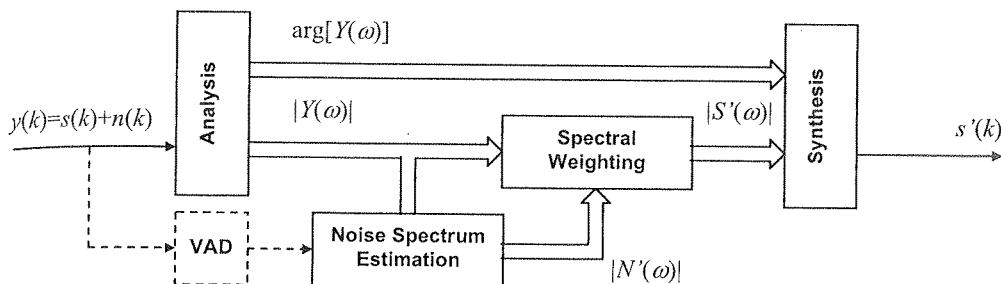


Fig. 1. Block diagram of simple speech enhancement system, based on spectral weighting technique.

Rys. 1. Schemat blokowy prostego systemu uzdatniania sygnału mowy opartego na metodzie wag widmowych.

All methods based on spectral weighting are very similar. The main difference between these methods is the using different ways to estimate the gain function $G(\omega)$ and different noise estimation techniques. In classical methods, the noise is estimated only in speech pauses using Voice Activity Detector (VAD) [15]. Another approaches use minimum statistics method [2, 16], which traces minimum noise level for each spectral component over specified time interval. Soft-decision approaches are also used where the probability of speech absence is estimated in order to modify spectral gain coefficients [17]. Because noise estimation is beyond scope of this review, we will not touch this subject. In the next part of this section, we will present some of the most popular weighting rules.

2.1. SPECTRAL SUBTRACTION

Method called spectral subtraction is one of the simplest methods of noise reduction [2, 3]. In this case, the enhanced speech spectrum $\hat{S}(\omega)$ is obtained from noisy speech spectrum $Y(\omega)$ using the following expression:

$$|\hat{S}(\omega)|^a = |Y(\omega)|^a - \lambda |\hat{N}(\omega)|^a, \quad (2)$$

where $\hat{N}(\omega)$ denotes an estimate of the noise spectrum. Parameter a defines type of the method, $1 \leq \lambda \leq 2$ is an over-subtraction factor. We can distinguish two types of spectral subtraction methods, depending on value of parameter a : Magnitude Spectral Subtraction (MMS) for $a = 1$ or Power Spectral Subtraction (PSS) for $a = 2$. In the context of spectral weighting technique, the gain function corresponding to (2) can be formulated as follows:

$$G(\omega) = \left(1 - \lambda \frac{|\hat{N}(\omega)|^a}{|Y(\omega)|^a} \right)^{1/a}. \quad (3)$$

2.2. WIENER RULE

Another frequently used spectral weighting method is a Wiener filtering [17, 18]. This method is based on the minimization of mean square error value, between clean and enhanced speech spectrum:

$$E \left[|S(\omega) - \hat{S}(\omega)|^2 \right]. \quad (4)$$

The minimization leads to the definition of the gain function as follows:

$$G(\omega) = \frac{R_{ss}(\omega)}{R_{yy}(\omega)} = \frac{R_{yy}(\omega) - R_{nn}(\omega)}{R_{yy}(\omega)}, \quad (5)$$

where $R_{ss}(\omega)$, $R_{yy}(\omega)$, $R_{nn}(\omega)$ are the Power Spectral Densities (PSD) of the clean speech, the noisy speech and the noise, respectively. On the other hand, the gain function can be expressed through a priori Signal to Noise Ratio (SNR) too:

$$G(\omega) = \frac{\text{SNR}_{\text{prio}}(\omega)}{1 + \text{SNR}_{\text{prio}}(\omega)}, \quad (6)$$

where

$$\text{SNR}_{\text{prio}}(\omega) = \frac{R_{ss}(\omega)}{R_{nn}(\omega)}. \quad (7)$$

Wiener rule is characterized by following properties: when SNR is high, the gain function tends asymptotically to one, otherwise, if SNR is low, the value of gain function is close to zero. The second effect is sometimes undesired, because enhanced speech may contain unnatural silence intervals. Hence, in order to keep background noise characteristic, the gain function is limited to any small value instead of zero.

2.3. EPHRAIM AND MALAH WEIGHTING RULE

Ephraim and Malah have proposed two spectral weighting rules for noise reduction [4, 5]. The first one (9) is based on the minimization of mean square error value, between clean and enhanced speech magnitude spectrum:

$$E \left[\left(|S(\omega)| - |\hat{S}(\omega)| \right)^2 \right]. \quad (8)$$

In this case, the minimization leads to the definition of the following gain function:

$$G(\omega) = \frac{\sqrt{\pi}}{2} \sqrt{\left(\frac{1}{1 + \text{SNR}_{\text{post}}(\omega)} \right) \left(\frac{\text{SNR}_{\text{prio}}(\omega)}{1 + \text{SNR}_{\text{prio}}(\omega)} \right)}, \quad (9)$$

$$\cdot M \left[\left(1 + \text{SNR}_{\text{post}}(\omega) \right) \frac{\text{SNR}_{\text{prio}}(\omega)}{1 + \text{SNR}_{\text{prio}}(\omega)} \right]$$

where M is defined as:

$$M(\theta) = \exp\left(-\frac{\theta}{2}\right) \left[(1 + \theta)I_0\left(\frac{\theta}{2}\right) + \theta I_1\left(\frac{\theta}{2}\right) \right], \quad (10)$$

where I_0 , I_1 are the modified Bessel functions of zero and first order, respectively. The second rule (12) proposed by Ephraim and Malah is based on the minimization of mean square error value, between clean and enhanced speech magnitude spectrum logarithm:

$$E \left[\left(\log |S(\omega)| - \log |\hat{S}(\omega)| \right)^2 \right]. \quad (11)$$

This criterion gives better results because a distortion measure, which is based on the mean square error of the log-spectra, is more suitable for speech processing. As a result, we obtain:

$$G(\omega) = \frac{\text{SNR}_{\text{prio}}(\omega)}{1 + \text{SNR}_{\text{prio}}(\omega)} \exp \left[\frac{1}{2} \int_{\Psi(\omega)}^{\infty} \frac{e^{-\tau}}{\tau} d\tau \right], \quad (12)$$

where

$$\Psi(\omega) = \frac{\text{SNR}_{\text{prio}}(\omega)}{1 + \text{SNR}_{\text{prio}}(\omega)} (\text{SNR}_{\text{post}}(\omega) + 1). \quad (13)$$

In both cases, it is necessary to estimate not only a priori SNR (7) but also a posteriori SNR, which is (14) defined as follow:

$$\text{SNR}_{\text{post}}(\omega) = \frac{R_{yy}(\omega) - R_{nn}(\omega)}{R_{nn}(\omega)} = \frac{R_{yy}(\omega)}{R_{nn}(\omega)} - 1. \quad (14)$$

2.4. ECHO CANCELLATION PROBLEM

Sometimes, background noise are not only one distortion in speech communication devices, another one is echo. This phenomenon occurs in hands-free telephony in situation when microphone is located some significant distance from the speaker and it is in field of loudspeaker. In other words, there is an effect of acoustic feedback between loudspeaker and microphone. Term of acoustic echo comes from the fact that far-end speaker can hear his own talk delayed by transmission time.

Echo cancellation problem [19] is originally solved by acoustic echo control systems. The main task of such systems is to avoid echo transmission to far-end speaker. They are based on adaptive filters, which model acoustic echo path. On this way, previous speech model (1) can be extended to the following form

$$y(k) = s(k) + n(k) + d(k), \quad (15)$$

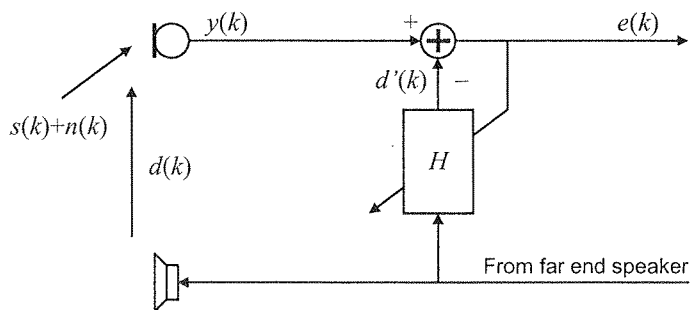


Fig. 2. Block diagram of typical acoustic echo cancellation system.

Rys. 2. Schemat blokowy typowego systemu eliminacji echa akustycznego.

where $d(k)$ denotes discrete-time acoustic echo signal. Echo is estimated by using adaptive filter H and then subtracted from noisy speech signal (see Fig. 2). As a result, the echo compensated signal can be expressed by

$$e(k) = s(k) + n(k) + d(k) - d'(k) = s(k) + n(k) + b(k). \quad (16)$$

The parameter $b(k)$ is so-called residual echo and can be additionally attenuated by noise suppressor. In order to calculate weighting factors of the filter H , adaptive algorithms are used. Most of these algorithms are based on minimization of Mean Square Error (MSE) [20].

Echo control and noise suppression have been investigated separately for many years. Pure echo cancellation problem is beyond scope of this review, so we only focus on combined approaches. Presently most of hand-free devices consist of integrated echo canceller and noise suppressor. Joint systems can be composed into different structures [21, 22]. The most popular structure is composition of the two cascaded, optimal filters where Acoustic Echo Canceller (AEC) precedes the Noise Reduction (NR) [11]. Furthermore, in such structures, noise reduction rules are often modified for suppressing not only background noise but also residual echo [23].

3. IMPROVEMENT OF SPECTRAL WEIGHTING TECHNIQUES

Some improved spectral weighting methods have been presented in [4–7]. The aim of these approaches is to eliminate musical tones and decrease distortions of useful signal by more accurate estimation of signal to noise ratio. In that way, so-called decision-directed approach [4, 5] is used to estimate a priori SNR:

$$\text{SNR}_{\text{prior},m}(\omega) = \alpha \frac{|\hat{S}_{m-1}(\omega)|^2}{R_{m,m}(\omega)} + (1 - \alpha) \max \{ \text{SNR}_{\text{post},m}(\omega), 0 \}, \quad (17)$$

where $0 < \alpha < 1$ corresponds to mixing factor between previous and present estimate of the SNR and m denotes frame index. High values of α are required to reduce musical

effect. U
Moreover
observed.
musical n
can be al
(2) param
obtained

More
propose to
of the fix
previous f

High value
distortions
of param
(but are no
same resid
slightly lo

The a
transient p
distance m

Parameter
this measu
sequence (
the variabl
parameter
Unfortunat
though in s
highly redu
periods or
of the para

Modern
choacoustic

effect. Unfortunately, if α is too high, distortions of speech signal are introduced. Moreover, unnatural sounds at the beginning and the end of voice activity intervals are observed. Parameter α is usually set to a value of 0.98 to achieve trade-off between musical noise and acceptable distortion [7]. It is proper to add that signal-to-noise ratio can be also used for time-varying modification of the generalized spectral subtraction (2) parameters λ , α . In that way, Nonlinear Spectral Subtraction (NSS) algorithm is obtained [24].

More advanced solutions of these problems have been presented in [6]. The authors propose to introduce in formula (17) a time-frequency varying parameter $\alpha_m(\omega)$ instead of the fixed one. First approach tends to link parameter α with SNR estimated in previous frame:

$$\alpha_m(\omega) = a_{\max} - \lambda \frac{\text{SNR}_{\text{prior},m-1}(\omega)}{1 + \text{SNR}_{\text{prior},m-1}(\omega)}. \quad (18)$$

High values of SNR give the values of parameter $\alpha_m(\omega)$ close to $\alpha_{\max} - \lambda$. Hence, speech distortions are reduced in voice activity periods. Otherwise, when SNR is low, a value of parameter $\alpha_m(\omega)$ is close to $\alpha_{\max}$. In this case, musical tones are strongly attenuated (but are not completely removed). As a result, proposed method is characterized by the same residual noise during no vocal periods as the method with a constant and only slightly lower speech distortions.

The alternative solution proposed in [6] tends to decrease parameter α during transient periods of noisy speech spectrum. Spectral changes are traced using spectral distance measure calculated between current and previous processing frame:

$$\Delta Y_m = \left[\frac{2}{\pi} \int_0^{\pi/2} (|Y_m(\omega)| - |Y_{m-1}(\omega)|)^2 d\omega \right]^{1/2}. \quad (19)$$

Parameter α is obtained by comparing current distance with smoothed version of this measure $\Delta \tilde{Y}_m$ updated in no vocal activity periods only. Due to fluctuations in sequence (19), the result of this comparison is rather erratic. In practical realizations, the variable α should be smoothed using an exponential averaging. In this case, the parameter α varies from 0 to 1, so completely musical noise elimination is possible. Unfortunately, the solution based on the spectral changes tracing is not perfect. Although in silence intervals, musical noises are eliminated and speech distortions are highly reduced, disappearing musical noise can be heard at the end of voice activity periods or during transitions of speech. This effect is caused by long-time adaptation of the parameter α .

4. PSYCHOACOUSTIC PRINCIPLES

Modern speech enhancement and speech coding systems, usually exploit the psychoacoustic models [14]. These approaches are based on the fact, that processed signal

often contains some irrelevant information, which is not perceived by human ear. For example, in the speech compression, omitting irrelevant information tends to optimize coding efficiency [25, 26]. On the other hand, exploiting psychoacoustic in the noise reduction systems tends to decrease speech distortions. This section presents several psychoacoustic principles, including absolute threshold of hearing, critical band frequency analysis, simultaneous masking, the spread of masking along the basilar membrane and temporal masking [14].

The frequency-to-place transformation occurs in cochlea (inner ear), along basilar membrane. Specific regions of the cochlea contain neural receptors, tuned to different frequencies. Generalized auditory model is based on a basilar membrane viewed as a filter bank consisting of highly overlapping band-pass filters. The magnitude responses are asymmetric and dependent on signal level. This comes from the fact that neural receptors are non-uniformly spaced along basilar membrane. The critical bandwidth is a function of frequency that quantifies the cochlear filter pass-bands and characterizes frequency selectivity of human ear (see Fig. 3 (left)). For an average listener, this measure is usually approximated by:

$$BW_c(f) = 25 + 75 \left[1 + 1.4 (f/1000)^2 \right]^{0.69} \quad [\text{Hz}], \quad (20)$$

where f is frequency in Hertz. A distance of one critical band is commonly referred to as one Bark in the literature.

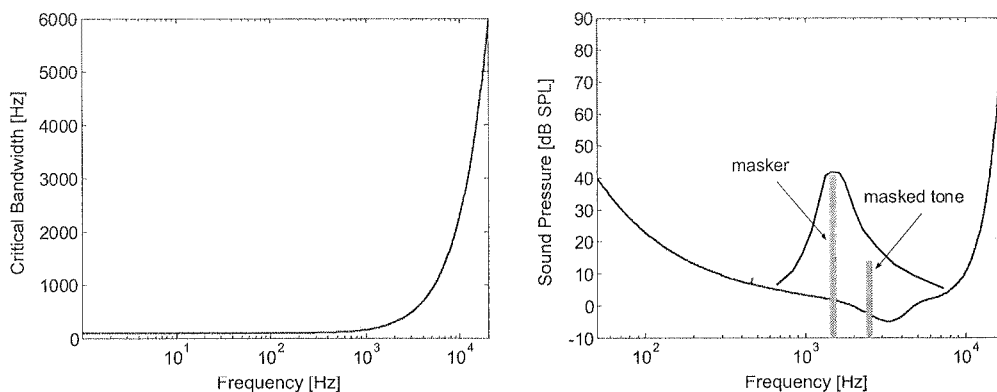


Fig. 3. Psychoacoustic principles: critical bandwidth (left) and an example of simultaneous masking — dotted line is an absolute threshold of hearing in noiseless environment.

Rys. 3. Podstawy psychoakustyki: szerokość pasma krytycznego w funkcji częstotliwości (z lewej) oraz przykład zjawiska maskowania równoczesnego — linia kropkowana oznacza absolutny próg słyszenia w ciszy.

The absolute threshold of hearing is defined by the power level of a pure tone such that a listener in a noiseless environment can detect it. In other words, if tone lies below the threshold, it is inaudible. This measure, typically expressed in terms of

dB Sound
of more
function:

$$T_q(f)$$

where f i

Two
temporal
where a w
Fig. 3 (rig
and simul
receive on
enough to
thresholds

Tempo
within a sl
stronger so
time of bas
occurs both
precedes m
Both phen
coding syst
ment only
the most of
ing function
performed

- frequen
- convolu
- subtract
- renorma

5. PSY

All prev
noise in the
gorithms sho
use only ma
speech disto
other hand, r
directly. The

dB Sound Pressure Level (SPL), depends on frequency and covers a dynamic range of more than 60dB. The absolute threshold of hearing is well approximated by the function:

$$T_q(f) = 3.64(f/1000)^{-0.8} - 6.5e^{-0.6(f/1000-3.3)^2} + 10^{-3}(f/1000)^4 \quad [\text{dB SPL}], \quad (21)$$

where f is frequency in Hertz.

Two kinds of perceptual masking are distinguished: simultaneous masking and temporal masking. Simultaneous masking refers to frequency domain phenomenon where a weaker signal can be inaudible because of presence another, stronger signal (see Fig. 3 (right)). If specific region of basilar membrane is stimulated by stronger signal and simultaneously, weaker signal stimulates the same region, the neural receptors receive only stronger signal. Hence, this phenomenon occurs if both signals are close enough to each other in frequency. In order to extract inaudible parts of signal masking thresholds are estimated.

Temporal masking refers to time domain phenomenon, where two sounds appear within a short time-interval. Depending on the individual sound pressure levels, the stronger sound may mask weaker one. Temporal masking is connected with a response time of basilar membrane, which is stimulated by acoustic impulse. This phenomenon occurs both prior to masker appearing as well as after its removing. If masked tone precedes masker we talk about pre-masking, otherwise we talk about post-masking. Both phenomena (temporal and simultaneous masking) are recently exploited in audio coding systems to minimize a bit-rate of the coder [27]. However, in speech enhancement only simultaneous masking is used (i.e. to discriminate audible distortions). In the most of systems an excitation along the basilar membrane is modeled by a spreading function and a corresponding masking threshold. Calculation of this threshold is performed with the following steps [28, 29]:

- frequency analysis along a critical bands,
- convolution with a spreading function,
- subtraction of a relative threshold offset,
- renormalization and comparison with the absolute threshold of hearing.

5. PSYCHOACOUSTIC MODIFICATION OF CONVENTIONAL NOISE SUPPRESSION RULES

All previously presented approaches only partially solve the problem of residual noise in the speech enhancement systems. Experiments performed with different algorithms show that there is a need for a further improvement. Conventional methods use only mathematical or experimental circumstances to reduce residual noise and speech distortions. In those cases, psychoacoustic can be used only indirectly. On the other hand, modern speech enhancement systems exploit the psychoacoustic principles directly. The simplest versions of these systems are based on a conventional spectral

weighting rules and simultaneous masking phenomenon. In these systems, weighting rules are modified to attenuate the signal at those frequencies where the noise is not completely masked by the speech. Sometimes, only post-processing after the implementation of classical suppression rules is used to eliminate the musical noise. This technique exploits the masking thresholds to identification of musical tones in output spectrum. More advanced approaches use psychoacoustic principles to define just new suppression rules. In this section, we present only these methods, which are based on the psychoacoustic modification of the conventional weighting rules or post-processing techniques.

5.1. PSYCHOACOUSTIC MODIFICATION OF SPECTRAL SUBTRACTION METHOD

A perfect example of the psychoacoustic modification of weighting rule is approach, which has been proposed in paper [8]. Main distinction between previously presented methods (see Section 3) and this system is direct using masking properties to estimate the parameters of generalized suppression rule. It is difficult to eliminate residual noise without introducing speech distortion, especially at very low SNR. Inaccurate adaptation of weighting rule parameters tends to situation where the enhanced speech can be even more disturbed (less intelligible) for a human listener than the original noisy speech signal. Moreover, some kinds of artifacts such as musical noise are perceived, again. Nevertheless, the authors prove that it is possible to minimize this problem by introducing knowledge on human perception in the enhancement system.

As mentioned earlier, the algorithm is based on the generalized spectral subtraction rule, but it is a possibility to use other weighting rules. In this case, gain function has been rewritten in the following form:

$$G(\omega) = \begin{cases} \left[1 - \alpha \left(\frac{|N(\omega)|}{|Y(\omega)|} \right)^\gamma \right]^{1/\gamma}, & \text{if } \left(\frac{|N(\omega)|}{|Y(\omega)|} \right)^\gamma < \frac{1}{\alpha + \beta} \\ \left[\beta \left(\frac{|N(\omega)|}{|Y(\omega)|} \right)^\gamma \right]^{1/\gamma}, & \text{otherwise,} \end{cases} \quad (22)$$

where α , β and γ are subtraction parameters, which control the tradeoff between noise suppression, residual noise and speech distortion. Parameter α determines amount of noise reduction, but the setting this factor to high values leads to an increased audible distortion. A level of background noise is controlled by parameter β . An increasing this factor can be very helpful in a masking residual noise and an elimination of musical tones. Unfortunately, increased level of background noise remains in the enhanced speech. Exponent γ determines the sharpness of the transition from the maximum value of gain function to its minimum value. This factor is usually set to one or two. The choice for value of γ is not as important as that of α and β . However, it is known that for $\gamma = 2$ less speech distortion and higher level of background noise is obtained.

Algorithms with fixed weighting rule parameters are unable to adapt to varying noise levels and characteristics, because at low SNR, it is impossible to simultane-

Fig. 4. BL

Rys. 4. S

In the
fixed to co
if the mas
values, bec
However,
maximal v
the subtra

weighting
noise is not
the imple-
noise. This
in output
the just new
based on
processing

HOD

rule is ap-
previously
properties
eliminate
SNR. Inac-
enhanced
r than the
sical noise
imize this
nt system.
subtraction
nction has

(22)

between noise
amount of
used audible
reasing this
of musical
e enhanced
maximum
one or two.
it is known
s obtained.
to varying
simultane-

ously minimize residual noise and speech distortion. Furthermore, the time-varying adaptation of the parameters is not easy task. Although it is possible to use SNR based solution [24], we will show that psychoacoustic approach is more efficient. In the presented method, authors have proposed to exploit a masking threshold to adapt subtraction parameters. This approach allows obtaining the best tradeoff between the amount of noise reduction, the speech distortion and level of residual noise. Because an average listener tolerates additive noise as long as it remains below masking threshold, there is no need for a complete removal of the noise, especially in speech presence. The speech signal should be strongly modified only at those frequencies where the noise is not completely masked by the speech. In other words, if a spectral component lies below the threshold the attenuation is not needed. As a result, the speech distortions can be decreased. Because, in the presented enhanced scheme only noisy signal is available, the masking threshold of the clean speech $T(\omega)$ must be estimated. This is done with simple spectral subtraction scheme where a preliminary spectral weighting is performed to obtain rough estimate of clean speech spectrum, and from this preliminary spectrum, masking threshold is calculated. The simplest scheme of the proposed algorithm is given in Fig. 4.

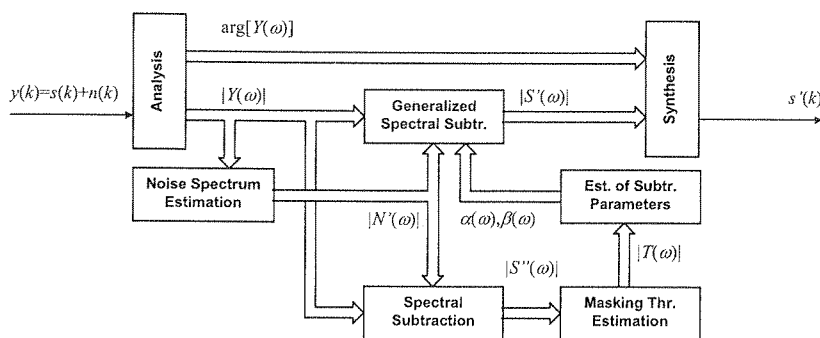


Fig. 4. Block diagram of speech enhancement system with psychoacoustic modification of generalized spectral subtraction rule.

Rys. 4. Schemat blokowy systemu uzdatniania sygnału mowy z psychoakustycznie zmodyfikowaną uogólnioną regułą odejmowania widmowego.

In the presented algorithm only parameters α , β are adapted, factor γ is kept fixed to constant value $\gamma = 2$. The adaptation is based on the following consideration: if the masking threshold is high the subtraction parameters are set to their minimal values, because residual noise will be masked by speech and attenuation is unnecessary. However, if the masking threshold is low, the subtraction parameters are set to their maximal values, due to annoying nature of residual noise. Therefore, the adaptation of the subtraction parameters is performed with the following relations:

$$\alpha(\omega) = F[\alpha_{\min}, \alpha_{\max}, T(\omega)], \quad (23)$$

$$\beta(\omega) = F[\beta_{\min}, \beta_{\max}, T(\omega)], \quad (24)$$

where the function F can be expressed as follow:

$$F[x, y, z] = x + (y - x) \frac{z_{\max} - z}{z_{\max} - z_{\min}}. \quad (25)$$

In the above equations, factors $\alpha_{\min}$, $\beta_{\min}$ and $\alpha_{\max}$, $\beta_{\max}$ are predefined minimal and maximal values of the subtraction parameters α , β respectively. Due to fluctuations in the sequences $\alpha(\omega)$ and $\beta(\omega)$, there is a possibility of discontinuities in the gain function. Therefore, the subtraction parameters should be additionally smoothed (for example by exponential averaging).

Main advantage of the psychoacoustically modified algorithm over conventional SNR-based methods lies in the use of the masking threshold. It clearly comes from the following observations:

- the estimation of the masking threshold in very noisy environment is more accurate than the SNR estimation,
- the masking threshold has smoother evolution than the SNR. It is important for suppression of residual noise,
- the adaptation based on the masking threshold is much better correlated with human auditory system than exploiting the SNR.

Moreover, presented approach shows that time-varying psychoacoustic modification of subtraction parameters gives better results than methods with fixed parameters. Although the noise attenuation, defined as the mean ratio between the input noise power and output noise power, is similar in both methods (see Fig. 5 (left)), the Itakura-Saito spectral distance measure clearly shows significant improvement of the psychoacoustically motivated approach (see Fig. 5 (right)). For more detailed performance evaluation, see [8].

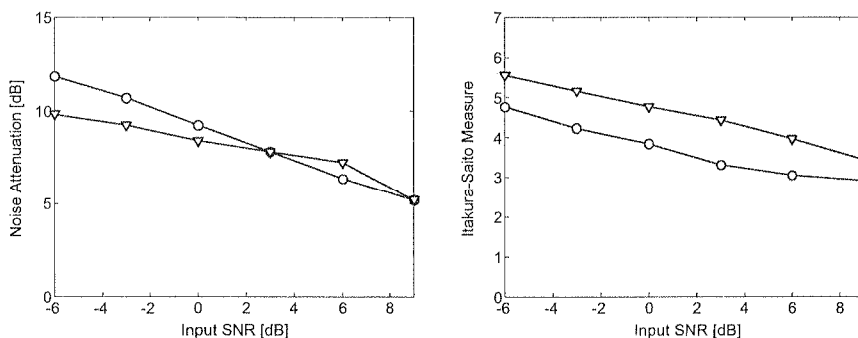


Fig. 5. Performance comparison of the psychoacoustically modified spectral subtraction rule (circles) and conventional spectral subtraction rule with fixed parameters (triangles).

Rys. 5. Porównanie właściwości zmodyfikowanej psychoakustycznie reguły odejmowania widmowego (koła) oraz reguły konwencjonalnej z optymalnie ustalonymi parametrami (trójkąty).

And
speech
tones
method,
mentation
This sol
achieve
new auc
The mai
noise su
proper t
to prese
and their

In c
tone mas
offset fac
ever, in
In fact,
backgrou
Because,
putting si

$$y(k)=s(k)$$

Fig. 6. Blo

Rys. 6. Sc

The m
Fig. 6). T

(24)

5.2. ELIMINATION OF THE MUSICAL TONES WITH POST-PROCESSING TECHNIQUES

(25)

imal and
 ctuations
 the gain
 thed (for
 ventional
 nes from
 accurate
 ortant for
 th human
 dification
 eters. Al-
 ise power
 ura–Saito
 oacousti-
 valuation,

Another approach to simultaneous elimination of musical noise and reduction of speech distortion exploits post-processing technique [9]. As mentioned earlier, musical tones are visible in output spectrum as narrow peaks of random frequencies. In this method, the masking thresholds are used to identification of these tones after implementation of conventional noise suppression algorithm, especially spectral subtraction. This solution is quite different from previously presented techniques but it tends to achieve similar results. However, the work [9] only concentrates on detecting any new audible spectral components at the output of the speech enhancement system. The main advantage of the proposed scheme is that it is independent of the selected noise suppression algorithm. There is no need to modify spectral weighting rule. It is proper to add that, there are some post-processing algorithms [30], which are similar to presented one. Nevertheless, those methods do not use psychoacoustic principles and their identification of musical tones is much more complicated and less efficient.

In calculation of the masking threshold, usually two scenarios are distinguished: tone masking noise or noise masking tone. Depending on the situation, corresponding offset factor is applied to the summarized excitation level of basilar membrane. However, in presented method, only second scenario (noise masking tone) is considered. In fact, it is correct during speech pauses where the musical tones are masked by background noise. Otherwise, an estimation of masking threshold may be inaccurate. Because, musical tones are especially annoying during speech pauses and due to computing simplification, authors have decided to exploit only noise maskers.

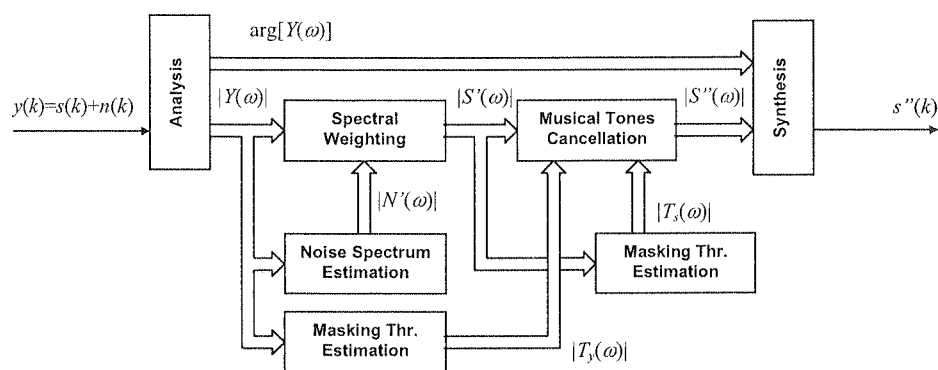


Fig. 6. Block diagram of speech enhancement system with post-processing based on masking properties for cancellation of musical tones.

Rys. 6. Schemat blokowy systemu uzdatniania sygnału mowy, opartego na metodzie usuwania tonów muzycznych w sygnale uzdatnionym, wykorzystującej zjawisko maskowania.

The masking thresholds are estimated at both the input and output signals (see Fig. 6). The audible components lie above these thresholds. If any spectral compo-

le (circles)
 widmowego
 y).

nent is only present at the output, it is a potential musical tone, which should be post-processed. Due to estimation errors, there is a possibility of qualifying a speech component as a musical tone. Therefore, in order to avoid wrong identification of components a short-time stationarity and bandwidth criteria are additionally applied. In the presented scheme, it has been assumed that musical noise bandwidth is not more than approximately 300Hz. In order to accomplish the stationarity criterion, present and previous values of the gain factors are compared with the predefined threshold. If one of these values is less than threshold, another one probably corresponds to musical noise component. After proper identification, musical components are set to the spectral floor.

6. CONCLUSIONS

In this review, conventional methods aimed at elimination of residual noise and their psychoacoustic modifications have been presented. Described methods have been classified into two groups: conventional non-psychoacoustically improved approaches and conventional schemes with psychoacoustic modification. The main advantage of perceptual methods over conventional ones lays in the use of the masking threshold instead SNR criterion to eliminate residual noise. Such approaches are much better correlated with human auditory system.

However, it is a possibility of further improvement. Presented approaches are not optimal in perceptual sense because psychoacoustic modification of weighting rules parameters has been obtained heuristically. Moreover, these systems exploit merely simultaneous masking properties. In Part II of the review, we will cover recent systems with the new psychoacoustically optimized spectral weighting rules.

7. ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewer for their helpful comments and suggestions.

8. REFERENCES

1. K. Kroschel: *Noise reduction: an old problem of actual concern*. 8th International Symposium on Sound Engineering and Mastering (ISSEM'99), Gdańsk, 1999, pp. 3–8
2. R. Martin: *Spectral subtraction based on minimum statistic*. VII European Signal Processing Conference (EUSIPCO'94), Edinburg, 1994, pp. 1182–1185
3. S. F. Boll: *Suppression of acoustic noise in speech using spectral subtraction*. IEEE Transactions on Acoustic Speech and Signal Processing, 1979, vol. ASSP-27, no. 2, pp. 113–120
4. Y. Ephraim, D. Malah: *Speech enhancement using a minimum mean-square error short-time spectral amplitude estimator*. IEEE Transactions on Acoustic Speech and Signal Processing, 1984, vol. ASSP-32, no. 6, pp. 1109–1121

5. Y. E. *amplit*
33, no.
6. C. B. *Confer*
1999,
7. O. C. *suppre*
8. N. V. *system*
9. T. H. *choaco*
1997, p
10. D. E. *based o*
no. 6, p
11. K. B. *cych wy*
Politech
12. S. G. *bined a*
Processi
13. A. A. *reductio*
Exhibitio
Russian)
14. E. Z. w
15. A. B. o
Internati
16. R. M. a
statistic
17. A. A. k
ner filter
SIPCO'9
18. J. L. i m
ceedings
19. K. B. i e
hands-fre
t. 44, z. 4
20. S. H. a y
21. B. A. y a
processing
and Signa
22. R. B. o u
and echo
cessing, 2
23. S. G. u s
on residu
U.K., 199

5. Y. Ephraim, D. Malah: *Speech enhancement using a minimum mean-square log-spectral amplitude estimator*. IEEE Transactions on Acoustic Speech and Signal Processing, 1985, vol. ASSP-33, no. 2, pp. 443–445
6. C. Beaupré, P. Scalart: *Noise reduction using perceptual spectral change*. 6th European Conference on Speech Communication and Technology (EUROSPEECH'99), Budapest, Hungary, 1999, pp. 2543–2546
7. O. Cappé: *Elimination of the musical noise phenomenon with the Ephraim and Malah noise suppressor*. IEEE Transactions on Speech and Audio Processing, 1994, vol. 2, pp. 345–349
8. N. Virag: *Single channel speech enhancement based on masking properties of the human auditory system*. IEEE Transactions on Speech and Audio Processing, 1999, vol. 7, no. 2, pp. 126–137
9. T. Haulick, K. Linhard, P. Schrogmeier: *Residual noise suppression using psychoacoustic criteria*. 5th European Conference on Speech Communication and Technology, Greece, 1997, pp. 1395–1398
10. D. E. Tsoukalas, J. N. Mourjopoulos, G. Kokkinakis: *Speech enhancement based on audible noise suppression*. IEEE Transactions on Speech and Audio Processing, 1997, vol. 5, no. 6, pp. 497–514
11. K. Bielawski: *System eliminacji echa akustycznego i szumów w urządzeniach głośnomówiących wykorzystujący modele psychoakustyczne*. Rozprawa doktorska (promotor prof. A.A. Petrovsky), Politechnika Warszawska, 2001
12. S. Gustafsson, R. Martin, P. Jax, P. Vary: *A psychoacoustic approach to combined acoustic echo cancellation and noise reduction*. IEEE Transactions on Speech and Audio Processing, 2002, vol. 10, no. 5, pp. 245–256
13. A. A. Petrovsky, A. E. Anoshenko: *Combined system for echo cancellation and noise reduction in frequency domain with psychoacoustic motivation*. 2nd International Conference and Exhibition, „Digital Signal Processing and its Applications”, 1999, Moscow, Russia, pp. 166–169 (in Russian)
14. E. Zwicker, H. Fasel: *Psychoacoustic. Facts and models*. Berlin, Springer Verlag, 1990
15. A. Borowicz, A. A. Petrovsky: *The comparative study of voice activity detectors*. VII International Conference of Modern Telecommunication Systems, Naroch, Belarus, 2002, pp. 148–152
16. R. Martin: *Noise power spectral density estimation based on optimal smoothing and minimum statistic*. IEEE Transactions on Speech and Audio Processing, 2001, vol. 9, no. 5, pp. 504–512
17. A. Akbari Azirani, R. Lebouquin, G. Faucon: *Speech enhancement using a Wiener filtering under signal presence uncertainty*. VIII European Signal Processing Conference (EU-SIPCO'96), Trieste, Italy, 1996, pp. 971–974
18. J. Lim, A. Oppenheim: *Enhancement and Bandwidth Compression of Noisy Speech*. Proceedings of the IEEE, 1979, vol. 67, no. 12, pp. 1586–1604
19. K. Bielawski, A. A. Petrovsky: *Joint system for acoustic echo and noise control in hands-free communication devices*. Kwartalnik Elektroniki i Telekomunikacji, PWN Warszawa, 1998, t. 44, z. 4, pp. 115–136
20. S. Haykin: *Adaptive filter theory*. Third edition, Prentice-Hall, 1996
21. B. Ayad, G. Faucon, R. Bouquin-Jeannes: *Optimization of a noise reduction pre-processing in an acoustic echo and noise controller*. International Conference on Acoustic, Speech and Signal Processing (ICASSP'96), Atlanta, USA, 1996, pp. 953–956
22. R. Bouquin Jeannes, P. Scalart, G. Faucon C. Beaupré: *Combined noise and echo reduction in hands-free systems: a survey*. IEEE Transactions on Speech and Audio Processing, 2001, vol. 9, no. 8, pp. 808–820
23. S. Gustafsson, R. Martin: *Combined acoustic echo control and noise reduction based on residual echo estimation*. International Workshop on Acoustic Echo and Noise Control, London, U.K., 1997, pp. 160–163

24. P. Lockwood, J. Boudy: *Experiments with a nonlinear spectral subtractor (NSS)*. Speech Communication, 1992, vol. 11, pp. 215–228
25. T. Painter, A. Spanias: *Perceptual coding of digital audio*. Proceedings of the IEEE, 2000, vol. 88, no. 4, pp. 451–513
26. P. Noll: *Digital audio coding for visual communications*. Proceedings of the IEEE, 1995, vol. 83, no. 6, pp. 925–943
27. A. I. Petrovsky, D. Krahe, A. A. Petrovsky: *Real-time wavelet packet-based low bit-rate audio coding on a dynamic reconfiguration system*. AES Convention paper 5778, presented at the 114th Convention, 2003, Amsterdam, The Netherlands, p. 22–27
28. J. D. Johnston: *Transform coding of audio signal using perceptual noise criteria*. IEEE Journal on Selected Areas in Communications, 1988, vol. 6, no. 2, pp. 314–323
29. ISO/IEC 11172-3, Information technology — Coding of moving pictures and associated audio for digital storage media at up to about 1,5 Mbit/s, Part 3: Audio, 1993
30. Z. Goh, K. Tan, B. T. G. Tan: *Postprocessing method for suppressing musical noise generated by spectral subtraction*. IEEE Transactions on Speech and Audio Processing, 1998, vol. 6, no. 3, pp. 287–292

A. BOROWICZ, A.A. PETROVSKY

METODY MOTYWOWANE PERCEPTUALNIE W UZDATNIANIU SYGNAŁU MOWY CZĘŚĆ I: ELIMINACJA ZJAWISKA TONÓW MUZYCZNYCH

Streszczenie

Przedstawiono problematykę związaną z projektowaniem oraz użytkowaniem systemów transmisji mowy opartych na metodzie wag widmowych. Omówiono rolę takich zjawisk psychoakustycznych jak maskowanie, pasma krytyczne i absolutny próg słyszenia w procesie uzdatniania sygnału mowy. Dokonano zestawienia klasycznych metod uzdatniania sygnału mowy ukierunkowanych na eliminację tonów muzycznych z systemami opartymi na psychoakustycznej modyfikacji metody wag widmowych. Porównano wyniki zastosowań rozwiązań psychoakustycznych pod kątem możliwości tłumienia szumu środowiskowego i zapobiegania generacji zniekształceń sygnału mowy. Uwypuklono korzyści wynikające z zastosowania podstawowych właściwości psychoakustycznych systemu słuchowego człowieka w metodach wag widmowych, jednocześnie wskazano ograniczenia tych metod i możliwości ich optymalizacji.

Słowa kluczowe: redukcja szumów, eliminacja echa, psychoakustyka.

Praca finansowana z grantu W/WI/3/02.

Perceptually Motivated Approaches to Speech Enhancement — Part II: Psychoacoustic Optimization of Spectral Weighting Rules

ADAM BOROWICZ, ALEXANDER A. PETROVSKY

*Wydział Informatyki, Politechnika Białostocka
ul. Wiejska 45A, 15-351 Białystok
E-mail: borowicz@ii.pb.bialystok.pl*

*Otrzymano 2003.06.11
Autoryzowano 2003.11.14*

This paper focuses on the class of speech enhancement systems, which capitalize on psychoacoustic properties of the human ear. More advanced psychoacoustically motivated spectral weighting rules are described. Presented systems are analyzed and classified according to their similarity with a human auditory model. Especially, a comparison of improvements in musical noise cancellation and increasing speech intelligibility is performed. Moreover, advantages of the perceptual approaches over conventional ones are focused. Finally, perspectives of integrated psychoacoustically motivated speech enhancement and coding systems are discussed. Paper shows that integration of subband coder with speech enhancement system based on non-uniformly spaced filter bank leads to most promising combined scheme.

Keywords: noise reduction, echo cancellation, psychoacoustic, speech coding.

1. INTRODUCTION

The problem of enhancing speech degraded by uncorrelated additive noise has been discussed for many years. Unfortunately, there is still no satisfactory solution available. However, such solutions are greatly needed in modern voice communication devices, especially hands-free telephones. This problem is also present on the field of speech coding where noise suppressor is used for improving a quality of the coding speech.

Spectral subtraction [1, 2] is a simple and efficient technique of noise reduction but major disadvantage of this approach is the residual noise known as musical tones and

the distortion of the speech signal. Musical tones produced by conventional spectral weighting algorithms are composed of sinusoidal components with random frequencies. This comes from the fact that the magnitude of the short-time spectrum contains strong fluctuations in noisy areas. Many modifications of classic algorithms have been proposed to reduce this undesirable effect [3–6]. However, in recent years, we observe growing tendency to use perceptual models [7] for problem solution. Although perception of audio signals is not fully understood yet, there are some psychoacoustic principles, which are successfully used in wideband audio coding [8, 9]. Simplest versions of perceptually motivated speech enhancement methods [10, 11] are based only on a psychoacoustic modification of the conventional weighting rule parameters. Examples of these systems have been described in Part I of the review [12].

From the perceptual point of view previously presented solutions are sub-optimal. There is a possibility of further improvements by complete exploitation of the human auditory system properties. In this paper, we present the most advanced, psychoacoustically optimized spectral weighting rules. In order to focus advantages of psychoacoustically motivated methods over conventional ones, we try to classify different approaches according to their similarity with a human auditory model. Simultaneously, we continue analysis of the improvements in musical noise cancellation and increasing speech intelligibility started in Part I [12] of the review.

The paper is organized as follows. In Section 2.1 we describe audible noise suppression technique aimed at minimization of the difference between audible spectrum of the noisy and clean speech signals. In Section 2.2 subband noise suppression approach is described. In Section 2.3 is presented another method based on attenuation of audible noise and speech distortions. In Section 3, we include a discussion on perspectives of integrated, psychoacoustically motivated speech enhancement and coding system. In that way, we tend to obtain the best combination of noise suppressor and speech coder. Finally, in Section 4 we summarize the paper.

2. PSYCHOACOUSTICALLY MOTIVATED NOISE SUPPRESSION RULES

Sometimes, a simple modification of weighting rule parameters is not sufficient and a further improvement is needed. In the next sections, we will present some of novel psychoacoustically motivated speech enhancement systems. In this case the term: 'psychoacoustically motivated' stands for a derivation of exactly new weighting rule by using properties of the human auditory system. In contrast to perceptually sub-optimal methods [10, 11], these ones do not use modified conventional spectral weighting rules, but exploit masking properties in direct manner to derive new rules. All those methods are based on the same simple framework, which has been presented on Fig. 1. The only difference between them is a various definition of the gain function (psychoacoustic weighting) and various realization of analysis/synthesis block. Although whole system can be implemented in time or frequency domain, the main scheme stays the same.

 $y(k)$

Fig. 1.

Rys. 1. S

Der
matical
speech
of the a
the noise
They we
derivatio
The
respecti

where R
 $T(\omega)$ is
as

The ma
is the fa
The pur
of the a
The mo
processi
and less

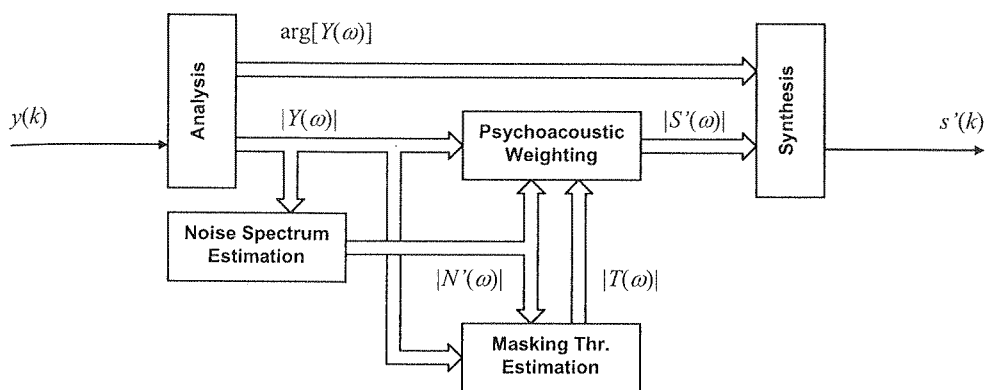


Fig. 1. Block diagram of speech enhancement system with psychoacoustically motivated suppression rule exploiting masking phenomenon.

Rys. 1. Schemat blokowy systemu uzdatniania sygnału mowy z motywowaną psychoakustycznie regułą tłumienia szumu, wykorzystującą zjawisko maskowania.

2.1. AUDIBLE NOISE SUPPRESSION

Derivations of conventional suppression rules are based on minimization of mathematical or statistical difference (e.g. mean square error), between a clean and enhanced speech spectrum. Tsoukalas et al. [13] developed a method aimed at minimization of the audible noise spectrum, defined as the difference between audible spectrum of the noisy and clean speech signals. In this case, perceptual difference is exploited. They were improving this approach for some years and finally extended and used it for derivation of a new optimal weighting rule that achieves audible noise suppression [14].

The audible spectra of the noisy and clean speech are defined as $A_y(\omega)$ and $A_s(\omega)$, respectively, using the following expressions

$$A_y(\omega) = \max \{R_{yy}(\omega), R_{TT}(\omega)\}, \quad (1)$$

$$A_s(\omega) = \max \{R_{ss}(\omega), R_{TT}(\omega)\}, \quad (2)$$

where $R_{yy}(\omega)$ and $R_{ss}(\omega)$ denote the power spectral densities of noisy and clean speech, $T(\omega)$ is a masking threshold. Therefore, the audible noise spectrum can be formulated as

$$A_d(\omega) = A_y(\omega) - A_s(\omega). \quad (3)$$

The main advantage of the above measure over the simple mathematical measures is the fact that it takes into the account only the perceptually significant differences. The pure spectral difference between clean and noisy speech is an overestimation of the audible noise spectrum (3) and includes components, which are not audible. The modification of the noisy speech spectrum to suppress the audible noise leads to processing improvements, because only selective frequency components are attenuated and less distortion is introduced into useful speech signal.

The main principle of the presented technique is to make the audible noise less than or equal to zero by proper modification of the noisy speech. By substituting $\hat{R}_{ss}(\omega)$ for $R_{yy}(\omega)$ into (1) we can express the minimization criterion as follow

$$0 \geq (A_d(\omega) = A_s(\omega) - A_s(\omega)), \quad (4)$$

$$0 \geq \max \{ \hat{R}_{ss}(\omega), T(\omega) \} - \max \{ R_{ss}(\omega), T(\omega) \}. \quad (5)$$

Theoretically, the enhanced speech power spectrum $\hat{R}_{ss}(\omega)$ can be obtained by the standard spectral weighting methods. However, a small flexibility of the classical gain functions is main disadvantage of these systems. It is an important problem even if psychoacoustic modification of the weighting rule parameters was performed, especially in very noisy conditions. In order to achieve more flexible gain control, a parametric nonlinear function was used instead other ones. This function is given by

$$G(\omega) = \frac{R_{yy}^{v(\omega)}(\omega)}{a^{v(\omega)}(\omega) + R_{yy}^{v(\omega)}(\omega)}, \quad (6)$$

where $a(\omega)$ and $v(\omega)$ are the time-frequency varying parameters (the time index is omitted for simplicity). The first parameter defines threshold below which spectral components are highly attenuated. The rate of this attenuation can be controlled by the second parameter. Note, that if both $a(\omega)$ and $v(\omega)$ are positive, the gain function (6) is always below than or equal to one. It is important to achieve an optimal spectral modification in the psychoacoustic sense. Therefore, there is a need for a proper derivation of the parameters, according to the minimization criterion. The effect of parameter $v(\omega)$ is not crucial to performance of the presented technique and a value of this factor may be considered constant for a specific frequency. Only parameter $a(\omega)$ should be adjusted during enhancement process. By combining (6) into (5) and solving inequality with respect to $a(\omega)$ and since $v(\omega)$ is positive, we obtain

$$a(\omega) \geq R_{yy}(\omega) \left[\frac{R_{yy}(\omega)}{\max \{ R_{ss}(\omega), T(\omega) \}} - 1 \right]^{1/v(\omega)}. \quad (7)$$

The presented system uses standard uniform spectral decomposition where each spectral component is modified using gain function in linear frequency domain. However, it is not desirable to estimate parameters $a(\omega)$ and $v(\omega)$ for every spectral component, due to non-linear frequency selectivity of the human ear. The perceptually significant frequency regions are defined using critical bands. Therefore, the parameters should be adjusted to the constant values over these regions. In the simplest version of the presented scheme, a value of parameter $a(\omega)$ for the specific critical band is obtained by selection of such a frequency within this band, that produces a maximum in (7):

$$a_b = R_{yy}(\omega_b) \left[\frac{R_{yy}(\omega_b)}{\max \{ R_{ss}(\omega_b), T(\omega_b) \}} - 1 \right]^{1/v_b}, \quad \omega_{lb} \leq \omega_b \leq \omega_{hb}. \quad (8)$$

The complete clean-speech power spectrum $R_{ss}(\omega)$ can be roughly estimated using simple spectral subtraction method. Unfortunately, a performance of such system is strongly dependent on a quality of clean-speech estimate. Therefore, in work [14], authors have proposed several solutions to achieve optimal psychoacoustic modification while only sparse clean signal components are known or can be easily estimated. These approaches are based on optimization of the clean speech estimation within subband regions by using only single value of the spectral components per critical band. On that way, number of parameters required for speech enhancement can be reduced. In both approaches it is assumed that the noise and speech are statistically independent, moreover, the masking threshold and background noise spectrum within each critical band are constant:

$$\begin{cases} R_{yy}(\omega) = R_{ss}(\omega) + R_{nn}(\omega) \\ T(\omega) = T_b \\ R_{nn}(\omega) = R_{nnb} \end{cases}, \quad \omega_{lb} \leq \omega \leq \omega_{hb}. \quad (9)$$

The first solution tends to estimate a_b' using the minimum speech power spectrum component $R_{ssb}^{\min}$ in the specific critical band. This parameter is usually selected from components lying above the masking threshold. However, it is possibility to estimate $R_{ssb}^{\min}$ using statistical methods. Under above assumptions (9), we can rewrite expression (8) in the following form

$$a_b' = [R_{nnb} + R_{ssb}^{\min}] \left[\frac{R_{nnb}}{R_{ssb}^{\min}} \right]^{1/\nu_b}, \quad (10)$$

where

$$R_{ssb}^{\min} = \min_{\omega_b} \{R_{ss}(\omega_b), \quad \omega_{lb} \leq \omega_b \leq \omega_{hb}\}. \quad (11)$$

Note that expression (10) is always greater than (8) and hence a_b' is solution, which satisfies minimization criterion (5). Similarly, in the second approach, clean speech power spectrum $R_{ss}(\omega)$ can be replaced by masking threshold $T(\omega)$.

$$a_b'' = [R_{nnb} + T_b] \left[\frac{R_{nnb}}{T_b} \right]^{1/\nu_b} \quad (12)$$

It is easy to see that a_b'' is also solution, which satisfies (5). An estimate of the clean-speech masking threshold $T(\omega)$ can be obtained from the noisy data. Unfortunately, at least rough estimate of the clean speech spectrum is necessary for threshold calculation. The simplest way to obtain this estimate is to use spectral subtraction method. Tsoukalas et al. [14] suggest that satisfactory estimate of the clean-speech masking threshold can be also obtained using an iterative procedure at some expense of computational efficiency. Each time, the noisy signal is passed through the filter (6) and better estimates of the clean speech and the masking threshold are obtained.

The audible noise suppression technique achieves significant improvements compared to algorithms based on the classical minimization criteria. It is especially evident in better speech intelligibility and smaller influence of SNR conditions to enhancement process. The main advantage of this method lies in more flexible modifying the noisy signal.

2.2. SUBBAND NOISE SUPPRESSION

As mentioned earlier, it may be uncomfortable to perform the perceptual speech enhancement using uniform spectral decomposition, due to non-linear frequency selectivity of the human ear. Note that in previous methods, although spectral weighting is applied to the uniformly spaced spectral components (obtained by STFT), the psychoacoustic parameters are calculated over critical bands. It is necessary to use the same non-uniform frequency scale in the whole system. On that way, some solutions have been developed where the spectral analysis and synthesis, are based on the human auditory system realized with the aid of filter banks or Discrete Wavelet Transform (DWT). A perfect example of such technique is system, which has been described in work [15–20]. Authors have proposed a new approach to the combined echo and noise reduction system. In this case, non-uniform DFT filter bank [21] has been exploited as a spectral analysis tool. For more detailed description of these techniques see [22]. Additionally, the algorithm for noise reduction used in this system is based on the psychoacoustically motivated suppression rule (6).

The non-uniform polyphase filter bank is a basis of the critical band operation system. Therefore, firstly we will present its framework. An effective structure of the uniform filter bank can be obtained by polyphase decomposition of a common prototype lowpass filter $H(e^{j\omega})$ and the Discrete Fourier Transform (DFT). The frequency response of this uniform bank is given by

$$H_m(e^{j\omega}) = H(e^{j(\omega - 2\pi m/M)}), \quad m = 0, \dots, M-1. \quad (13)$$

The polyphase uniform filter bank can be converted into filter bank with non-uniform frequency resolution through all-pass transformation [23]. It is based on replacing all delays with the following all-pass filter

$$H_{AP}(z) = \frac{z^{-1} - a}{1 - az^{-1}}, \quad (14)$$

where the parameter a is a frequency warping factor. The phase response of the all-pass filter H_{AP} is nonlinear function of frequency:

$$\varphi(\omega) = 2 \arctan \left(\frac{1-a}{1+a} \tan \frac{\omega}{2} \right). \quad (15)$$

By proper modification of the parameter, it is possible to get different frequency band distribution especially approximation of the Bark scale can be obtained. The frequency

response of
is given by

It is proper
ment system
(Fig. 2), fir
used for ech
suppression
dible noise
system wor
the time-sa
required for

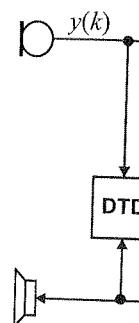


Fig. 2. Block d

Rys. 2. Schem

Presented
systems for a
in the superio
which implici
on the sub-b
faster converg
characteristic
the masking p
speech has be
threshold are
audible noise.

response of the resulting bank, which is now an infinite impulse response filter bank, is given by

$$H_m'(e^{j\omega}) = H(e^{j(\omega - \varphi(\omega) - 2\pi m/M)}). \quad (16)$$

It is proper to add that similar structures are widely used not only in speech enhancement systems but also in audio coding. In the presented speech enhancement scheme (Fig. 2), firstly, a transversal structure FIR filter with normalized LMS algorithm is used for echo cancellation. Then the echo compensated signal is forwarded to the noise suppression unit (AEC+NR configuration). The noise filtering is performed using audible noise suppression technique described in previous section. Because the whole system works in time domain, spectral weighting rule has been modified to work with the time-sampled bandlimited signals. In order to reduce the number of parameters required for speech enhancement, the approach (12) has been chosen.

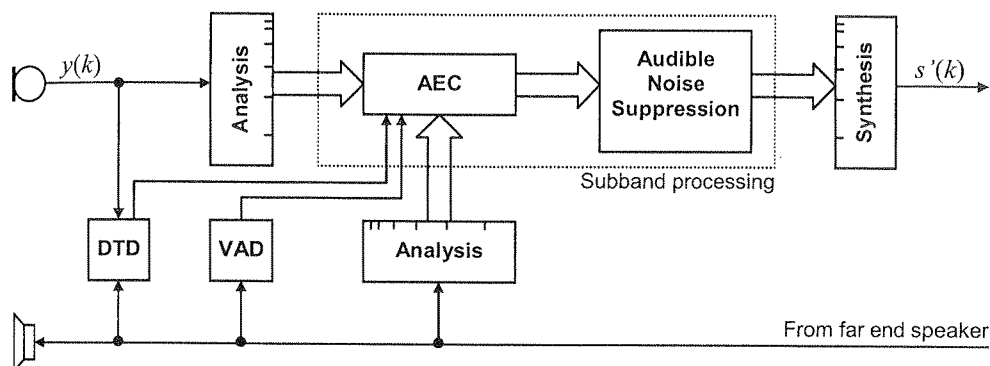


Fig. 2. Block diagram of subband processing system for acoustic echo cancellation and noise reduction.

Rys. 2. Schemat blokowy połączonego systemu eliminacji echa i tłumienia szumów akustycznych z przetwarzaniem podpasmowym.

Presented system is one of the most advanced psychoacoustically motivated joint systems for acoustic echo and noise control. The main advantage of this approach lies in the superior representation of the human auditory model. There are many benefits, which implicitly come from this fact. Because echo cancellation procedure is performed on the sub-band signal, the length of adaptive filters can be truncated. Moreover, faster convergence of LMS algorithm is obtained due to more flatter sub-band spectral characteristic compared to the wideband signal. Apart from this, due to exploiting the masking properties in the noise suppression rule an intelligibility of the enhanced speech has been increased. The residual noise components, which lay below masking threshold are not attenuated because are inaudible. This technique suppresses only audible noise. So there are no conditions for generating a musical noise.

2.3. INAUDIBLE NOISE DISTORTION RULE

Another psychoacoustically motivated algorithm for noise reduction has been presented in work [24]. The main goal of this approach is not a complete removal of the noise, but instead to only attenuate audible noise by constant factor. This comes from the fact that in some situations like a phone conversation the complete noise removal is not desirable. In these cases, the intervals of silence generated by a common noise suppressor may cause the impression of an interrupted transmission.

The new method is also based on the fact, that a human listener does not perceive any additive signal components as long as they lay completely below the masking threshold. In order to make a difference between desired signal and enhanced signal to be inaudible it is compared with the masking threshold. Similarly to previously described system also in this case an original method [24] has been extended to use it in the combined echo and noise reduction system [25, 26]. The presented system consists of a conventional normalized LMS echo canceller and a second frequency domain postfilter (see Fig. 3). The main goal of this postfilter is to attenuate both residual echo and noise.

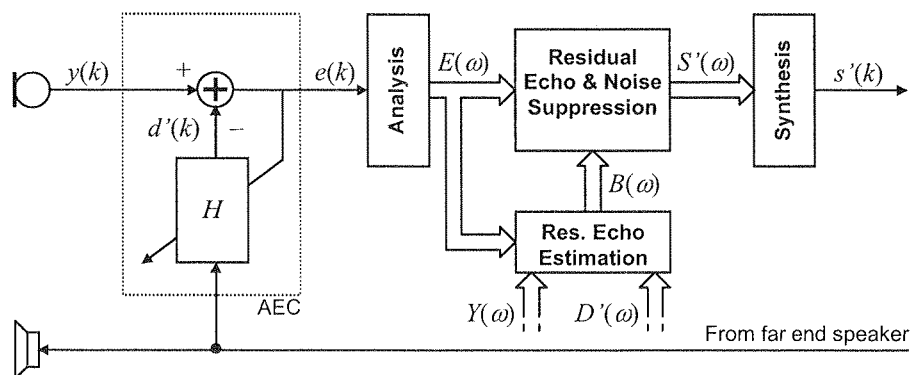


Fig. 3. Speech enhancement system with echo canceller and postfilter for residual echo and noise attenuation.

Rys. 3. System uzdatniania sygnału mowy wyposażony w moduł usuwania echa oraz filtr do jednoczesnego tłumienia szumów środowiskowych i echa szczątkowego.

The derivation of the extended suppression rule is briefly described here, although a more detailed description can be found in [25, 26]. In order to preserve background noise characteristic it is necessary to define preferable residual noise level ζ_n . Similarly, residual echo level ζ_b should be specified. Then the desired enhanced speech spectrum $\bar{S}(\omega)$ can be expressed as sum of clean speech, the attenuated residual echo and noise

$$\bar{S}(\omega) = S(\omega) + \zeta_b B(\omega) + \zeta_n N(\omega). \quad (17)$$

By using spectral weighting technique, we obtain the following estimate

$$\bar{S}'(\omega) = H(\omega)[S(\omega) + B(\omega) + N(\omega)], \quad (18)$$

where the gain function $0 \leq H(\omega) \leq 1$. Hence, difference between the desired and estimated spectrum can be written as

$$Q(\omega) = \bar{S}(\omega) - \bar{S}'(\omega). \quad (19)$$

If we assume that speech, noise and residual echo are statistically independent, we can express the power spectral density of $Q(\omega)$ in the following form

$$R_{qq}(\omega) = [1 - H(\omega)]^2 R_{ss}(\omega) + [\zeta_b - H(\omega)]^2 R_{bb}(\omega) + [\zeta_n - H(\omega)]^2 R_{nn}(\omega), \quad (20)$$

$$R_{qq}(\omega) = R_{qsqs}(\omega) + R_{qbqb}(\omega) + R_{qnqn}(\omega). \quad (21)$$

The particular components of (21) correspond to the speech distortion, the residual echo distortion and the noise distortion, respectively. In order to minimize the sum (21) in the perceptual sense it is appropriate to mask distortions to be inaudible for listener. Ideally, all distortions should be masked. However, in most real cases this demand cannot be accomplished due to fact that sometimes minimum of $R_{qq}(\omega)$ may be greater than masking threshold (distortions will be audible). From these reasons, authors propose to only mask the distortions of the residual echo and the noise, allowing a variable speech distortion. Hence, the minimization criterion is formulated as follow

$$R_{qbqb}(\omega) + R_{qnqn}(\omega) = T(\omega). \quad (22)$$

Solving this equation for $H(\omega)$ with the constraint $0 \leq H(\omega) \leq 1$ (and with some other simplifications) leads to the weighting rule

$$\tilde{H}^{\text{IND}}(\omega) = \min \left\{ 1, \sqrt{\frac{T(\omega)}{R_{bb}(\omega) + R_{nn}(\omega)}} + \frac{\zeta_n R_{nn}(\omega) + \zeta_b R_{bb}(\omega)}{R_{nn}(\omega) + R_{bb}(\omega)} \right\}, \quad (23)$$

where IND stands for Inaudible Noise Distortion. It is easy to see that if residual echo and noise are laid bellow the masking threshold, the term under square root is greater than one and the speech is not distorted because $\tilde{H}^{\text{IND}}(\omega) = 1$. Otherwise, residual noise and echo are optimally attenuated to be perceptually undistorted. In particular, when neither speech nor residual echo is present, the equation (23) can be reduced to $\tilde{H}^{\text{IND}}(\omega) = \zeta_n$. Therefore, since the parameter ζ_n is not too small the musical noise is not generated.

In order to calculate weighting factors (23) it is a need for efficient estimating the masking threshold and the power spectral densities of the noise and residual echo. The first two parameters can be obtained by using standard methods. In particular, an estimate of the masking threshold $T(\omega)$ is computed from rough estimate of the clean

speech spectrum obtained by conventional spectral weighting method. The third parameter $R_{bb}(\omega)$ can be estimated using coherence function [27] and equivalent transfer function [28] method.

In work [26], authors have compared this technique with the conventional (non-psychoacoustically motivated) combined echo and noise reduction system. They report the psychoacoustic approach is mostly preferred since it leads to more natural near end speech and to less annoying residual noise. The presented technique seems to be alternative to one described in the previous section. Also in this case masking threshold is used in direct manner to calculate the weighting coefficients. Both approaches have similar performance of the echo canceller, but ways to achieve such results were quite different. Similarly, the length of adaptive filter for echo canceller can be truncated. However, in this case, it comes from the fact that the echo cancellation problem is partially solved by optimized noise suppressor. Unfortunately, the IND rule is still computed only partly in the critical band domain. It is possible to achieve further improvements (i.e. faster convergence of LMS algorithm) by placing whole system in the same domain. Therefore, subband approach seems to be more advanced in psychoacoustic sense. On the other hand, there are more differences between both methods in objective measures. The comparison of both methods has been performed in [20]. Two measures have been used for performance evaluation: Noise Attenuation and distortion factor $SEGSNR_{s-s'}$. Results of comparison show that in case of IND rule, stronger noise attenuation and less speech distortion are achieved (see Fig. 4).

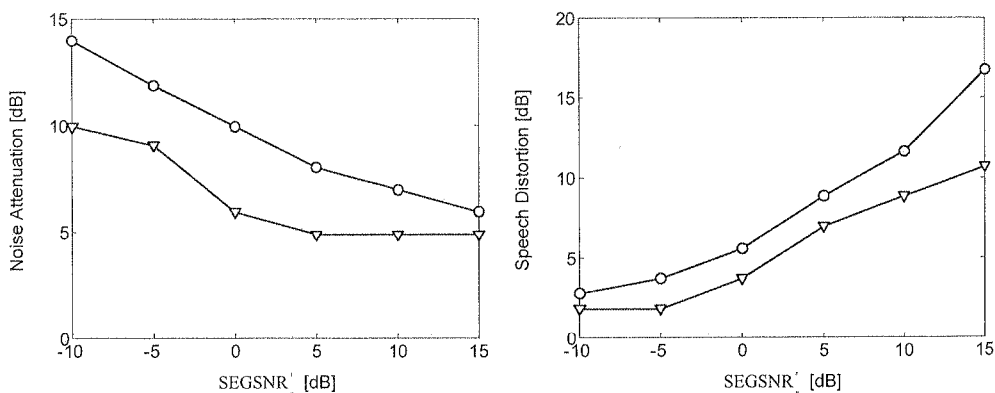


Fig. 4. Performance comparison of IND rule (circles) and subband ANS technique (triangles).

Rys. 4. Porównanie właściwości systemu opartego na regule wagowej IND (koła) i przetwarzania podpasemowego z wykorzystaniem techniki ANS (trójkąty).

Another example of combined system of echo cancellation and noise reduction in the frequency domain can be found in [29, 30]. The use of frequency domain adaptive filters in the context of acoustic echo cancellation has been extensively studied in the

literatury
correlacji
joint sy
choacoustic
realizacji
putation
properties

3. T

The
ing (LPC

Most
performa
These te
speech i
synthesis

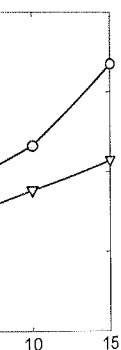
The
of a reco
signal. In
a good q

Subb
input sig
ally quan
for speech
speech by
is much
suppressio
noise red
this soluti
example,
coder. In
LPC spee
soft-decis
report tha
serting no
background
a priori S
by using p
that, the re
techniques

the third pa-
lent transfer

tional (non-
They report
natural near
seems to be
ing threshold
oaches have
s were quiet
e truncated.

problem is
rule is still
ieve further
hole system
advanced in
etween both
n performed
Attenuation
ase of IND
e Fig. 4).



riangles).

etwarzania

reduction in
ain adaptive
udied in the

literature. These algorithms are well suited for long impulse response modeling and correlated input signals like speech [29, 30]. In papers [31, 32], authors have proposed joint systems for acoustic echo and noise control in the frequency domain with psychoacoustic motivation. The complete system appears to be valuable candidate to the realization of car hands-free function, characterized by a good trade-off between computational complexity, low transmission delay, echo cancellation and noise reduction properties.

3. THE PERSPECTIVES OF COMBINED SPEECH ENHANCEMENT AND CODING SYSTEM

There are several different speech-coding methods such as Linear Predictive Coding (LPC), Waveform Coding and Subband Coding [8, 9].

Most schemes use some form of linear predictive coding, which gives high bit-rate performance. Satisfactory quality of speech signals is achieved for bit rates of 16kb/s. These techniques (i.e. CELP, MELP) make use of a speech generation model where speech is represented by an excitation signal feeding through an adaptive filter (LPC synthesis filter).

The main goal of waveform coders (i.e. PCM, DPCM, ADPCM) is a production of a reconstructed signal whose waveform is as close as possible to the original input signal. In this case, speech generation model is not exploited. Waveform coders achieve a good quality of speech signals only for high bit rates such as 48 kb/s.

Subband coders (i.e. ISO/MPEG) exploit filter bank analysis to split wideband input signal into several frequency bands. Spectral components in each band are usually quantized using psychoacoustic rules. Subband coders are widely used not only for speech coding but also for audio coding. It is very difficult to create high quality speech by using low bit rates with this technique. Nevertheless, bit rate performance is much better than for waveform coders. There are two ways for combining noise suppressor and speech coder. One leads to simple arranging two systems in row where noise reduction system is followed by common speech coder. The main advantage of this solution is that both components can be designed independently of each other. For example, frame size of enhancement system can be different from the frame size speech coder. In paper [33], authors have studied influence of such noise pre-processing on LPC speech codecs. In these studies, conventional (non-psychoacoustically motivated) soft-decision Wiener and soft-decision MMSE estimators have been tested. Authors report that in some cases a performance of coding system can be improved by inserting noise reduction pre-processor. This improvement strongly depends on type of background noise and quality of suppression rule parameters estimation, especially a priori SNR. On this way, we believe that further improvements can be obtained by using psychoacoustically motivated methods instead conventional ones. Apart from that, the results of work [33] suggest the use of different optimized speech enhancement techniques for different parameter extraction modules of coder. An example of such

system can be found in [34] where two optimized MMSE estimators are used as pre-processor in conjunction with MELP codec. One pre-processor is optimized for LPC estimation while another for computation of the residual signal. This approach and previous studies only confirm thesis that in many cases, simple pre-processing is not sufficient and further integration is required. Moreover, modular systems consisting of independent noise suppressor and speech coder should not be used for real time processing due to delays introduced by spectral analysis/synthesis [35]. Unfortunately, better integration of some systems is not easy task, especially in case of LPC approaches. Although subband-coding algorithms have now even worse performance than LPC codecs, they can be more easily integrated with speech enhancement systems, especially with psychoacoustically motivated ones. In this case, several common processing-units can be distinguished from both noise suppressor and coder. For example, calculation of masking threshold can be implemented in one module. Moreover, in case of speech enhancement methods based on subband processing [15–20] a spectral analysis and synthesis can be shared between two systems. This may leads to significant improvements in the computation complexity. In work [36], authors propose joint system for perceptual coding and speech enhancement based on fast wavelet packet transform algorithm. Mentioned system uses noise suppressor (adopted to work in critical band domain) similar to that described in [10]. Results of this approach are very promising. From this point of view, also using method presented in Section 2.2 seems to be suitable in order to design new combined speech enhancement and coding system. Furthermore, an additional exploitation of echo canceller votes in favor of this method. This feature may be very useful in some applications, especially in teleconferencing.

4. CONCLUSIONS

More advanced psychoacoustically motivated speech enhancement systems have been presented. The paper shows different approaches to derivation of the perceptually optimal spectral weighting rules. The main advantage of all perceptual methods over conventional ones lies in improvements of musical noise suppression and increase speech intelligibility. Moreover, in case of psychoacoustic joint systems (AEC+NR), decreased requirements for echo canceller are reported.

Non-uniform spectral decomposition and subband processing determine further evolution courses of perceptual noise reduction. It comes from the fact that such approaches are based on superior representation of the human auditory model. It is proper to add that there are some modern speech enhancement techniques where the psychoacoustic is not used, yet. Examples of such systems are approaches based on Blind Signal Separation (BSS) and Independent Component Analysis (ICA) [37].

Discussion on perspectives of integrated, psychoacoustically motivated noise suppressor and coding system shows that integration of subband coder and speech enhancement system based on non-uniformly spaced filter bank leads to most promising combined scheme.

TI
and su

1. K. .
on 3
2. R. .
Con
3. S. .
on A
4. Y. .
time
1984
5. Y. .
amp
33, 1
6. C. .
Conf
1999
7. O. .
supp
8. E. .
9. T. F.
vol. 8
10. P. N.
no. 6
11. N. V.
system
12. T. F.
choac
1997,
13. D. F.
ligibil
Speech
14. D. E.
based
no. 6,
15. K. E.
cychn
Politec
16. K. B.
uniform
IEEE
17. K. B.
based

5. ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewer for their helpful comments and suggestions.

6. REFERENCES

1. K. Kroschel: *Noisereduction: an old problem of actual concern*. 8th International Symposium on Sound Engineering and Mastering (ISSEM'99), Gdańsk, 1999, pp. 3-8
2. R. Martin: *Spectral subtraction based on minimum statistic*. VII European Signal Processing Conference (EUSIPCO'94), Edinburg, 1994, pp. 1182-1185
3. S. F. Boll: *Suppression of acoustic noise in speech using spectral subtraction*. IEEE Transactions on Acoustic Speech and Signal Processing, 1979, vol. ASSP-27, no. 2, pp. 113-120
4. Y. Ephraim, D. Malah: *Speech enhancement using a minimum mean-square error short-time spectral amplitude estimator*. IEEE Transactions on Acoustic Speech and Signal Processing, 1984, vol. ASSP-32, no. 6, pp. 1109-1121
5. Y. Ephraim, D. Malah: *Speech enhancement using a minimum mean-square log-spectral amplitude estimator*. IEEE Transactions on Acoustic Speech and Signal Processing, 1985, vol. ASSP-33, no. 2, pp. 443-445
6. C. Beaugéant, P. Scalart: *Noise reduction using perceptual spectral change*. 6th European Conference on Speech Communication and Technology (EUROSPEECH'99), Budapest, Hungary, 1999, pp. 2543-2546
7. O. Cappé: *Elimination of the musical noise phenomenon with the Ephraim and Malah noise suppressor*. IEEE Transactions on Speech and Audio Processing, 1994, vol. 2, pp. 345-349
8. E. Zwicker, H. Fastl: *Psychoacoustic. Facts and models*, Berlin Springer Verlag, 1990
9. T. Painter, A. Spanias: *Perceptual coding of digital audio*. Proceedings of the IEEE, 2000, vol. 88, no. 4, pp. 451-513
10. P. Noll: *Digital audio coding for visual communications*. Proceedings of the IEEE, 1995, vol. 83, no. 6, pp. 925-943
11. N. Virag: *Single channel speech enhancement based on masking properties of the human auditory system*. IEEE Transactions on Speech and Audio Processing, 1999, vol. 7, no. 2, pp. 126-137
12. T. Haulick, K. Linhard, P. Schrogmeier: *Residual noise suppression using psychoacoustic criteria*. 5th European Conference on Speech Communication and Technology, Greece, 1997, pp. 1395-1398
13. D. E. Tsoukalas, J. N. Mourjopoulos, G. Kokkinakis: *Improving the intelligibility of noisy speech using an audible noise suppression technique*. 5th European Conference on Speech Communication and Technology, Greece, 1997, pp. 1415-1418
14. D. E. Tsoukalas, J. N. Mourjopoulos, G. Kokkinakis: *Speech enhancement based on audible noise suppression*. IEEE Transactions on Speech and Audio Processing, 1997, vol. 5, no. 6, pp. 497-514
15. K. Bielawski: *System eliminacji echa ekustycznego i szumów w urządzeniach głośnomówiących wykorzystujący model psychoakustyczne*. Rozprawa doktorska (promotor prof. A. A. Petrovsky), Politechnika Warszawska, 2001
16. K. Bielawski, A. A. Petrovsky: *Acoustic echo control and noise reduction in non-uniform filter bank: an application of oversampling multirate systems and all-pass transformation*. IEEE Nordic Signal Processing Symposium (NORSIG'2000), Kolmarden, Sweden, 2000, pp. 41-44
17. K. Bielawski, A. A. Petrovsky: *Speech enhancement system for hands-free telephone based on the psychoacoustically motivated filter bank with all-pass frequency transformation*. 6th

- European Conference on Speech Communication and Technology (EUROSPEECH'99), Budapest, Hungary, 1999, pp. 2555–2558
18. A. A. Petrovsky, K. Bielawski: *Combined system for echo cancellation and noise reduction: subbands adaptive filtering approach*. 3rd International Conference, „Sampling Theory” (SAMPTA'99), 1999, Loen, Norway, (invited paper) pp. 1–6
 19. K. Bielawski, A. A. Petrovsky: *Proposition of minimum bands multirate noise reduction system which exploits properties of the human auditory system and all-pass transformed filter bank*. IEEE Workshop Signal Processing, Poznań, Poland, 2001, pp. 65–70
 20. A. A. Petrovsky, K. Bielawski: *Auditory model based enhancement system for hands-free devices*. XI European Signal Processing Conference (EUSIPCO'2002), Toulouse, France, 2002, vol. 1 pp. 487–490
 21. M. Parfieniuk, A. A. Petrovsky: *Struktury polifazowe w cyfrowych bankach filtrów*. Elektronika. SEP Warszawa, Maj 2001, pp. 17–21
 22. A. A. Petrovsky, M. Parfieniuk, K. Bielawski: *Psychoacoustically motivated nonuniform cosine modulated polyphase filter bank*. International TICSP Workshop on Spectral Methods and Multirate Signal Processing (SMMSP'02), Toulouse, France, 2002. pp. 95–101
 23. K. Bielawski, J. Baszun, A. A. Petrovsky: *Cochlear spaced filter bank using allpass frequency transformation*. EURASIP Conference DSP for Multimedia Communication and Services (ECMCS'99), Krakow, Poland, 1999, CD
 24. S. Gustafsson, P. Jax, P. Vary: *A novel psychoacoustically motivated audio enhancement algorithm preserving background noise characteristic*. IEEE International Conference on Acoustic, Speech, and Signal Processing (ICASSP'98), Seattle, USA, 1998, vol. 1, pp. 397–400
 25. S. Gustafsson, P. Jax, A. Kamphausen, P. Vary: *A postfilter for echo and noise reduction avoiding the problem of musical tones*. IEEE International Conference on Acoustic, Speech and Signal Processing (ICASSP'99), Phoenix, USA, 1999, vol. 2, pp. 873–876
 26. S. Gustafsson, R. Martin, P. Jax, P. Vary: *A psychoacoustic approach to combined acoustic echo cancellation and noise reduction*. IEEE Transactions on Speech and Audio Processing, 2002, vol. 10, no. 5, pp. 245–256
 27. V. Turbin, A. Gilloire, P. Scalart, C. Beaugeant: *Using psychoacoustic criteria in acoustic echo cancellation algorithms*. IEEE Workshop on Acoustic Echo and Noise Control, London, 1997, pp. 53–56
 28. S. Gustafsson, R. Martin: *Combined acoustic echo control and noise reduction based on residual echo estimation*. International Workshop on Acoustic Echo and Noise Control, London, U.K., 1997, pp. 160–163
 29. G. P. M. Egelmeers: *Decoupling of partition factors in partitioned block FDAF*. European Conference on Circuits Theory and Design (ECCTD'93), Davos, Switzerland, 1993, pp. 323–329
 30. F. Capman, J. Boudy, P. Lockwood: *Acoustic echo cancellation and noise reduction in the frequency domain: a global optimization*. 8th European Signal Processing Conference (EUSIPCO'96), Trieste, Italy, 1996, pp. 29–32
 31. A. Petrovsky, K. Bielawski, A. Anoshenko: *Hands-free radiotelephony communication devices with combine front end processing systems: global approaches in the time and frequency domain*. Journal of the University of Applied Sciences, Mittweida (FH), Kommunikationstechnik, band C, no. 3, 1998, pp. 135–142
 32. A. A. Petrovsky, A. E. Anoshenko: *Combined system for echo cancellation and noise reduction in frequency domain with psychoacoustic motivation*. 2nd International Conference and Exhibition, „Digital Signal Processing and its Applications”, 1999, Moscow, Russia, pp. 166–169 (in Russian)
 33. D. Virette, P. Scalart, C. Lamblin: *Analysis of background noise reduction techniques for robust speech coding*. XI European Signal Processing Conference (EUSIPCO'2002), Toulouse, France, 2002, vol. 3, pp. 297–300

34. T.
Spee
35. R.
spee
catic
36. B.
chron
vol.
37. A.
Chic

MI
CZ

Doko
alnie. Wsk
wag widm
izacji tych
człowieka.
wości tłum
ieniu uwzg
iono persp
uwydatniaj

Słowa klucz

Praca finan

34. T. Agarwal, P. Kabal: *Pre-processing of noisy speech for voice coders*. IEEE Workshop on Speech Coding, Tsukuba, Japan, 2002, pp. 169–171
35. R. Martin, H. G. Kang, R. V. Cox: *Low delay analysis/synthesis scheme for joint speech enhancement and low bit rate speech coding*. 6th European Conference on Speech Communication and Technology (EUROSPEECH'99), Budapest, Hungary, 1999, vol. 3, pp. 1463–1466
36. B. Carnero, A. Drygajło: *Perceptual speech coding and enhancement using frame-synchronized fast wavelet packet transform algorithms*. IEEE Transactions on Signal Processing, 1999, vol. 47, no. 6, pp. 1622–1635
37. A. Cichocki, S. Amari: *Adaptive Blind Signal and Image Processing*. Wiley & Sons, Ltd., Chichester, UK, 2002

A. BOROWICZ, A.A. PETROVSKY

METODY MOTYWOWANE PERCEPTUALNIE W UZDATNIANIU SYGNAŁU MOWY CZĘŚĆ II: PSYCHOAKUSTYCZNA OPTIMALIZACJA METOD WAG WIDMOWYCH

Streszczenie

Dokonano przeglądu oraz porównania metod uzdatniania sygnału mowy motywowanych perceptualnie. Wskazano na niedoskonałość rozwiązań psychoakustycznych wykorzystujących klasyczne metody wag widmowych. Opierając się na literaturze zaprezentowano różne sposoby psychoakustycznej optymalizacji tych metod. Prezentowane systemy sklasyfikowano według stopnia zgodności z modelem słuchowym człowieka. Jednocześnie zestawiono wyniki zastosowań rozwiązań psychoakustycznych pod kątem możliwości tłumienia szumu środowiskowego i zapobiegania generacji zniekształceń sygnału mowy. W zestawieniu uwzględniono także połączone systemy eliminacji echa i redukcji szumów. Ostatecznie przedstawiono perspektywę integracji systemu uzdatniania sygnału mowy z systemem kodowania podpasmowego uwytłumiając wykorzystanie modeli psychoakustycznych jako element wspólny obu systemów.

Słowa kluczowe: redukcja szumów, eliminacja echa, psychoakustyka, kodowanie mowy.

Praca finansowana z grantu W/WI/3/02.

Mode

wzi
kac
mał
wyj
ster
wyr
waż
cy v

mak
znac
scale
ściow
sowa
w m
zależ
lepsz
firmę

Słowa

Progr
zawierają

Modelowanie zależności maksymalnego napięcia wyjściowego wzmacniaczy operacyjnych od prądu wyjściowego

JÓZEF STANCLIK

*Politechnika Wrocławska, Instytut Telekomunikacji i Akustyki,
50-370 Wrocław, ul. Wybrzeże Wyspiańskiego 27,
E-mail: jozef.stanclik@pwr.wroc.pl*

*Otrzymano 2003.09.16
Autoryzowano 2003.12.12*

Zmodyfikowano strukturę i podano sposób wyznaczania parametrów makromodelu wzmacniaczy operacyjnych, przeznaczonego do stosowania w programie PSpice. Modyfikację przeprowadzono w taki sposób, aby uwzględnić zależność (liniową lub nieliniową) maksymalnego napięcia wyjściowego oraz napięcia nasycenia układu scalonego od prądu wyjściowego. Modyfikacja polega na dodaniu do pierwotnego makromodelu dwóch źródeł sterowanych. Dzięki temu uzyskano lepszą zgodność wyników symulacji z danymi katalogowymi w szerokim przedziale zmian obciążenia, zasilania iysterowania. Jest to szczególnie ważne w przypadku symulacyjnych obliczeń właściwości energetycznych (maksymalnej mocy wyjściowej, mocy strat, sprawności energetycznej) wzmacniaczy mocy.

Modyfikację można stosować do makromodeli różnych wzmacniaczy operacyjnych, tak makromodeli standardowych, jak i ulepszonych (enhanced macromodels). Podstawą do wyznaczania parametrów dodatkowych źródeł sterowanych są typowe dane katalogowe układu scalonego. Podano przykład użycia liniowych źródeł napięciowych sterowanych prądem wyjściowym w makromodelu operacyjnego wzmacniacza mocy OPA502, oraz przykład zastosowania źródeł nieliniowych, w których zależności SEM od prądu są opisane wielomianami, w makromodelu wzmacniacza operacyjnego OPA130. Uzyskana dokładność modelowania zależności maksymalnych napięć wyjściowych od prądów wyjściowych okazała się o rząd lepsza niż w przypadku makromodeli dostarczanych przez producenta układów scalonych, firmę Burr-Brown.

Słowa kluczowe: makromodel, wzmacniacz operacyjny, program PSpice

1. WSTĘP

Programy komputerowe, przeznaczone do projektowania układów elektronicznych, zawierają biblioteki makromodeli, często opracowane przez producentów układów sca-

lonych. Szeroko rozpowszechniony program PSpice jest wyposażony w biblioteki makromodeli różnych wzmacniaczy scalonych: od mikromocowych wzmacniaczy operacyjnych do operacyjnych wzmacniaczy mocy, makromodeli standardowych (standard macromodels), jak i ulepszonych (enhanced macromodels) [1]. Do identyfikacji makromodeli służy specjalistyczne oprogramowanie (np. program PARTS w pakiecie Design Center [4]). Uzyskane w ten sposób makromodele uwzględniają podstawowe parametry i charakterystyki układów scalonych w nominalnych warunkach pracy (zasilanie, temperatura, itp.).

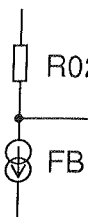
Wspólną cechą wymienionych makromodeli wzmacniaczy operacyjnych jest pominięcie zależności maksymalnej amplitudy napięcia wyjściowego od prądu wyjściowego układu scalonego (napięcie nasycenia jest przyjmowane jako stałe). Z tego powodu wyniki obliczeń symulacyjnych mogą się znacznie różnić się od danych katalogowych. Niedokładne modelowanie zależności napięcia nasycenia od prądu wyjściowego scalonych wzmacniaczy mocy małej częstotliwości może prowadzić do różnic pomiędzy wynikami symulacyjnych obliczeń maksymalnej mocy wyjściowej, mocy zasilania i mocy strat oraz sprawności energetycznej, a odpowiednimi danymi katalogowymi, sięgających 50% [7].

W pracy [5] oceniono dokładność makromodeli scalonych operacyjnych wzmacniaczy mocy opracowanych przez ich producenta, firmę BurrBrown. Wykazano, że błędy modelowania zależności napięcia nasycenia od prądu wyjściowego są znaczne (maksymalne sięgają 37,5%). Możliwość zwiększenia dokładności makromodeli tych układów scalonych przedstawiono w pracy [6]. Jednak zaproponowany tam makromodel ma strukturę klasycznego modelu Boyle'a [2] i sposobu zwiększenia jego dokładności, polegającego na zastosowaniu rezystancji szeregowych w modelach diod ograniczników napięcia, nie można użyć w przypadku makromodeli stosowanych obecnie (generowanych przez program PARTS).

W niniejszym opracowaniu przedstawiono modyfikację makromodeli wzmacniaczy operacyjnych, prowadzącą do uwzględnienia zależności maksymalnego napięcia wyjściowego od prądu wyjściowego (liniowej bądź nieliniowej). Podano strukturę zmodyfikowanego makromodelu i sposób doboru parametrów dodatkowych źródeł sterowanych na podstawie typowych danych katalogowych. Procedurę modyfikacji makromodeli zilustrowano przykładami makromodelu scalonego operacyjnego wzmacniacza mocy typu OPA502, w którym zastosowano liniowe źródła sterowane, oraz makromodelu wzmacniacza operacyjnego OPA130, gdzie użyto źródeł nieliniowych.

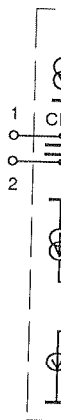
2. ZASADA MODYFIKACJI MAKROMODELI

Zarówno w makromodelach standardowych, jak i ulepszonych, maksymalne i minimalne napięcie wyjściowe jest określone przez parametry obwodów ograniczników napięcia wyjściowego. W obwodach tych, pokazanych na rys. 1a (por. także pełny schemat makromodelu wzmacniacza operacyjnego — rys. 2), występują jednakowe,



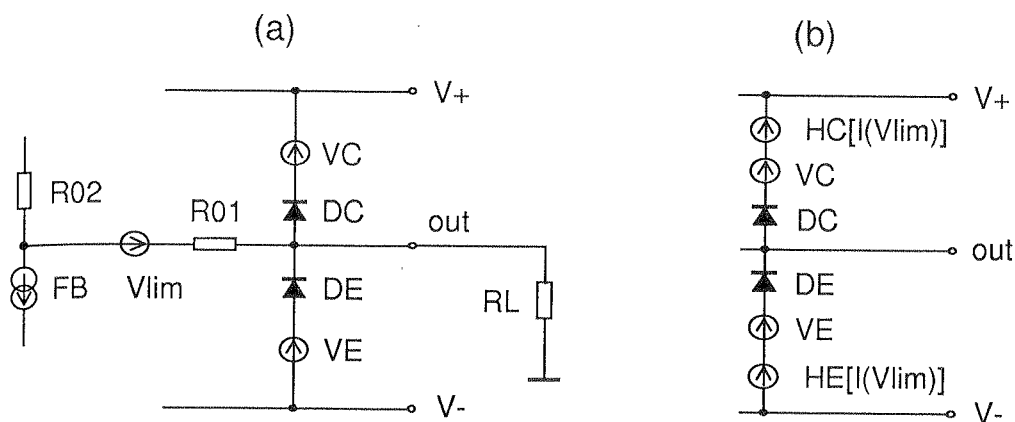
Rys. 1.

F



Diod
 V_{out}^+ . Prąd
jego prąd

idealne diody półprzewodnikowe DC i DE . Napięcia progowe ograniczników są ustalane przez napięcia zasilające V^+ i V^- oraz SEM źródeł pomocniczych VC i VE .

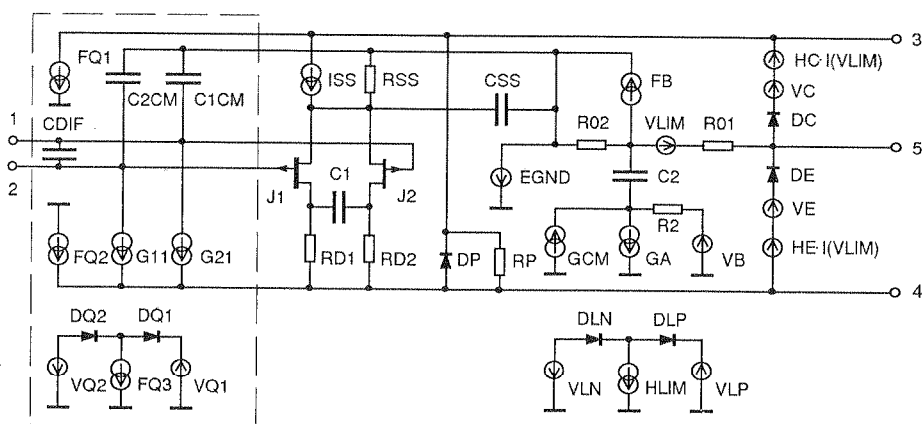


Rys. 1. Schematy ograniczników napięcia wyjściowego w makromodelach wzmacniaczy scalonych:

a) makromodel pierwotny, b) makromodel zmodyfikowany

Fig. 1. Circuit diagrams of output voltage limiters in integrated amplifier macromodels:

a) original macromodel, b) modified macromodel



Rys. 2. Struktura zmodyfikowanego makromodelu wzmacniacza operacyjnego

Fig. 2. Structure of modified macromodel for operational amplifier

Dioda DC przewodzi gdy napięcie wyjściowe przekracza wartość maksymalną V_{out}^+ . Prąd diody steruje źródłem prądowym FB , ograniczając możliwość zwiększania jego prądu. Maksymalne napięcie wyjściowe makromodelu jest dane zależnością

$$V_{out}^+ = V^+ - VC + V_{DC} \quad (1)$$

gdzie V_{DC} jest spadkiem napięcia na przewodzącej diodzie DC. Zależność napięcia nasycenia makromodelu V_{sat}^+ od maksymalnego napięcia wyjściowego wyraża wzór

$$V_{sat}^+ = V^+ - V_{out}^+ = VC - V_{DC} \quad (2)$$

Działanie obwodu ograniczającego ujemne napięcie wyjściowe makromodelu (dioda DE, źródło VE i napięcie V^-) jest analogiczne. Minimalne ujemne napięcie wyjściowe makromodelu V_{out}^- jest opisane zależnością

$$V_{out}^- = V^- + VE - V_{DE}, \quad (3)$$

a związane z nim napięcie nasycenia V_{sat}^- określa wzór

$$V_{sat}^- = -V^- + V_{out}^- = VE - V_{DE} \quad (4)$$

W makromodelach dostępnych w bibliotekach programu PSpice napięcie wyjściowe jest najczęściej ograniczane symetrycznie ($V_{out}^+ = -V_{out}^- = V_{out}$), również oba napięcia nasycenia są sobie równe ($V_{sat}^+ = V_{sat}^- = V_{sat}$). W przypadku niektórych wzmacniaczy scalonych (np. operacyjnych wzmacniaczy mocy) ograniczanie napięcia wyjściowego nie jest symetryczne, a zróżnicowanie wartości napięć V_{out}^+ i $-V_{out}^-$ (oraz odpowiednio V_{sat}^+ i V_{sat}^-) jest uzyskiwane w wyniku przyjęcia różnych wartości SEM źródeł pomocniczych VC i VE.

W tak zbudowanych ogranicznikach napięcia wyjściowego makromodelu wzmacniacza scalonego amplituda maksymalnego nie zniekształconego napięcia wyjściowego i wartość napięcia nasycenia praktycznie nie ulegają zmianie przy zmianach prądu wyjściowego (zmiana napięć w makromodelach jest rzędu miliwoltów, gdyż wynika wyłącznie ze zmian prądu diod ogranicznika w czasie przesterowania układu scalonego, pomijalnie małych w stosunku do prądu wyjściowego).

Napięcie wyjściowe rzeczywistego wzmacniacza scalonego zmniejsza się w miarę zwiększania jego prądu wyjściowego (prądu obciążenia), a więc jest zależne odysterowania, rezystancji obciążenia i napięć zasilających. Producenci wzmacniaczy operacyjnych często zamieszczają w danych katalogowych zależności maksymalnej i minimalnej wartości napięcia wyjściowego od prądu wyjściowego (por. dane katalogowe wzmacniacza operacyjnego OPA130 pokazane na rys. 4). Dla operacyjnych wzmacniaczy mocy podawana jest raczej zależność napięć nasycenia od prądu wyjściowego (por. dane operacyjnego wzmacniacza mocy OPA502 — rys. 3).

Różnice między odpowiednimi napięciami układu scalonego i jego makromodelu mogą sięgać kilku woltów w całym zakresie zmienności prądu wyjściowego. Może to prowadzić do dużych różnic między wynikami obliczeń symulacyjnych a parametrami katalogowymi wzmacniacza scalonego przy pełnymysterowaniu i przesterowaniu wzmacniacza. W szczególności wyniki obliczeń mocy: dostarczonej, wyjściowej i strat, oraz sprawności energetycznej mogą się różnić od danych katalogowych nawet o kilkanaście do kilkudziesięciu procent (nawet o 50% w przypadku maksymalnej mocy wyjściowej wzmacniacza mocy m.cz. [7]).

Rys. 3

Fig.

Rys

Fig

Zwię
z zastoso
można u
sterowane
lu. Sposó
rys. 1b. 9

napięcia
wzór

(2)

elu (dio-
cie wyj-

(3)

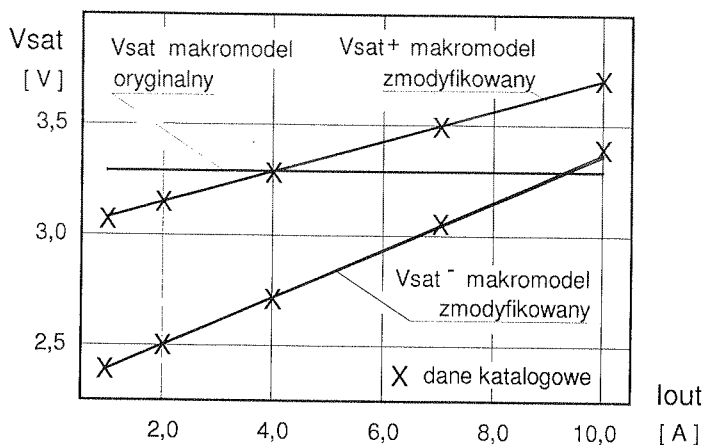
(4)

cie wyj-
nież oba
tektórych
napięcia
 V_{out}^- (oraz
ści SEM

a wzmac-
ściowego
ich prądu
ż wynika
calonego,

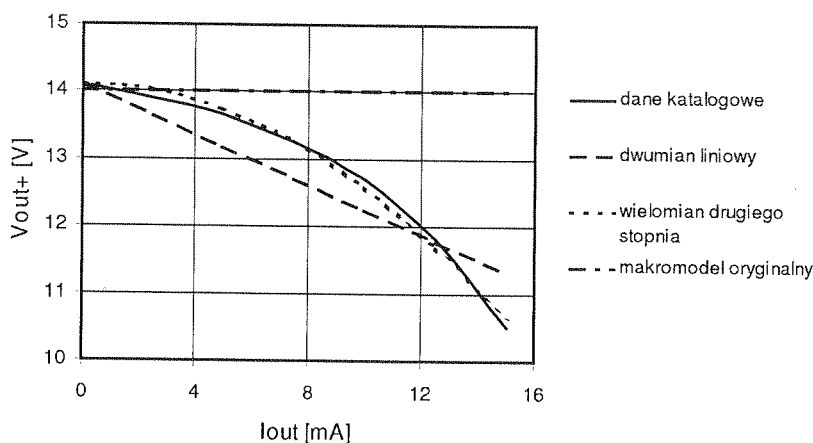
ię w mia-
ależne od
nadcniaczy
ksymalnej
dane kata-
racyjnych
prądu wyj-

makromodelu
o. Może to
parametra-
sterowaniu
wej i strat,
awet o kil-
lnej mocy



Rys. 3. Zależność napięć nasycenia operacyjnego wzmacniacza mocy OPA502 i jego makromodeli od prądu wyjściowego

Fig. 3. Saturation voltages against output current for OPA502 high-power operational amplifier and its macromodels



Rys. 4. Zależność maksymalnych napięć wyjściowych wzmacniacza operacyjnego OPA130 i jego makromodeli od prądu wyjściowego

Fig. 4. Maximum output voltages against output current for OPA130 operational amplifier and its macromodels

Zwiększenie zgodności danych katalogowych i wyników symulacji wykonanej z zastosowaniem użytkowanych powszechnie makromodeli wzmacniaczy scalonych można uzyskać w wyniku uzupełnienia obwodów ograniczników napięcia o źródła sterowane *HC* i *HE*, których SEM są zależne od prądu wyjściowego makromodelu. Sposób modyfikacji struktury makromodelu wzmacniacza scalonego pokazano na rys. 1b. SEM źródeł *HC* i *HE* zmieniają się w zależności od prądu wyjściowego

makromodelu i przesuwają napięcia progowe ograniczników. Do próbkowania prądu wyjściowego można wykorzystać źródło V_{lim} makromodelu.

Podstawą do wyznaczenia parametrów źródeł HC i HE może być zespół wzorów określających maksymalne i minimalne napięcia wyjściowe wzmacniacza w zależności od prądu wyjściowego, w n i m (odpowiednio) punktach charakterystyki $V_{out}(I_{out})$:

$$V_{out i}^+ = f(I_{out i}), \quad 1 \leq i \leq n, \quad (5a)$$

$$V_{out j}^- = f(I_{out j}), \quad 1 \leq j \leq m \quad (5b)$$

Analogicznie określa się zależność dodatniego i ujemnego napięcia nasycenia wzmacniacza od prądu wyjściowego zespołem wzorów:

$$V_{sat i}^+ = f(I_{out i}), \quad 1 \leq i \leq n, \quad (6a)$$

$$V_{sat j}^- = f(I_{out j}), \quad 1 \leq j \leq m, \quad (6b)$$

SEM źródła HC przy danym prądzie wyjściowym iysterowaniu wzmacniacza, powodującym pracę na progu ograniczania, spełnia równanie oczka (rys. 1b):

$$HC = V^+ - VC + V_{DC} - V_{out}^+ \quad (7a)$$

Analogicznie SEM źródła HE spełnia równanie:

$$HE = -V^- - VE + V_{DE} + V_{out}^- \quad (7b)$$

Po podstawieniu danych układu scalonego (5a) i (5b) do tych zależności, odpowiednio, uzyskujemy

$$HC(I_{out i}) = V^+ - VC + V_{DC} - V_{out i}^+(I_{out i}), \quad 1 \leq i \leq n, \quad (8a)$$

$$HE(I_{out j}) = -V^- - VE + V_{DE} + V_{out j}^-(I_{out j}), \quad 1 \leq j \leq m \quad (8b)$$

Jeśli podstawą modyfikacji są zależności napięć nasycenia od prądu, to zależności przyjmują postać

$$HC(I_{out i}) = V_{sat i}^+(I_{out i}) - VC + V_{DC}, \quad 1 \leq i \leq n, \quad (9a)$$

$$HE(I_{out j}) = V_{sat j}^-(I_{out j}) - VE + V_{DE}, \quad 1 \leq j \leq m \quad (9b)$$

Po przyjęciu postaci funkcji $HC(I_{out})$ i $HE(I_{out})$, wzory (8a) i (8b), lub (9a) i (9b), stanowią podstawę do wyznaczenia parametrów tych funkcji.

3. MODYFIKACJA MAKROMODELU Z ZASTOSOWANIEM LINIOWYCH ŹRÓDEŁ STEROWANYCH

Strukturę zmodyfikowanego makromodelu wzmacniacza operacyjnego, odtworzoną na podstawie listy połączeń makromodeli z bibliotek programu PSpice, pokazano na rys. 2 (w makromodelu pierwotnym nie występują źródła HC i HE). Część zaznaczona linią przerywaną modeluje pobór mocy ze źródeł zasilających oraz impedancję wejściową układu scalonego i występuje jedynie w ulepszonej wersji makromodelu (enhanced macromodel).

Jeśli zmiany napięć wyjściowych lub napięć nasycenia wzmacniacza operacyjnego są w przybliżeniu proporcjonalnie do prądu wyjściowego, to do modelowania tych zależności wystarczy zastosowanie w obwodach ograniczników napięcia liniowych źródeł HC i HE sterowanych prądem wyjściowym. Wtedy na podstawie wzorów (8) uzyskujemy

$$HC(I_{out}) = a_C I_{out} \approx V^+ - V_C + V_{DC} - V_{out}^+(I_{out}) \quad (10a)$$

$$HE(I_{out}) = a_E I_{out} \approx -V^- - V_E + V_{DE} + V_{out}^-(I_{out}) \quad (10b)$$

Do wykonania modyfikacji makromodelu jest konieczne wyznaczenie transrezystancji a_C i a_E dodatkowych źródeł sterowanych. Może też być niezbędna zmiana wartości SEM źródeł pomocniczych V_C i V_E . Do wyznaczenia transrezystancji a_C i siły elektromotorycznej V_C wystarczy znajomość charakterystyki (5a) w dwóch punktach ($n = 2$). Na tej podstawie:

$$a_C = \frac{V_{out\ 1}^+ - V_{out\ 2}^+}{I_{out\ 2} - I_{out\ 1}} \quad (11a)$$

$$V_C = V^+ + V_{DC} - V_{out\ 1}^+ - \frac{V_{out\ 1}^+ - V_{out\ 2}^+}{I_{out\ 2} - I_{out\ 1}} I_{out\ 1} \quad (11b)$$

Analogicznie a_E i V_E można obliczyć z zależności:

$$a_E = \frac{V_{out\ 2}^- - V_{out\ 1}^-}{I_{out\ 2} - I_{out\ 1}} \quad (12a)$$

$$V_E = -V^- + V_{out\ 1}^- + V_{DE} - \frac{V_{out\ 2}^- - V_{out\ 1}^-}{I_{out\ 2} - I_{out\ 1}} I_{out\ 1} \quad (12b)$$

We wzorach (11b) i (12b) przyjęto spadki napięć V_{DC} i V_{DE} takie jak w makromodelu przed modyfikacją.

W przypadku, gdy w danych katalogowych układu scalonego są podane napięcia nasycenia jako funkcje prądu wyjściowego, to należy wykorzystać następujące zależności określające parametry zmodyfikowanego ogranicznika napięcia wyjściowego makromodelu

$$a_C = \frac{V_{sat\ 2}^+ - V_{sat\ 1}^+}{I_{out\ 2} - I_{out\ 1}} \quad (13a)$$

Tabela 1

Zależność napięć nasycenia operacyjnego wzmacniacza mocy OPA502 i jego makromodeli od prądu wyjściowego

Dependence of saturation voltages on output current for OPA502 high-power operational amplifier and its macromodels

I_{out}	[A]	1,0	2,0	4,0	7,0	10,0
$V_{sat\ IC}^+$	[V]	3,1	3,17	3,3	3,5	3,7
$V_{sat\ mm}^+$	[V]	3,101	3,167	3,299	3,498	3,697
$V_{sat\ BB}^+$	[V]	3,3	3,3	3,3	3,3	3,3
$V_{sat\ IC}^-$	[V]	2,4	2,5	2,7	3,0	3,4
$V_{sat\ mm}^-$	[V]	2,393	2,501	2,716	3,038	3,358
$V_{sat\ BB}^-$	[V]	3,3	3,3	3,3	3,3	3,3

$$VC = V_{sat\ 1}^+ + V_{DC} - \frac{V_{sat\ 2}^+ - V_{sat\ 1}^+}{I_{out\ 2} - I_{out\ 1}} I_{out\ 1} \quad (13b)$$

$$a_E = \frac{V_{sat\ 2}^- - V_{sat\ 1}^-}{I_{out\ 2} - I_{out\ 1}} \quad (14a)$$

$$VE = V_{sat\ 1}^- + V_{DE} - \frac{V_{sat\ 2}^- - V_{sat\ 1}^-}{I_{out\ 2} - I_{out\ 1}} I_{out\ 1} \quad (14b)$$

Zwykle dane katalogowe układu scalonego zawierają charakterystyki $V_{out}^+ = f(I_{out})$ i $V_{out}^- = f(I_{out})$ lub $V_{sat}^+ = f(I_{out})$ i $V_{sat}^- = f(I_{out})$. W takim przypadku podstawą do wyznaczania parametrów ograniczników napięcia wyjściowego mogą być uśrednione wartości odpowiednich napięć i prądów. Można także zastosować procedurę liniowej aproksymacji dostępną w różnych programach komputerowych.

Przykład

Dla zilustrowania przedstawionej modyfikacji wyznaczono wartości SEM źródeł autonomicznych VC i VE oraz transrezystancji źródeł sterowanych HC i HE ograniczników napięcia wyjściowego makromodelu scalonego, operacyjnego wzmacniacza mocy OPA502. Jego dane katalogowe [3] zawierają informacje o zależności napięć nasycenia wzmacniacza od prądu wyjściowego w formie wykresów pokazanych na rys. 3. Wybrane punkty odpowiednich charakterystyk przytoczono w tabeli 1 (napięcia odczytane z danych katalogowych oznaczono jako $V_{sat\ IC}^+$ i $V_{sat\ IC}^-$). Na tej podstawie, po uśrednieniu, wyznaczono z wzorów (13a) i (14a) transrezystancje źródeł $a_C = 0,0667$ V/A i $a_E = -0,107$ V/A, a z wzorów (13b) i (14b) SEM źródeł $VC = 3,76$ V i $VE = 3,01$ V. Zmodyfikowany makromodel wzmacniacza OPA502 zapisany w formacie bibliotek podukładów liniowych programu PSpice zamieszczono w tabeli 2. Tłustym drukiem zaznaczono zmiany w stosunku do makromodelu pierwotnego.

Tabela 1

od prądu

plifier and

Tabela 2

Zmodyfikowany makromodel operacyjnego wzmacniacza mocy OPA502

Modified macromodel subcircuit for OPA502 high-power operational amplifier

```

* OPA502 OPERATIONAL AMPLIFIER "MACROMODEL" SUBCIRCUIT
*
* CONNECTIONS:      NON-INVERTING INPUT
*                   | INVERTING INPUT
*                   | | POSITIVE POWER SUPPLY
*                   | | | NEGATIVE POWER SUPPLY
*                   | | | | OUTPUT
*                   | | | | |
.SUBCKT OPA502/BB 1 2 3 4 5
C1 11 12 21.45E-12
C2 6 7 12.00E-12
CSS 10 99 10.00E-12
DC 5 53 DX
DE 54 5 DX
DLP 90 91 DX
DLN 92 90 DX
DP 4 3 DX
EGND 99 0 POLY(2) (3,0) (4,0) 0 .5 .5
FB 7 99 POLY(5) VB VC VE VLP VLN 0 886.3E6 -90E6
+90E6 90E6 -90E6
GA 6 0 11 12 113.1E-6
GCM 0 6 10 99 1.131E-9
ISS 3 10 DC 240.0E-6
HLIM 90 0 VLIM 1E3
J1 11 2 10 JX
J2 12 1 10 JX
R2 6 9 100.0E3
RD1 4 11 8.842E3
RD2 4 12 8.842E3
RO1 8 5 5
RO2 7 99 1.5
RP 3 4 4.000E3
RSS 10 99 833.3E3
VB 9 0 DC 0
VC 3 55 DC 3.76 ;VC 3 53 DC 4
VE 56 4 DC 3.01 ;VE 54 4 DC 4
HC 55 53 VLIM 66.7E-3
HE 54 56 VLIM -0.107
VLIM 7 8 DC 0
VLP 91 0 DC 15.00E3
VLN 0 92 DC 15.00E3
.MODEL DX D(IS=800.0E-18)
.MODEL JX PJF(IS=6.000E-12 BETA=106.6E-6 VTO=-1)
.ENDS

```

Zmodyfikowany makromodel poddano weryfikacji. Wykonano za pomocą programu PSpice obliczenia zależności napięć nasycenia od prądu wyjściowego. Na rysunku 3 pokazano zależność tych napięć od prądu wyjściowego dla zmodyfikowanego makromodelu, a w tabeli 1 zestawiono odpowiednie wartości napięć, oznaczone jako V_{sat}^+ i V_{sat}^- . Widać, że napięcia nasycenia zmieniają się ze zmianami prądu wyjściowego a różnice między wartościami odczytanymi z danych katalogowych i obliczonymi dla zmodyfikowanego makromodelu nie przekraczają 42 mV (1,2%) dla napięcia V_{sat}^- oraz 3 mV (0,1%) dla napięcia V_{sat}^+ .

Dla porównania na rys. 3 pokazano napięcie nasycenia obliczone w takich samych warunkach z użyciem makromodelu, pochodzącego z biblioteki BURR_BRN.LIB programu PSpice. Napięcia nasycenia mają jednakową wartość 3,3 V i praktycznie nie zależą od prądu. Błędy modelowania dodatniego napięcia nasycenia sięgają 0,4 V a napięcia ujemnego -0,9 V. Przy pracy z małymi prądami wyjściowymi oznacza to błąd około +37,5%. Jak widać zaproponowana modyfikacja prowadzi do kilkudziesięciokrotnego zmniejszenia błędów modelowania zależności napięć nasycenia od prądu wyjściowego.

4. MODYFIKACJA MAKROMODELU Z UŻYCIEM ŹRÓDEŁ NIELINIOWYCH

W przypadku wzmacniaczy operacyjnych, w których zmiany maksymalnego napięcia wyjściowego lub napięcia nasycenia nie są proporcjonalnie do prądu wyjściowego, do modelowania właściwości wzmacniaczy w obszarze nasycenia można użyć dostępnych w programie PSpice nieliniowych źródeł sterowanych HC i HE o opisie wielomianowym

$$HC = a_{0C} + a_{1C}I_{out} + a_{2C}I_{out}^2 + \dots, \quad (15a)$$

$$HE = a_{0E} + a_{1E}I_{out} + a_{2E}I_{out}^2 + \dots \quad (15b)$$

Stopień wielomianu może być dowolny. Jego zwiększenie prowadzi do zmniejszenia błędów aproksymacji, rośnie jednak przy tym liczba parametrów modelu. Dlatego stosowanie wielomianów stopnia wyższego niż drugi nie wydaje się celowe.

Po przyjęciu opisu źródeł za pomocą trójmianów kwadratowych z połączenia zależności (8a) i (15a) oraz (8b) i (15b) otrzymujemy:

$$HC = a_{0C} + a_{1C}I_{out i} + a_{2C}(I_{out i})^2 \approx V^+ - VC + V_{DC} - V_{out i}^+, \quad 1 \leq i \leq n, \quad (16a)$$

$$HE = a_{0E} + a_{1E}I_{out j} + a_{2E}(I_{out j})^2 \approx -V^- - VE + V_{DE} + V_{out j}^-, \quad 1 \leq j \leq m \quad (16b)$$

a jeśli podstawą modyfikacji są zależności napięć nasycenia od prądu, to odpowiednie zależności przyjmują postać

$$HC = a_{0C} + a_{1C}I_{out i} + a_{2C}(I_{out i})^2 \approx V_{sat i}^+ - VC + V_{DC}, \quad 1 \leq i \leq n, \quad (17a)$$

Podobnie
jak w m

$$HE = a_{0E} + a_{1E}I_{outj} + a_{2E}(I_{outj})^2 \approx V_{satj}^- - VE + V_{DE}, \quad 1 \leq j \leq m \quad (17b)$$

przy czym liczby punktów są większe od stopni wielomianów ($n > 2, m > 2$).

Do wykonania modyfikacji makromodelu jest konieczne wyznaczenie parametrów a_{0C} , a_{1C} i a_{2C} oraz a_{0E} , a_{1E} i a_{2E} źródeł sterowanych (wartości SEM źródeł pomocniczych VC i VE pozostają niezmiennione). Wystarczy do tego znajomość charakterystyk (5a) i (5b) w trzech punktach ($n = m = 3$). Wtedy:

$$a_{0C} = V^+ - VC + V_{DC} - V_{out1}^+ - \frac{I_{out1}}{I_{out3} - I_{out1}} \left[(V_{out3}^+ - V_{out2}^+) \frac{I_{out2}}{I_{out3} - I_{out2}} + \right. \\ \left. - (V_{out2}^+ - V_{out1}^+) \frac{I_{out3} + I_{out2} - I_{out1}}{I_{out2} - I_{out1}} \right] \quad (18a)$$

$$a_{1C} = \frac{1}{I_{out3} - I_{out1}} \left[(V_{out3}^+ - V_{out2}^+) \frac{I_{out2} + I_{out1}}{I_{out3} - I_{out2}} + \right. \\ \left. - (V_{out2}^+ - V_{out1}^+) \frac{I_{out3} + I_{out2}}{I_{out2} - I_{out1}} \right] \quad (18b)$$

$$a_{2C} = \frac{1}{I_{out3} - I_{out1}} \left[\frac{V_{out2}^+ - V_{out1}^+}{I_{out2} - I_{out1}} - \frac{V_{out3}^+ - V_{out2}^+}{I_{out3} - I_{out2}} \right] \quad (18c)$$

(15a)

(15b)

$$a_{0E} = -V^- - VE + V_{DE} + V_{out1}^- + \frac{I_{out1}}{I_{out3} - I_{out1}} \left[(V_{out3}^- - V_{out2}^-) \frac{I_{out2}}{I_{out3} - I_{out2}} + \right. \\ \left. - (V_{out2}^- - V_{out1}^-) \frac{I_{out3} + I_{out2} - I_{out1}}{I_{out2} - I_{out1}} \right] \quad (19a)$$

$$a_{1E} = \frac{1}{I_{out3} - I_{out1}} \left[(V_{out2}^- - V_{out1}^-) \frac{I_{out3} + I_{out2}}{I_{out2} - I_{out1}} + \right. \\ \left. - (V_{out3}^- - V_{out2}^-) \frac{I_{out2} + I_{out1}}{I_{out3} - I_{out2}} \right] \quad (19b)$$

$$a_{2E} = \frac{1}{I_{out3} - I_{out1}} \left[\frac{V_{out3}^- - V_{out2}^-}{I_{out3} - I_{out2}} - \frac{V_{out2}^- - V_{out1}^-}{I_{out2} - I_{out1}} \right] \quad (19c)$$

Podobnie jak poprzednio, przyjęto napięcia V_{DC} i V_{DE} we wzorach (18a) i (19a) takie jak w makromodelu przed modyfikacją.

Gdy dane katalogowe wzmacniacza zawierają zależności napięć nasycenia od prądu wyjściowego, to wzory określające parametry źródeł *HC* i *HE* przyjmują postać

$$a_{0C} = V_{sat1}^+ - VC + V_{DC} + \frac{I_{out1}}{I_{out3} - I_{out1}} \left[\left(V_{sat3}^+ - V_{sat2}^+ \right) \frac{I_{out2}}{I_{out3} - I_{out2}} + \right. \\ \left. - \left(V_{sat2}^+ - V_{sat1}^+ \right) \frac{I_{out3} + I_{out2} - I_{out1}}{I_{out2} - I_{out1}} \right] \quad (20a)$$

$$a_{1C} = \frac{1}{I_{out3} - I_{out1}} \left[\left(V_{sat2}^+ - V_{sat1}^+ \right) \frac{I_{out3} + I_{out2}}{I_{out2} - I_{out1}} + \right. \\ \left. - \left(V_{sat3}^+ - V_{sat2}^+ \right) \frac{I_{out2} + I_{out1}}{I_{out3} - I_{out2}} \right] \quad (20b)$$

$$a_{2C} = \frac{1}{I_{out3} - I_{out1}} \left[\frac{V_{sat3}^+ - V_{sat2}^+}{I_{out3} - I_{out2}} - \frac{V_{sat2}^+ - V_{sat1}^+}{I_{out2} - I_{out1}} \right] \quad (20c)$$

$$a_{0E} = V_{sat1}^- - VE + V_{DE} + \frac{I_{out1}}{I_{out3} - I_{out1}} \left[\left(V_{sat3}^- - V_{sat2}^- \right) \frac{I_{out2}}{I_{out3} - I_{out2}} + \right. \\ \left. - \left(V_{sat2}^- - V_{sat1}^- \right) \frac{I_{out3} + I_{out2} - I_{out1}}{I_{out2} - I_{out1}} \right] \quad (21a)$$

$$a_{1E} = \frac{1}{I_{out3} - I_{out1}} \left[\left(V_{sat2}^- - V_{sat1}^- \right) \frac{I_{out3} + I_{out2}}{I_{out2} - I_{out1}} + \right. \\ \left. - \left(V_{sat3}^- - V_{sat2}^- \right) \frac{I_{out2} + I_{out1}}{I_{out3} - I_{out2}} \right] \quad (21b)$$

$$a_{2E} = \frac{1}{I_{out3} - I_{out1}} \left[\frac{V_{sat3}^- - V_{sat2}^-}{I_{out3} - I_{out2}} - \frac{V_{sat2}^- - V_{sat1}^-}{I_{out2} - I_{out1}} \right] \quad (21c)$$

Jeżeli dane katalogowe układu scalonego zawierają charakterystyki $V_{out}^+ = f(I_{out})$ i $V_{out}^- = f(I_{out})$ lub $V_{sat}^+ = f(I_{out})$ i $V_{sat}^- = f(I_{out})$, współczynniki wielomianów można wyznaczyć za pomocą uniwersalnego programu matematycznego (Matlab, Mathcad, itp.) tak, aby zminimalizować błędy aproksymacji.

Przykład.

Zaproponowana modyfikacja została wprowadzona do makromodelu wzmacniacza operacyjnego OPA130, dostępnego w bibliotece BURR_BRN.LIB programu PSpice. Obwody ograniczników napięcia wyjściowego uzupełniono o nieliniowe źródła sterowane *HC* i *HE*. Informacje o zależności maksymalnego i minimalnego napięcia wyjściowego od prądu wyjściowego są podawane przez producenta układu scalonego

Współczynniki wielomianu (15a) wyznaczono za pomocą programu Excel. Narzucono zgodność maksymalnego napięcia wyjściowego dla zerowego prądu wyjściowego. Uzyskano następującą postać wielomianu drugiego stopnia:

$$HC = -0,1 - 9,5313I_{out} + 16,152 \cdot 10^3 I_{out}^2 \quad (22a)$$

Analogiczne obliczenia wykonano dla minimalnej wartości napięcia wyjściowego uzyskując opis źródła HE za pomocą wielomianu:

$$HE = 7,8273I_{out} + 20,752 \cdot 10^3 I_{out}^2 \quad (22b)$$

Zmodyfikowany makromodel wzmacniacza OPA130 ze źródłami HC i HE opisanymi wielomianami (22a) i (22b) w formacie stosowanym w bibliotekach programu PSpice zamieszczono w tabeli 4. Wprowadzone zmiany zaznaczono tłustym drukiem.

Makromodel ten poddano weryfikacji. Wykonano za pomocą programu PSpice obliczenia zależności napięć wyjściowych od prądu wyjściowego i porównano z danymi katalogowymi. Na rysunkach 4 i 5 pokazano zależności napięć wyjściowych od prądu dla zmodyfikowanego makromodelu, a w tabeli 3 przytoczono wartości napięcia maksymalnego uzyskane w wyniku symulacji z użyciem źródeł opisanych wielomianami (22a) i (22b), oznaczone jako $V_{sat sq}^+$. Napięcie wyjściowe jest ograniczane na poziomie zależnym od prądu wyjściowego i błędy nie przekraczają kilkuset miliwoltów.

Podobne obliczenia wykonano z zastosowaniem dwumianów liniowych. Źródła HC i HE opisano funkcjami

$$HC = -0,1 + 186,15I_{out}, \quad (23a)$$

$$HE = 248,49I_{out} \quad (23b)$$

Zależności napięć wyjściowych od prądu wykreślono na rys. 4 i 5 a wartości maksymalnego napięcia wyjściowego (oznaczonego jako $V_{sat lin}^+$) podano w tabeli 3. Błędy przy opisie źródeł dwumianami liniowymi nie przekraczają 1 V.

Dla porównania na rys. 4 i 5 pokazano napięcia wyjściowe obliczone w takich samych warunkach z użyciem makromodelu pierwotnego. Napięcia wyjściowe praktycznie nie zależą od prądu. Błędy modelowania maksymalnego napięcia wyjściowego sięgają 3,5 V a napięcia minimalnego $-4,5$ V. Jak widać zaproponowana modyfikacja prowadzi do zmniejszenia błędów modelowania zależności maksymalnego i minimalnego napięcia wyjściowego od prądu wyjściowego: kilkukrotnego w przypadku zastosowania do opisu źródeł dwumianów liniowych oraz o jeden rząd w przypadku zastosowania trójmianów kwadratowych.

Tabela 4

Zmodyfikowany makromodel wzmacniacza operacyjnego OPA130

Modified macromodel subcircuit for OPA130 operational amplifier

```

* OPA130 OPERATIONAL AMPLIFIER "MACROMODEL" SUBCIRCUIT
*
.SUBCKT OPA130
+ 1      ; non-inverting input
+ 2      ; inverting input
+ 3      ; positive power supply
+ 4      ; negative power supply
+ 5      ; output
*
C1      11 12 4.330E-12
C2      6  7 15.00E-12
CSS     10 99 1.000E-30
DC      5 53 DX
DE      54 5 DX
DLP     90 91 DX
DLN     92 90 DX
DP      4  3 DX
EGND    99 0 POLY(2) (3,0) (4,0) 0 .5 .5
FB      7 99 POLY(5) VB VC VE VLP VLN 0 2.386E9 -2E9
+2E9 2E9 -2E9
GA      6  0 11 12 94.25E-6
GCM     0  6 10 99 530.1E-12
ISS     3 10 DC 30.00E-6
HLIM    90 0 VLIM 1E3
J1      11 2 10 JX
J2      12 1 10 JX
R2      6  9 100.0E3
RD1     4 11 10.61E3
RD2     4 12 10.61E3
RO1     8  5 50
RO2     7 99 25
RP      3  4 56.25E3
RSS     10 99 6.667E6
VB      9  0 DC 0
VC    55 53 DC 1.500 ;VC 3 53 DC 1.500
VE    54 56 DC 1 ;VE 54 4 DC 1
HC    3 55 POLY(1) VLIM -.1 -9.5313 16.152E3
HE    56 4 POLY(1) VLIM 0 7.8273 20.752E3
VLIM    7  8 DC 0
VLP     91 0 DC 18
VLN     0 92 DC 18
.MODEL  DX D(IS=800.0E-18)
.MODEL  JX PJF(IS=2.500E-12 BETA=296.1E-6 VTO=-1)
.ENDS

```

5. WNIOSKI

Zaproponowana modyfikacja makromodeli wzmacniaczy operacyjnych zapewnia modelowanie zależności maksymalnego i minimalnego napięcia wyjściowego (a także napięć nasycenia) układu scalonego od prądu wyjściowego. W wyniku jej wprowadzenia zmniejszono różnice między odpowiednimi wynikami obliczeń symulacyjnych a parametrami katalogowymi wzmacniacza w szerokim zakresie zmienności warunków pracy (przedziale zmian rezystancji obciążenia, napięć zasilających i sterujących). Ma to szczególne znaczenie w przypadku symulacyjnych obliczeń właściwości energetycznych (maksymalnej mocy wyjściowej, mocy strat, czy sprawności energetycznej) operacyjnych wzmacniaczy mocy.

Wprowadzenie modyfikacji nie powoduje istotnego zwiększenia złożoności makromodelu, polega bowiem na dodaniu do pierwotnego makromodelu dwóch sterowanych źródeł napięciowych (liniowych lub nieliniowych, opisanych wielomianami). Do identyfikacji parametrów tych źródeł nie są potrzebne dodatkowe pomiary układów scalonych. Wystarczającą podstawę stanowią typowe dane katalogowe. Wprowadzenie modyfikacji do listy połączeń makromodelu w bibliotece symulatora układów elektronicznych polega na zmianie zapisu dwóch wierszy oryginalnej listy połączeń i dopisaniu dwóch wierszy dodatkowych.

Zaproponowaną modyfikację można stosować do makromodeli różnych wzmacniaczy scalonych, w tym również do makromodeli wzmacniaczy mocy m.cz.

6. BIBLIOGRAFIA

1. H. Biagi, R. M. Stitt, B. Baker, S. Baier: *Burr-Brown Spice Based Macromodels*. Burr-Brown Application Bulletin, AB-020F, Texas Instruments Incorporated, 2000
2. G. R. Boyle, B. M. Cohn, D. O. Pederson, J. E. Solomon: *Macromodeling of Integrated Circuit Operational Amplifiers*. IEEE Journ. of Solid-State Circuits 1974, Vol. SC-9, pp. 353-364
3. BURR-BROWN IC Data Book — Linear Products 1996. Burr-Brown Corporation, 1996/1997, pp. 3.2.41-3.2.50, 2.53-2.59
4. MicroSim PSpice A/D. Ver.6.3. Reference Manual. MicroSim Corporation, April 1996
5. J. Stanclik: *Comparison of BiFET High-Power Operational Amplifier Macromodels*. Proc. of the ICSES'2000, Gliwice-Ustroń, Poland, 17-20 Oct. 2000, pp. 409-414
6. J. Stanclik: *Modyfikacja makromodelu operacyjnych wzmacniaczy mocy przeznaczonego do wspomagającego komputerem projektowania elektroakustycznych wzmacniaczy dużej mocy*. Materiały XLVII OSA, Rzeszów-Jawor, 19-22.09.2000, ss. 597-602
7. J. Stanclik: *Zwiększenie dokładności makromodelu scalonych wzmacniaczy mocy małej częstotliwości*. Kwartalnik Elektroniki i Telekomunikacji, 2001, T. 47, z. 4, ss. 463-476

J. STANCLIK

MODELLING THE RELATION BETWEEN MAXIMUM OUTPUT VOLTAGE AND
OUTPUT CURRENT OF OPERATIONAL AMPLIFIERS

Summary

The structure of operational amplifier macromodels, intended to be used in PSpice Programme, was modified and a way of deriving additional parameters for these macromodels was also given. Macromodel was modified so as to allow for relation (of both linear and non-linear nature) between maximum output voltage and IC saturation voltage on the one side, and the output current on the other side. The modification made consists in supplementing the original macromodel with two controlled sources. When they were introduced, macromodel accuracy was improved and operating condition area (load, power supply, control signal), wherein the macromodel ensures high accuracy, was extended. This modification may be applied to macromodels of various operational amplifiers, standard macromodels and enhanced macromodels. Typical IC's data-sheet information is the sufficient basis to calculate additional macromodel parameters. Examples were provided where simple current-controlled voltage sources were applied in IC's macromodel for the OPA502 Power Operational Amplifier and where non-linear controlled sources were introduced to macromodel of the OPA130 Operational Amplifier. The identified macromodels were subjected to verification. The accuracy of modelling the maximum output voltages and saturation voltages proved to be one order better than attained for original macromodels developed by Burr-Brown Company, the manufacturer of integrated circuits under consideration.

Keywords: macromodel, operational amplifier, PSpice Programme

For

the
ga
the
re
an
wh
me
tio
am

Ke

Dir
conditio
of the tw
limits. F
formula

Four-State Measurement Method for Direct-Reading Noise Figure Measurement of Low-Gain Two-Ports

WŁADYSŁAW ADAMSKI

*Gdańsk University of Technology
Faculty of Electronics, Telecommunications & Informatics
Narutowicza 11/12, 80- 952 Gdańsk, Poland
e-mail: wla@eti.pg.gda.pl*

*Otrzymano 2003.05.14
Autoryzowano 2003.11.27*

Noise figure F_m displayed by an automatic noise figure indicator should be equal to the true noise figure F_x value of the two-port under test, regardless of its available power gain G_{Ax} and noise figure of the measurement receiver F_r . Technical problems related to the F_x direct readings and well-known methods for such measurements have been briefly recalled. A new method that allows the elimination of the influence of both noise figure F_r and gain G_{Ax} values on the noise figure reading F_m is presented. Formulas have been given which allow the estimation of the technical requirements for additional components of the measurement system or to evaluate the method measurement error. Theoretical considerations have been supported by an example of noise figure measurements of the microwave amplifier.

Keywords: noise measurements, noise figure, automatic meters.

1. INTRODUCTION

Direct-reading noise figure measurement means achieving such a measurement condition in which the measured noise figure F_m and the true value of noise figure F_x of the two-port under test (TUT) satisfy the equation $F_m = F_x$ within acceptable error limits. For conventional noise figure measurement methods this relation satisfies Friis' formula

$$F_m = F_x (\Gamma_g) + \frac{F_r (\Gamma_{2x}) - 1}{G_{Ax} (\Gamma_g)} \quad , \quad (1)$$

where F_r is the noise figure of the measurement receiver formed by blocks connected to the TUT output in cascade, G_{Ax} is an available power gain of the TUT, Γ_g is an output reflection coefficient of the noise generator connected to the TUT input and Γ_{2x} is the output reflection coefficient of the TUT. Neglecting the second term in right hand of (1) leads to significant measurement error. It equals 0.06dB even for $F_x = 4$ dB and $F_x G_{Ax} = 20$ dB. For the easily achieved value of $F_r = 10$ dB this error increases to 0.37dB. Calculating the F_x value by means of (1) requires knowledge of the output impedance (or reflection coefficient Γ_g) of the noise generator, the scattering matrix $[S_x]$ of the TUT (for calculating the values of $G_{Ax}(\Gamma_g)$ and the output reflection coefficient Γ_{2x} of the TUT) and the four noise parameters of the measurement receiver (for the calculation of $F_r(\Gamma_{2x})$). By integrating S-parameter measurements with the noise figure measurement (preferably in the same test set), all of the necessary impedances and parameters can be measured to enable the calculation of a more correct noise figure [1]—[6].

The simultaneous determination of both noise figure and available gain of the TUT through noise figure and reflection coefficient measurements only is described in many of Sannino's works, presented in [7] and [8] and references therein. Expanded computer-controlled measuring system contains a programmable step attenuator inserted into the well-matched receiver input. The measurements of F_m are performed for different values of $F_r(\Gamma_g)$, obtained by acting on the high-repeatability step attenuator. The values of $F_x(\Gamma_g)$ and $G_{Ax}(\Gamma_g)$ are computed from linear equation system derived from (1) [8]. The four noise parameters and four gain parameters as well as all the scattering parameters of the TUT can be determined by means of this measurement methodology [7].

A direct-reading noise figure meter, where true value of TUT noise figure is determined using the relation (1), without S-parameter measurements, is described in [9]. The method's error is caused by the fact that insertion gain $G_{Ix}(\Gamma_g)$ and noise figure $F_r(\Gamma_g)$ are utilized instead of available gain $G_{Ax}(\Gamma_g)$ and $F_r(\Gamma_{2x})$ [9], [10]. A way to decrease the error consists of inserting a low-noise, high-gain preamplifier at the receiver's input [10].

For measurement methods with elimination of error caused by measurement receiver noise figure [11] the relation (1) simplifies to the formula

$$F_m = F_x \sqrt{(\Gamma_g)} - \frac{1}{G_{Ax} \sqrt{(\Gamma_g)}} \quad (2)$$

The difference $F_x [dB] - F_m [dB]$ does not exceed 0.1 dB as long as $F_m G_{Ax} \geq 16.4$ dB. Measurement system contains a two-state output power regulator (a switch) at the receiver input. As a result of the gated operation of the noise generator and the switch, there are different states created in the measurement system, when different power levels appear in receiver of measurement system. Two or three power levels can be used in a principle of operation of the measurement system [11]. An automatic direct-reading noise figure meter, where the F_x value is calculated using relation (2), and the

insert
The r
meter

F
as we
in the
figure
metho
this re

TI
on me
can be
noise
excess
combin
by a sp
actual
the add
than 0.

In
ment w
fourth s
is appli

The
used as
front-en
the swit
„c” state
ter micro
appropri
from „p
switch M
receiver
„c” repre
OFF (ma
while the

insertion gain instead of the available gain of the TUT has been used, can be made [12]. The method error caused by an inaccurate available gain determination exists, but the meter has a lower ($F_r - 1$ times) noise figure sensitivity to the TUT gain uncertainty.

Fulfilling the equality $F_m = F_x$, regardless of the measurement receiver noise figure as well as TUT gain value, can be achieved with any method where equation (2) is used in the case while at the TUT output a preamplifier is inserted in cascade, having noise figure and available power gain satisfying the following relation: $F_p = 1 + G_{Ap}^{-1}$. The method is limited by technical difficulties in realization of the preamplifier, to fulfil this relation in the frequency band.

Three-state or two-state methods with elimination of the receiver noise influence on measurement results as well as the conventional constant output power method can be modified to achieve the equality $F_m = F_x$, when two (main and additional) noise generators are used in the measurement systems [11]. A signal with a certain excess noise ratio (ENR) from additional noise generator is directed by a passive power combiner to the receiver input. It increases the noise figure of measurement system by a specified value in some operation states, depending on the measurement method actually used. For the system calibration error equal to 0.5 dB, while the ENR value of the additional noise generator is adjusted, the difference $F_x [dB] - F_m [dB]$ is smaller than 0.1 dB for $F_m G_{Ax} \geq 6.8 \text{ dB}$ [11].

In this paper, a new method is presented for direct-reading noise figure measurement with the equality $F_m = F_x$ fulfilled. It is a modified three-state method where a fourth state is introduced in the measurement procedure and a new calculation formula is applied.

2. DESCRIPTION OF THE NOISE MEASUREMENT SET-UP

2.1. CONFIGURATION AND PRINCIPLE OF OPERATION

The configuration of the measurement set-up is shown in Fig. 1. Two switches are used as the two-state power level regulators. One of them is inserted at the receiver front-end input and the second — at the detector two-port input. Their „p”, or „i” (for the switch P) and „m”, or „n” (for the switch M) operating states, as well as „h” or „c” states of the noise generator, are changed in cycles according to the output meter microprocessor program. For „p” and „m” states, the signals are passed through appropriate switches with small attenuation values. If the switch P state is changed from „p” to the „i” state, the attenuation will be increased α times. Similarly, the switch M, in its „n” state, increases the attenuation of the signal transmitted from the receiver front-end output to the detector two-port input β times. The states „h” and „c” represent the states when the noise generator is turned ON (hot-noise source) and OFF (matched termination at room temperature as a cold-noise source), respectively, while the ENR of the noise generator is changed from Ψ_h to Ψ_c values, respectively.

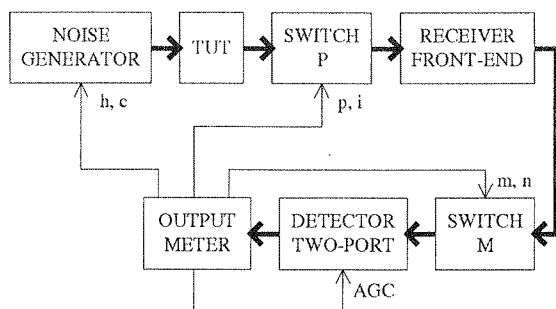


Fig. 1. Simplified scheme of the noise figure measurement set-up.

—— measurement signals, ——— control signals

Rys. 1. Uproszczony schemat zestawu do pomiaru współczynnika szumu

—— sygnały pomiarowe, ——— sygnały sterujące

Four measurement set-up states are used in a measurement procedure. The noise power levels appearing at the noise power detector input of the detector two-port are shown in Fig. 2. The subscripts refer to the noise generator as well as the switch states. For example, the power level N_{cim} occurs when the noise generator is turned OFF, the switch P is in its „i” state, and the switch M is in its „m” state.

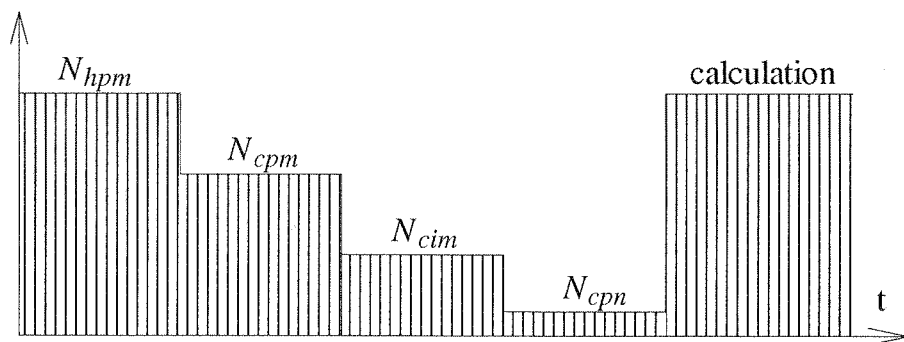


Fig. 2. Noise power levels appearing in cycles at the noise power detector input

Rys. 2. Poziomy mocy szumu pojawiające się cyklicznie na wyjściu detektora mocy szumu

The measured value of noise figure is calculated from the formula

$$F_m = \Psi_h \frac{N_{cpm} - N_{cim} + F_{rk}^{-1}(N_{cim} - N_{cpm})}{N_{hpm} - N_{cpm}}, \quad (3)$$

where F_{rk} is the noise figure of the switch P and receiver front-end in cascade. The F_{rk} value is measured and entered into the microcomputer memory during the calibration procedure of the measurement system. To perform the system calibration, the noise generator is directly connected to the input of the switch P which is in its „p” state

throughout the calibration procedure. The three-state noise figure measurement method is used and the noise figure F_{rk} value is calculated from the formula [11], [12]

$$F_m = \Psi_h \frac{N_{cpm} - N_{cpn}}{N_{hpm} - N_{cpm}}. \quad (4)$$

The receiver front-end has a sufficient high available power gain, so $F_m = F_{rk}$.

2.2. REALIZATION OF THE MEASUREMENT SET-UP BLOCKS

A typical microwave noise generator can be used in measurement set-up.

It is profitable to use a p-i-n diode SPDT switch as a fast and broad-band microwave switch. Fig. 3 shows a schematic diagram of the microwave part of this switch. A pole-port is used as an RF output port, while one of the throw-ports is used as an

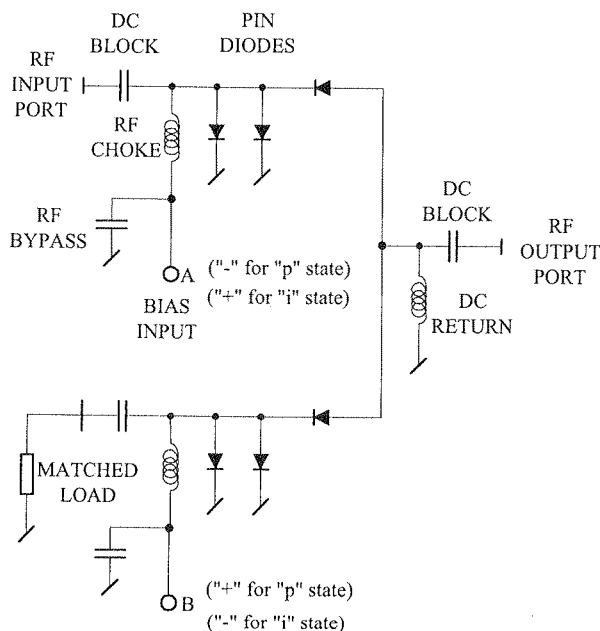


Fig. 3. Schematic diagram of the microwave part of the switch P

Rys. 3. Schemat części mikrofalowej przełącznika P

RF input port. The second throw-port is terminated with an impedance-matched load. Just the same numbers of the p-i-n diodes are forward biased and reverse biased for the „p” state as well as for the „i” state. Due to such configuration of the diode circuits and biasing of them, the noise signal generated by p-i-n diode circuits is just the same in both the „p” and „i” switch states. The switch can be presented as a two-state passive two-port having an additional noise generator in its RF output port. The ENR of this

generator has a constant and smaller than 0.1 value [13] for both operating states of the switch.

The receiver front-end is a part of the microwave receiver superheterodyne system with the bulk of the gain at the intermediate frequency (IF).

The switch M can be constructed in the same way as the switch P described above. It is easy to realize the switch M, because it is a narrow-band two-port inserted between the IF amplifiers. A typical gain modulator, predestined for the receiver's IF chain, is suitable for this application.

The detector two-port consists of an IF amplifier controlled by automatic gain control loop (AGC), a noise power detector and an analog-to-digital converter (ADC) at the end. The AGC has been applied to control the maximum input signal of the noise power detector. Typical noise power detector circuit [9] converts RF noise signals to dc levels that the microprocessor can interpret by means of the ADC.

The output meter contains a noise figure indicator, a microprocessor, and power supplies for the entire measuring system. The indicator displays the measurement results of $F_m[dB] = 10\log F_m$ in digital form, computed according to expression (3), on the front panel of the meter.

3. MEASUREMENT METHOD ERROR ANALYSIS

A method error is caused by the actual parameters of the measurement system blocks and a system calibration error. To calculate the error components, the power levels in formula (3) are described with a generalized expression for output noise power of the cascade of n two-ports at ambient temperature $T_a = T_0 = 290K$ [11].

$$N = \left[(\Psi_1 + F_1) \prod_{i=1}^n + \sum_{i=2}^n (\Psi_i + F_i - 1) \prod_{j=i}^n G_{A_j} \right] k T_0 B \quad (5)$$

Here symbols $\prod$ and $\sum$ stand for the product and sum of the components, k is Boltzmann's constant, T_0 is the standard noise temperature, and B is a noise bandwidth of the cascade. F_i and G_{A_i} are the spot noise figure and available power gain of the two-port, respectively, with number i at its input frequency. They are the values for generator impedance equal to the output impedance of the two-port with number $i - 1$. Each Ψ_i is the sum of the ENR of the noncorrelated noise signal directed to the input of the two-port number i , excluding the signals generated by two-port number i .

The power level N_{hpm} can be expressed as follows:

$$\begin{aligned} N_{hpm} = & [(\Psi_h + F_x) \frac{G_{Ax} G_{Arp} G_{Adm}}{A_{Ap} A_{Am}} + \\ & + (A_{Ap} - 1) \frac{G_{Arp} G_{Adm}}{A_{Ap} A_{Am}} + (\Psi_p + F_{rp} - 1) \frac{G_{Arp} G_{adm}}{A_{Am}} + \end{aligned}$$

$$+ (A_{Am} - 1) \frac{G_{Adm}}{A_{Am}} + (\Psi_m + F_{dm} - 1) G_{Adm} k T_0 B, \quad (6)$$

where $G_{Arp} = G_{Ar}(\Gamma_{2p})$ and $F_{rp} = F_r(\Gamma_{2p})$ are the available power gain and noise figure (single-or double-sideband, according to the construction of the receiver), respectively, of the receiver front-end for „p” state of the switch P. Γ_{2p} is an output reflection coefficient of the switch P in its „p” state. Similarly, $G_{Adm} = G_{Ad}(\Gamma_{2m})$ and $F_{dm} = F_d(\Gamma_{2m})$ are the available power gain and noise figure, respectively, of the IF amplifier of the detector two-port while the switch M is in its „m” state. Γ_{2m} is an output reflection coefficient of the switch M in its „m” state. A_{Ap} and Ψ_p are the available power attenuation of the switch P and ENR of the noise generator appearing at the RF output port of the switch P in its „p” state, respectively. Similarly, A_{Am} and Ψ_m stand for the same values of the switch M in its „m” state.

The power level N_{cpm} appears when the noise generator is turned OFF and it can be expressed by (6) for $\Psi_h = \Psi_c = 0$.

The power level N_{cim} appears when the switch P state „p” is changed to the „i” state. This power level can be expressed by (6) for $\Psi_h = 0$ and when the values of A_{Ap} , Ψ_p , G_{Arp} , and F_{rp} , are replaced by A_{Ai} , Ψ_i , $G_{Ari} = G_{Ar}(\Gamma_{2i})$, and $F_{ri} = F_r(\Gamma_{2i})$, respectively, where Γ_{2i} , is the output reflection coefficient of the switch P in its „i” state.

The power level N_{cpn} appears when the switch M state „m” is changed to the „n” state. This power level can be expressed by (6) for $\Psi_h = 0$ and when the values of A_{Am} , Ψ_m , G_{Adm} , and F_{dm} , are replaced by A_{An} , Ψ_n , $G_{Adn} = G_{Ad}(\Gamma_{2n})$, and $F_{dn} = F_d(\Gamma_{2n})$, respectively, where Γ_{2n} is the output reflection coefficient of the switch M in its „n” state.

Inserting, expressed as above, the power levels N_{hpm} , N_{cpm} , N_{cim} , and N_{cpn} into (3) we obtain

$$\begin{aligned} F_m &= F_x \left[1 - \frac{1}{\alpha} + \frac{1}{F_{rk}} \left(\frac{1}{\alpha} - \frac{1}{\beta} \right) \right] + \frac{\frac{1}{\beta} - \frac{1}{\alpha} - \frac{A_{Ap}}{\beta} (\Psi_p + F_{rp})}{F_{rk} G_{Ax}} + \\ &+ \frac{\frac{1}{\alpha} + A_{Ap} (\Psi_p - \Psi_i g_r) + A_{Ap} (F_{rp} - F_{ri} g_r)}{G_{Ax}} + \\ &+ A_{Ap} \frac{\frac{1}{\beta} - 1 + A_{Am} (\Psi_m - \Psi_n g_d) + A_{Am} (F_{dm} - F_{dn} g_d)}{F_{rk} G_{Ax} G_{Arp}} = \\ &= \frac{F_{rk} - A_{Ap} (\Psi_i + F_{ri}) g_r}{F_{rk} G_{Ax}}, \end{aligned} \quad (7)$$

$$\text{with } \alpha = \frac{A_{Ai}}{A_{Ap} g_r}, \quad g_r = \frac{G_{Ari}}{G_{Arp}}, \quad \beta = \frac{A_{An}}{A_{Am} g_d}, \quad g_d = \frac{G_{Adn}}{G_{Adm}}.$$

It is easy to get α and β values higher than 1000 [14], thus the expression in square brackets of the first term of 7) can be assumed as equal to 1. For the same reasons the second term of (7) can be neglected as a very small value.

The term g_r denotes a ratio of the available gains of the receiver front-end while the output reflection coefficient of the switch P is changed from Γ_{2i} to the Γ_{2p} value. Thus

$$g_r = \frac{(1 - |\Gamma_{2i}|^2)1 - \Gamma_{2p}S_{11r}|^2}{(1 - |\Gamma_{2p}|^2)1 - \Gamma_{2i}S_{11r}|^2}, \quad (8)$$

where S_{11r} equals an input reflection coefficient of the high-gain receiver front-end. For a well impedance-matched receiver front-end (or having a ferrite isolator at its input) the value of $S_{11r} \cong 0$. From the schematic diagram of the switch P, shown in Fig. 3, and its principle of operation, it is evident, that the output reflection coefficients Γ_{2p} and Γ_{2i} of the switch P depend on the output reflection coefficient Γ_{2x} of the TUT and the reflection coefficient Γ_l of the matched load, which terminate the second throw-port of the switch, respectively. For $\Gamma_l \cong \Gamma_{2x}$, we have $\Gamma_{2i} \cong \Gamma_{2p}$, and we have seen, that we get $g_r \cong 1$, and $F_{ri} \cong F_{rp}$. We have $\Psi_p \cong \Psi_i < 0.1$. Thus, the numerator of the third term of (7) has a very small value, and the third term of (7) can be neglected as a very small value.

Similarly as above, discussion of the parameters of the switch M and detector two-port shows that the numerator of the fourth term of (7) can be assumed as equal to $-A_{Ap}$, but the $F_{rk}G_{Ax}G_{Arp}$ value can be much larger than 1000. If so, the fourth term of (7) can also be neglected as a very small value.

Finally the last term of (7) has to be analyzed. It contains the noise figure F_{rk} which is measured during the calibration procedure with some measurement error $\delta_k \cong 1$. Noise figure F_{rk} can be described as follows:

$$F_{rk} = \delta_k A_{Ap}(\Psi_p + F_{rpg}), \quad (9)$$

where $A_{Ap} = A_{Ap}(\Gamma_{2pg})$ and $F_{rpg} = F_r(\Gamma_{2pg})$. Γ_{2pg} is an output reflection coefficient of the switch P in its „p” state, when noise generator is directly connected to the switch P.

$$A_{Ap} = \frac{|1 - S_{11p}\Gamma_g|^2(1 - |\Gamma_{2pg}|^2)}{(1 - |\Gamma_g|^2)S_{21p}|^2}, \quad (10)$$

$$\Gamma_{2pg} = S_{22p} + \frac{S_{12p}S_{21p}\Gamma_g}{1 - S_{21p}\Gamma_g}, \quad (11)$$

where S_{ijp} are the scattering parameters of the switch P in its „p” state. Usually $A_{Ap} = \mu A_{Ap}$, where $\mu \cong 1$ when $\Gamma_g \cong \Gamma_{2x}$. Schematic diagram of the switch P (Fig. 3) shows, that, if $\Gamma_g \cong \Gamma_l$, we have $\Gamma_{2pg} \cong \Gamma_{2i}$ and $F_{rpg} \cong F_{ri}$. For $\Psi_p \cong \Psi_i$ the relation (9) can be rewritten as

$$F_{rk} = \delta A_{Ap}(\Psi_i + F_{ri})g_r, \quad (12)$$

where $\delta = \delta_k \mu / g_r$. Placing (12) into (7), neglecting in it the previously discussed terms of very small values, we obtain

$$F_m = F_x - \frac{1 - \delta^{-1}}{G_{Ax}}. \quad (13)$$

while
value.

The measurement method error is given by

$$\Delta F_{\delta}[\text{dB}] = 10 \log \left(1 - \frac{1 - \delta^{-1}}{F_x G_{Ax}} \right) \quad (14)$$

For $\delta = -0.5\text{dB}$, and $F_x G_{Ax} \geq 7.2\text{dB}$, the error $\Delta F_{\delta}[\text{dB}]$ is smaller than 0.1dB .

4. VERIFICATION AND EXPERIMENTAL RESULTS

To assess the reliability of this noise figure measurement method, the noise figure of a low-gain amplifier has been measured in the 2-3.9GHz frequency range. The insertion gain G_{Ix} [dB] value as well as the output reflection coefficient module $|\Gamma_{2x}|$ value of the amplifier are shown in Table I.

Table 1

Insertion gain and output reflection coefficient of the amplifier and its noise figure measurement results displayed while using the different measurement methods

Wzmocnienie wtarceniowe i wyjściowy współczynnik odbicia wzmacniacza oraz jego wyniki pomiarów współczynnika szumu, pokazywane na wskaźniku miernika wykorzystującego różne metody pomiarowe

f[GHz]	2.0	2.3	2.6	3.3	3.9
$G_{Ix}[\text{dB}]$	9.6	8.9	8.5	6.1	3.4
$ \Gamma_{2x} $	.065	.060	.060	.090	.135
$F_{rk}[\text{dB}]$	8.4	7.6	8.5	8.4	9.2
$F_{mc}[\text{dB}]$	5.2	5.3	5.7	6.6	8.9
$F_{xc}[\text{dB}]$	4.2	4.3	4.5	5.1	6.6
$F_{m3}[\text{dB}]$	3.9	4.1	4.2	4.7	5.9
$F_{x3}[\text{dB}]$	4.1	4.3	4.4	5.1	6.4
$F_{xg}[\text{dB}]$	4.1	4.3	4.4	5.1	6.2
$F_{x4}[\text{dB}]$	4.2	4.3	4.4	5.1	6.4

A suitable computer program for four-state method has been implemented into noise figure measurement system described in [12]. Direct-reading measurement results of the noise figure values, measured for the different measurement methods used, are presented in Table I, where the noise figure indicator readings are marked as follows:

- F_{rk} — while performing the calibration procedure,
 F_{mc} — while using the conventional method without elimination of the influence of the measurement receiver noise figure as well as low TUT gain on noise figure value displayed by the noise figure indicator,

- F_{xc} — while using the conventional method with automatic correction of the second stage error as described in [9]
- F_{m3} — while using the three-state method with elimination of error caused by measurement receiver noise figure only (see the relation (2)),
- F_{x3} — while using the three-state method with automatic correction of the error caused by low gain of the TUT [12],
- F_{xg} — while using the three-state method with additional noise generator providing elimination of measurement system noise influence as well as low TUT gain influence on the measurement results [11],
- F_{x4} — while using the verified four-state method.

The F_{mc} value made the true noise figure values too high whereas the F_{m3} values made them too low. We have seen a very good agreement of the F_{x4} values with the F_{x3} as well as F_{xg} values.

5. CONCLUSION

The measured noise figure F_m is determined by means of the noise power level difference calculation and the ratio calculation of these differences. Due to such calculations the measurement system has an improved external noise immunity. Additionally, the construction of the noise power detector is simplified. It is so because the direct current component of the detector output signal, which is independent of its input noise signal, is eliminated by subtracting the signals.

The adaptation of the four-state method into the measurement system, usually used for the three-state method, does not require any additional devices to develop this system. The switch M in the three-state method has to be employed while the noise figure of the microwave mixer under test is measured in this measurement system.

The system calibration procedure is simple and moderately time-consuming. The broad-band and continuous in the frequency band calibration is performed due to computer linear approximation of the F_{rk} value between the calibration points.

Mathematical relations presented in the paper allow us to calculate the error components caused by parameters of the nonideal blocks of the measuring system. It is easy to satisfy the suitable values of the noise signals and attenuation increases (α and β) of the switches (P and M, respectively) as well as the gain G_{Ar} value of the receiver front-end, so that the error components, caused by the actual parameters of these blocks, can be neglected. The error caused by the acceptance of $\Psi_c = 0$ for the true temperature conditions as well as the other known error components (caused by the ENR uncertainty, microwave connector losses, and impedance mismatch at the TUT input) are the same as in all well-known measurement methods.

The results show that a simple measurement method can be employed with a good degree of accuracy whenever rapid noise testing of low-gain TUT is needed. This

method can be applied to determinate the four noise parameters, while a tuner is inserted into noise generator output and well-known measurement and calculating methodology with measurement of reflection coefficient is used.

6. REFERENCES

1. A. C. Davidson, B. W. Leake, E. Strid: *Accuracy improvements in microwave noise parameter measurements*. IEEE Trans. Microwave Theory Tech. 1989, vol. 37, pp. 1973–1977
2. J. W. Archer, R. A. Batchelor: *Fully automated on-wafer noise characterization of GaAs MESFET's and HEMT's*. IEEE Trans. Microwave Theory Tech. 1992, vol. 40, pp. 209–216
3. R. Benelbar, B. Huyart, R. G. Bosisio: *Microwave noise characterization of two-port devices using an uncalibrated tuner*. IEEE Trans. Microwave Theory Tech. 1996, vol. 44, pp. 1725–1728
4. P. Beand, L. Roy, S. Labonte, M. Stubbs: *An enhanced on-wafer millimeter-wave noise parameter measurement system*. IEEE Trans. Instrum. Meas. 1999, vol. 48, pp. 825–829
5. A. Lazaro, L. Pradell, J. M O'Callaghan: *FET noise-parameter determination using a novel technique based on 50-Ω noise-figure measurements*. IEEE Trans. Microwave Theory Tech. 1999, vol. 47, pp. 315–324
6. S. Van den Bosch, L. Martens: *Experimental verification of pattern selection for noise characterization*. IEEE Trans. Microwave Theory Tech. 2000, vol. 48, pp. 156–158
7. G. Martinez, M. Sannino: *The determination of the noise, gain and scattering parameters of microwave transistors (HEMT's) using only an automatic noise figure test-set*. IEEE Trans. Microwave Theory Tech. 1994, vol. 42, pp. 1105–1113
8. A. Di Poala, M. Sannino: *A novel noise figure and gain test set for microwave devices*. IEEE Trans. Instrum. Meas. 1999, vol. 48, pp. 921–926
9. H. L. Swain, R. M. Cox: *Noise figure meter sets records for accuracy, repeatability, and convenience*. Hewlett-Packard Jour. 1983, vol. 34, pp. 23–34
10. Noise figure measurement accuracy. Hewlett-Packard application note 57–2, Nov. 1988
11. W. Adamski: *Methods for direct-reading noise measurements of low-gain two-ports*. IEEE Trans. Instrum. Meas. 1994, vol. 43, pp. 411–414
12. W. Adamski, J. Dąbrowski, A. Janeczko, W. Rydzkowski: *The set-up for automatic noise figure measurement*. 9th Microwave Conf. MIKON'91, Rydzyna (Poland), 1991, pp. 408–412
13. Solid state SPDT microwave switch. Hewlett-Packard tentative date 15, Mar. 1968
14. R. Tenenholz: *Broadband MIC multithrow PIN diode switches*. Microwave Jour. 1973, vol. 16, pp. 25–30

W. ADAMSKI

CZTEROSTANOWA METODA BEZPOŚREDNIEGO POMIARU WSPÓŁCZYNNIKA SZUMU DWUWROTNIKÓW O MAŁYM WZMOCNIENIU

Streszczenie

Współczynnik szumu F_m pokazywany przez wskaźnik zestawu automatycznego miernika współczynnika szumu powinien być równy wartości prawdziwej współczynnika szumu F_x dwuwrotnika badanego, niezależnie od jego wzmocnienia mocy dysponowanej G_{Ax} i współczynnika szumu odbiornika pomi-

arowego F_r . We wstępie omówiono krótko problemy techniczne związane z realizacją pomiaru zapewniającego bezpośrednimi odczyt wartości prawdziwej współczynnika szumu F_x dwuwrotników o niewielkim wzmacnieniu przy stosowaniu znanych metodach pomiarowych. W miernikach wykorzystujących automatyczne wprowadzanie poprawek wskazań, wynikających z małego wzmacnienia, jest mierzone i wykorzystywane w obliczeniach tych poprawek wzmacnienie wtrąceniowe G_{I_x} zamiast wymaganego w tym przypadku wzmacnienia mocy dysponowanej G_{A_x} .

W pracy przedstawiono nową metodę, która umożliwia eliminowanie wpływu zarówno wartości współczynnika szumu F_r jak i wartości wzmacnienia G_{A_x} na wartość współczynnika szumu F_m pokazawanego na wskaźniku miernika. Jest to zmodyfikowana metoda trzystanowa, gdzie w procedurze pomiarowej wprowadzono czwarty stan układu pomiarowego oraz zastosowano nową postać równania służącego do wyznaczania wartości mierzonej współczynnika szumu F_m . Wartość ta jest obliczana (zależność 3) na podstawie pomiarów czterech pojawiających się cyklicznie w układzie pomiarowym poziomów mocy szumu oraz znajomości, wprowadzonych podczas kalibracji układu do pamięci procesora, wartości współczynnika szumu F_{rk} odbiornika pomiarowego układu i względnego przyrostu szumu Ψ_h zastosowanego w układzie generatora szumu. Cztery stany układu pomiarowego są ustalane na podstawie sterowania stanami (włączenie, wyłączenie) generatora szumu oraz dwóch regulatorów poziomu sygnału w torze pomiarowym, realizowanych najkorzystniej w postaci dwustanowych (przenoszenie, tłumienie) przełączników. Omówiono możliwości realizacji poszczególnych bloków układu pomiarowego oraz szczególnie zasady pracy i budowy diodowego przełącznika typu SPDT zapewniającego stałość w obu jego stanach pracy poziomu generowanego w nim sygnału szumowego, co jest pożądane ze względu na minimalizację błędów analizowanej metody pomiaru współczynnika szumu. Wyprowadzono wyrażenie (zależność 7), które służy do oszacowania wymagań technicznych dla dodatkowych podzespołów układu pomiarowego oraz do oceny błędów metody pomiaru. Łatwo jest zrealizować przełączniki diodowe, w których zachowany jest stały poziom generowanego w nich sygnału szumowego a przyrost tłumienia (α i β) przy przejściu od stanu przenoszenia do stanu izolacji jest wystarczająco duży oraz wykonać wzmacniacz p.c.z. zapewniający wystarczająco duże wzmacnienie $G_{A_{rp}}$ stopnia wejściowego odbiornika pomiarowego układu w celu zapewnienia w równaniu (7) wartości wyrażenia w nawiasach kwadratowych pierwszego składnika praktycznie równej jedności oraz zaniedbania jako wielkości bardzo małych składników drugiego, trzeciego i czwartego.

Rozważania teoretyczne zostały wsparte przykładowymi pomiarami współczynnika szumu wzmacniacza mikrofalowego. Wykonano je w zrealizowanym zestawie automatycznego miernika współczynnika szumu, opisanego w pracy [12], w którym zaimplementowano oprogramowanie zapewniające stosowanie różnych metod pomiaru współczynnika szumu. Uzyskano bardzo dobrą zgodność wyników pomiarów (tabela I) otrzymanych przy stosowaniu weryfikowanej metody czterostanowej z wynikami innych metod bezpośredniego pomiaru współczynnika szumu.

Zastosowanie czterostanowej metody pomiaru posiada wszystkie zalety wykorzystywania w układach pomiarowych właściwości obliczania wyników pomiarów na podstawie ilorazu różnic mierzonych sygnałów. Nie wymagało to również rozbudowy układowej zestawu miernika, gdyż przełącznik M, znajdujący się w torze wzmacniacza p.c.z. miernika, jest przewidziany dla stosowania trzystanowej metody pomiaru współczynnika szumu mieszaczy mikrofalowych.

Słowa kluczowe: pomiary szumu, współczynnik szumu, mierniki automatyczne

U
takimi
które b
często
ny w m
eksper
(ang. A
nych p
technik
padają
w spos
wych u

Optymalizacja szumowa wielokanałowych mieszanych układów scalonych na przykładzie układu RX64

PAWEŁ GRYBOS

Wydział Fizyki i Techniki Jądrowej, AGH,
al. Mickiewicza 30, 30-059 Kraków
E-mail: grybos@ftj.agh.edu.pl

Otrzymano 2003.10.03

Autoryzowano 2003.12.20

W artykule przedstawiono metodologię projektowania wielokanałowych specjalizowanych układów scalonych w technologii CMOS, ze zwróceniem szczególnej uwagi na optymalizację szumową stopni wejściowych i minimalizację przesłuchów w układach analogowo-cyfrowych. Szczegółową analizę sposobu projektowania przeprowadzono w oparciu o 64-kanałowy układ scalony RX64 do odczytu paskowych detektorów krzemowych używanych do detekcji niskoenergetycznego promieniowania X.

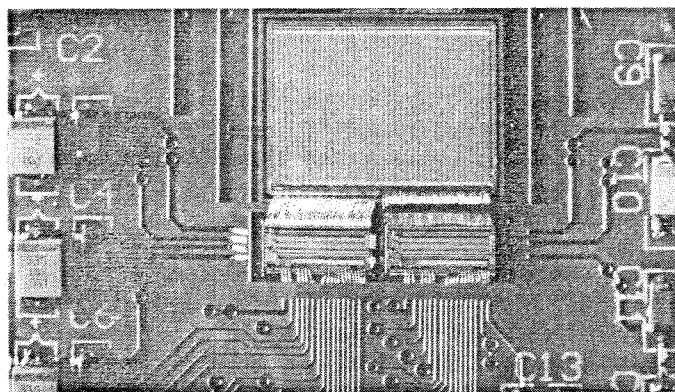
Słowa kluczowe: wielokanałowe układy CMOS VLSI, ASIC, szумы, układy mieszane.

1. WSTĘP

Układy scalone są dzisiaj najczęściej kojarzone z szybkimi układami cyfrowymi, takimi jak mikroprocesory czy bloki pamięci, oraz wielkimi firmami elektronicznymi, które będąc w posiadaniu technologii oraz wysokiej klasy zaplecza laboratoryjnego, często wyznaczają trendy rozwoju współczesnej elektroniki [1]. Jednak z drugiej strony w nauce i technice istnieje szereg gałęzi, które aby poszerzyć swoje możliwości eksperymentalne czy pomiarowe potrzebują tzw. specjalizowanych układów scalonych (ang. ASIC — Application Specific Integrated Circuits), wykonanych i zoptymalizowanych pod ich konkretne zastosowanie. Przykładem takich aplikacji są różnego rodzaju techniki obrazowania wykorzystujące matryce mikrosensorów (konwertujące kwanty padającego promieniowania na sygnał elektryczny), z których sygnał odczytywany jest w sposób równoległy, dla każdego mikrosensora niezależnie przy pomocy wielokanałowych układów typu ASIC. Zaletą tego typu obrazowania jest możliwość pracy z dużym

natężeniem promieniowania, a zatem uzyskiwanie obrazu w przeciągu krótkiego okresu czasu, z zachowaniem jego bardzo wysokich parametrów jeżeli chodzi o przestrzenną zdolność rozdzielczą i zakres dynamiczny, z jednoczesną możliwością dalszej obróbki cyfrowej obrazu [2,3].

W takich systemach, z punktu widzenia rozwiązań technicznych, krytycznym punktem są parametry wielokanałowego specjalizowanego układu scalonego służącego do odczytu matrycy mikrosensorów. Jako przykład omawianych technik obrazowania rozważmy zastosowanie paskowych detektorów krzemowych do pozycjoczułej detekcji niskoenergetycznego promieniowania X. Do odczytu detektorów stosuje się opracowane na Wydziale Fizyki i Techniki Jądrowej AGH wielokanałowe układy scalone o nazwie RX64 [4]. Fragment modułu stosowanego w omawianym przykładzie składającego się ze 128-paskowego detektora krzemowego i dwóch 64-kanałowych układów RX64 przedstawia rys. 1. Generowanie obrazu odbywa się w następujący sposób. Wpadający



Rys. 1. Fragment modułu zbudowanego z paskowego detektora krzemowego i dwóch układów scalonych RX64

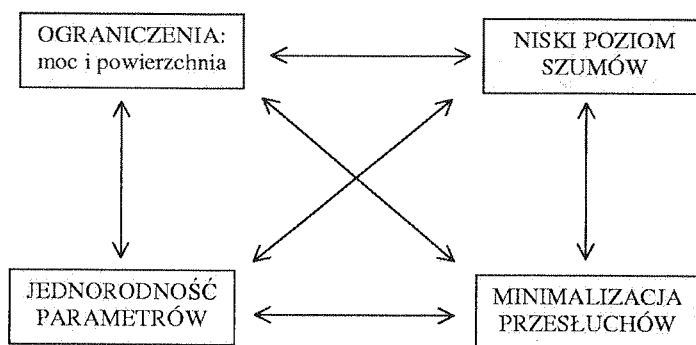
Fig. 1. Fragment of module consisting of silicon strip detector and two RX64 integrated circuits

do detektora niskoenergetyczny kwant, np. o energii 8 keV, generuje w detektorze krzemowym ładunek elektryczny, który zbierany jest na paskach detektora, w okolicach których wpadł kwant. Z elektronicznego punktu widzenia każdy pasek detektora stanowi impulsowe źródło sygnałów prądowych o czasie trwania około 20 ns i niosących ładunek około 2000 elektronów każdy. Impulsy z poszczególnych pasków są wzmacniane, formowane, a następnie zliczane w układzie scalonym RX64. Dokonując pomiaru w określonym okresie czasu, uzyskujemy obraz jednowymiarowy [5], a przy jednoczesnym zastosowaniu przesuwu mechanicznego obraz dwuwymiarowy [6,7].

Nawet tak krótka charakterystyka wskazuje na wymagania i ograniczenia które muszą spełniać omawiane tu specjalizowane układy scalone. Układy scalone służące do odczytu tego typu mikrosensorów powinny charakteryzować się następującymi wyróżnikami:

- wielokanałową architekturą, która narzuca bardzo ostre ograniczenia na rozpraszaną moc i powierzchnię pojedynczego kanału odczytowego,
- ze względu na małą amplitudę sygnałów wejściowych, konieczny jest bardzo niski poziom szumów własnych elektroniki odczytu oraz minimalizacja efektów przesłuchu w związku z mieszanym analogowo-cyfrowym charakterem układu scalonego,
- wymagana jest jednorodność parametrów analogowych (wzmocnienie, napięcia niezrównoważenia, szumy) dla wszystkich kanałów, co wymaga gruntownej analizy efektów niedopasowania.

W literaturze można znaleźć wiele opracowań dotyczących osobno każdego z wymienionych powyżej aspektów. Jednak w omawianej klasie układów wszystkie te aspekty są krytyczne i nakładają wzajemnie sprzeczne ograniczenia (patrz rys. 2), dlatego też należy rozwiązywać je łącznie, mając na uwadze parametry końcowe układu scalonego. Przykładowo, z punktu widzenia optymalizacji szumowej korzystnie jest, aby stopnie wejściowe toru odczytowego pracowały z dużym prądem, co niestety limitowane jest dopuszczalną rozpraszaną mocą w pojedynczym kanale. Pewne rozwiązania układowe korzystne z punktu widzenia optymalizacji szumów są wrażliwe na efekty przesłuchów. Z kolei ograniczenia na powierzchnię układu scalonego są sprzeczne z wymaganiami odnośnie minimalizacji szumów $1/f$, jak i redukcji efektów niedopasowania w układzie wielokanałowym. Krótkiego komentarza wymagają w tym kontekście



Rys. 2. Wzajemne zależności w omawianej klasie układów scalonych

Fig. 2. Relations between different aspects of ASIC design

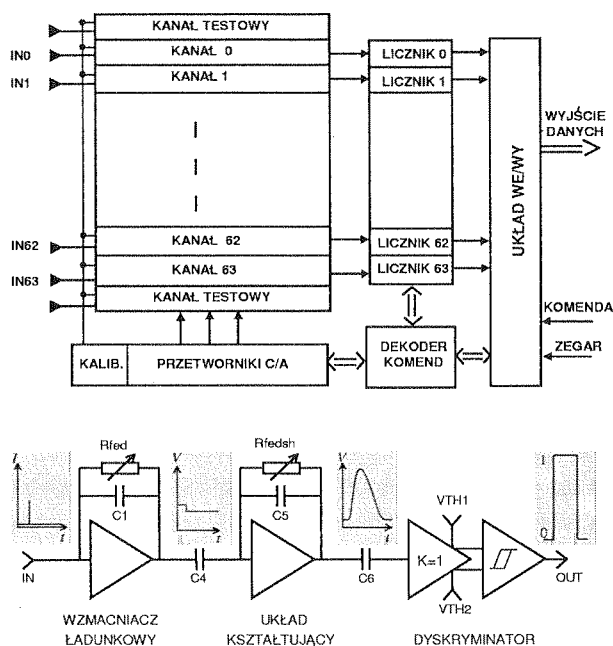
spotykane w literaturze przedmiotu rozwiązania układów wielokanałowych do odczytu tego typu czujników stosowanych w fizyce wysokich energii [8,9]. Układy te są dalekie od spełnienia wymagań dla zastosowań do obrazowania z użyciem niskoenergetycznego promieniowania rentgenowskiego ze względu na poziom szumów, wymagania dotyczące jednorodności parametrów analogowych w układach wielokanałowych i architekturę odczytu. W zastosowaniach detektorów krzemowych do detekcji promieniowania rentgenowskiego dostępny sygnał jest 10-krotnie mniejszy niż w zastosowaniach do detekcji cząstek relatywistycznych, co znacznie zaostrza wymagania dotyczące poziomu szumów i minimalizacji efektów przesłuchu. Po drugie, w dotychczasowych

układach stosowanych w eksperymentach fizyki cząstek i opisanych w literaturze korekty rozrzutów wzmacnień i napięć niezrównoważenia są wykonywane poza układem scalonym w systemie zbierania danych. Przy detekcji promieniowania rentgenowskiego rozwiązanie takie nie jest możliwe ze względu na statystyczny charakter sygnałów, a zatem konieczne jest rozwiązanie problemu efektów niedopasowania na poziomie układu scalonego.

Niniejsza praca ze względu na ograniczenia objętości koncentruje się na aspektach dotyczących optymalizacji szumowej i minimalizacji efektów przesłuchu w układzie mieszanym, z uwzględnieniem jednak ograniczeń na moc i powierzchnię, oraz wymagań jednorodności parametrów w układzie wielokanałowym omówionych we wcześniejszych pracach [10,11].

2. ARCHITEKTURA UKŁADU SCALONEGO RX64

Schemat blokowy układu RX64 przedstawia rys. 3a. Układ ten wykonano w technologii $0.8\ \mu\text{m}$ CMOS stosującej podłoże typu epi. Składa się on z 64 kanałów analogowych, 64 liczników, dekodera komend, układu kalibracyjnego, przetworników cyfrowo-analogowych i układów wejścia/wyjścia.



Rys. 3. Schematy blokowe układu RX64: a) cały układ, b) pojedynczy kanał analogowy

Fig. 3. Block diagram of RX64 IC: a) whole chip, b) single analogue channel

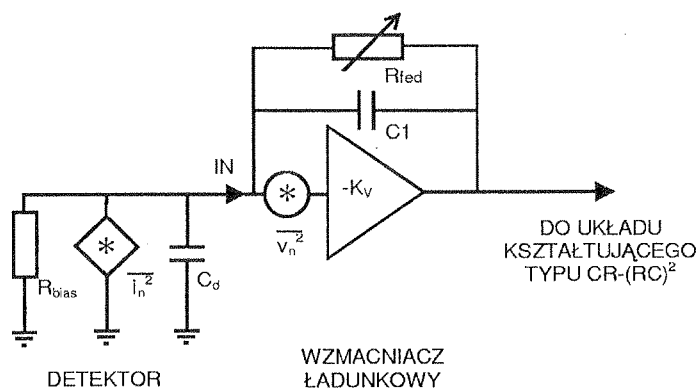
W pojedynczym kanale analogowym (patrz rys. 3b) wzmacniacz ładunkowy, w odpowiedzi na niosący ładunek Q_{in} kwazidirakowski, prądowy impuls z detektora, daje na swoim wyjściu skok napięcia o amplitudzie równej Q_{in}/C_1 . Kolejny stopień, jakim jest układ kształtujący, formuje impuls stosownie do wymagań czasowych, wzmacniając go i filtrując. Tak uformowany impuls podawany jest na dyskryminator napięcia, na wyjściu którego otrzymujemy odpowiedź binarną. Impulsy pojawiające się na wyjściu dyskryminatora są zliczane przez 20-bitowe liczniki. Dodatkowo zastosowane w układzie przetworniki cyfrowo-analogowe pozwalają na ustawienie parametrów filtracji w torach analogowych oraz progu dyskryminacji powyżej którego impulsy są zliczane. Wewnętrzny układ kalibracyjny pozwala na wstrzykiwanie do wejść kanałów analogowych impulsów testowych o regulowanej amplitudzie. Działanie całego układu kontrolowane jest przy pomocy wewnętrznego dekodera komend.

3. OPTYMALIZACJA SZUMOWA STOPNIA WEJŚCIOWEGO

Rozpoczynając analizę szumową, należy odpowiedzieć sobie na pytanie, które elementy stopnia wejściowego (przy założeniu odpowiednio dużego wzmocnienia tego stopnia) są istotne z punktu widzenia szumów oraz jakie szczególne ograniczenia dotyczące tego stopnia są narzucone przez konkretną aplikację — chodzi tu głównie o typ i parametry sensora, częstość impulsów wejściowych itp. W przypadku omawianego układu RX64 stopniem wejściowym jest wzmacniacz ładunkowy, po którym dla poprawy stosunku sygnału do szumu stosuje się dodatkowe układy kształtujące w postaci filtrów pasmowo-przepustowych. Rozważmy zatem przedstawiony schematycznie na rys. 4 szumowy schemat zastępczy wzmacniacza ładunkowego, który posiada w pętli sprzężenia zwrotnego kondensator C_1 i rezystor R_{fed} , z uwzględnieniem istotnych z punktu widzenia szumów elementów składowych detektora. Sygnał wyjściowy wzmacniacza ładunkowego, dla poprawy stosunku sygnału do szumu, poddany jest filtracji w układzie filtrów typu CR-(RC)². W schemacie zastępczym detektora uwzględniono rezystor polaryzujący R_{bias} oraz pojemność C_d , będącą sumą pojemności detektora i pojemności pasożytniczej połączeń detektora ze wzmacniaczem ładunkowym. Szumy prądowe, przedstawione na rys. 4, związane są z szumami termicznymi rezystorów R_{fed} i R_{bias} oraz szumami śrutowymi prądu upływu detektora I_{DET} i można je wyrazić formułą

$$\overline{di_n^2} = \frac{4kT}{R_{fed}} + \frac{4kT}{R_{bias}} + 2qI_{DET} \quad (1)$$

gdzie q oznacza ładunek elementarny, k stałą Boltzmana, a T temperaturę bezwzględną. Przy odpowiednim doborze parametrów detektora (mały prąd upływu detektora I_{DET} oraz duża wartość rezystora polaryzującego R_{bias}) dominujący wkład do szumów pochodzi od rezystora R_{fed} w pętli sprzężenia zwrotnego wzmacniacza i dwa pozostałe składniki pojawiające się we wzorze (1) mogą zostać pominięte. Natomiast przedsta-



Rys. 4. Szumowy schemat zastępczy wzmacniacza ładunkowego i detektora

Fig. 4. Noise schematic diagram of charge amplifier and detector

wione na rys. 4 napięciowe źródło szumów reprezentuje głównie szumy wzmacniacza, które można wyrazić formułą

$$\frac{dv_n^2}{df} = A + \frac{B}{f} + 4kTR_{con} \quad (2)$$

gdzie A jest gęstością widmową szumu białego wzmacniacza, B związana jest z szumem $1/f$ tego stopnia, natomiast R_{con} jest szeregową rezystancją pasożytniczą połączenia pomiędzy wzmacniaczem a sensorem. Zanim przejdziemy do szczegółowych wyrażeń na A i B prześledzimy jak wygląda ogólna analiza szumowa stopnia przedstawionego na rys. 4.

Zakładając powyższe gęstości widmowe szumów oraz określoną przepustowość filtru, można obliczyć wartość średniokwadratową napięcia szumów na wyjściu układu filtrującego $CR-(RC)^2$ [12]. W przypadku wzmacniacza ładunkowego przedstawionego na rys. 4 szumy odniesione do wejścia wyraża się jako tzw. równoważny ładunek szumów ENC (ang. *Equivalent Noise Charge*) mierzony w elektronach rms. Definicja ENC określa go jako taki ładunek, który wprowadzony na wejście wzmacniacza w postaci prądowego impulsu dirakowskiego daje na wyjściu filtru odpowiedź napięciową o amplitudzie równej średniokwadratowej wartości szumów [12]. W literaturze można spotkać tego typu obliczenia dla różnych sposobów filtracji [13,14]. Dla układu przedstawionego na rys. 4 ENC wyraża formułą

$$ENC^2 = ENC_w^2 + ENC_f^2 + ENC_i^2 \quad (3)$$

gdzie ENC_w reprezentuje wkład od napięciowych szumów białych

$$ENC_w^2 = \frac{0.847}{q^2} \frac{(C_d + C_1 + C_g)^2}{T_p} \times (A + 4kTR_{con}) \quad (4)$$

ENC_f odpowiada wkładowi od napięciowych szumów typu $1/f$

$$ENC_f^2 = \frac{3.41}{q^2} (C_d + C_1 + C_g)^2 \times B \quad (5)$$

natomiast ENC_i związane jest z wkładem od szumów prądowych

$$ENC_i^2 = \frac{0.635}{q^2} T_p \times \frac{4kT}{R_{fed}} \quad (6)$$

W powyższych równaniach T_p jest czasem kształtowania, definiowanym jako czas, po którym odpowiedź na wyjściu układu kształtującego osiąga maksimum (ang. *peaking time*), a C_g jest pojemnością bramki dla wejściowego tranzystora MOS. Nie znając jeszcze wewnętrznej struktury samego wzmacniacza ładunkowego (a zatem wyrażen na A i B oraz wartości C_g), zastanówmy się, jakie ograniczenia nałożone są na przedział zmienności parametrów C_d , C_1 , T_p i R_{fed} pojawiających się w równaniach (3)–(6).

Pojemność C_d — wartość tej pojemności jest zdeterminowana przez wymiary geometryczne sensora (od których zależy przestrzenna zdolność rozdzielcza całego elementu detekcyjnego) oraz sposób jego połączenia z układem scalonym. W przypadku czujnika, do którego odczytu jest przeznaczony omawiany układ scalony, pojemność pojedynczego sensora wynosi 2.5 pF. Ponieważ sensor łączony jest z układem scalonym przy pomocy cienkiego drutu aluminium metodą ultrakompresji, należy uwzględnić jeszcze pojemność do podłoża pola kontaktowego na układzie scalonym, która w tej technologii wynosi 0.8 pF, co daje łączną wartość pojemności $C_d = 3.3$ pF.

Czas kształtowania T_p — przy długich czasach T_p dominujący czynnik w wyrażeniu (3) na ENC pochodzi od szumów prądowych ENC_i (wzór (6)), które są wprost proporcjonalne do $\sqrt{T_p}$, podczas gdy przy krótkich stałych czasowych dominuje składowa od napięciowych szumów białych ENC_w , która jest odwrotnie proporcjonalna do $\sqrt{T_p}$ (wzór (4)). Zatem przy ustalonej konfiguracji detektora, można by poszukiwać takiego T_p , przy którym ENC przyjmuje wartość minimalną. Jednak zazwyczaj wymagania odnośnie natężenia rejestrowanego promieniowania limitują dopuszczalny czas kształtowania T_p , bowiem od niego zależy maksymalna dozwolona częstość impulsów wejściowych, przy której nie dochodzi do spiętrzenia impulsów na wyjściu układu kształtującego. W przypadku układu RX64, ze względu na natężenie promieniowania X rejestrowanego przez detektor półprzewodnikowy, czas kształtowania T_p zawężony jest do przedziału od 500 ns do 1 μ s. Pojemność C_1 — od jej wartości zależy wzmocnienie wzmacniacza ładunkowego. Wzmacniacz ładunkowy w odpowiedzi na kwazidierakowski impuls prądowy niosący ładunek Q_{in} daje na swoim wyjściu w przybliżeniu skok napięcia o amplitudzie równej Q_{in}/C_1 . Zatem im mniejsza wartość pojemności w sprzężeniu zwrotnym, tym większe wzmocnienie ładunkowe układu, a w konsekwencji wkład kolejnych stopni następujących po wzmacniaczu w całkowite szумы układu staje się do pominięcia. Pojemność ta wchodzi zresztą bezpośrednio w skład wyrażenia na ENC_w i ENC_f . Z drugiej jednak strony zmniejszanie tej pojemności limitowane

jest przez stabilność wzmacniacza oraz jednorodność wzmocnienia w układzie wielokanałowym. Mała wartość C_1 oznacza kondensator o niewielkiej powierzchni, co z punktu widzenia rozrzutów parametrów wielu takich kondensatorów rozmieszczonych na powierzchni wielokanałowego układu scalonego jest niekorzystne. Symulacje dotyczące stabilności i efektów niedopasowania pozwoliły na ustawienie pojemności kondensatora C_1 na poziomie 200 fF.

Rezystancja R_{fed} — pojawiająca się w pętli sprzężenia zwrotnego wzmacniacza, służy do rozładowania kondensatora C_1 po przyjściu impulsu prądowego. Jej duża wartość korzystnie wpływa na poziom szumów prądowych, natomiast nadmiernie wydłuża czas rozładowania kondensatora C_1 w pętli wzmacniacza po pojawieniu się impulsu na wejściu układu. Jako kompromis pomiędzy stałą czasową pętli sprzężenia zwrotnego wzmacniacza a wkładem R_{fed} w szumy prądowe, ustalono nominalną wartość rezystancji R_{fed} na poziomie 80 MΩ, z zaimplementowaną w układzie możliwością jej regulacji w zakresie od 20 MΩ do 100 MΩ. Rolę tej rezystancji pełni tranzystor PMOS o wymiarach $W_5/L_5 = 2\mu\text{m}/384\mu\text{m}$ pracujący w zakresie liniowym.

Jak widać z powyższej dyskusji, wartości pojemności C_d oraz czasu kształtowania T_p są bezpośrednio narzucone przez wymagania przyszłej aplikacji, natomiast w przypadku parametrów w pętli sprzężenia zwrotnego wzmacniacza C_1 i R_{fed} pole manewru jest mocno ograniczone. Przejdźmy zatem do struktury samego wzmacniacza, przy optymalizacji szumów którego projektant ma najwięcej do powiedzenia.

3.1. OPTYMALIZACJA STOPNIA WZMACNIACZA

Ze względu na wymagane duże pole wzmocnienia w otwartej pętli wzmacniacza, w przypadku scalonych wzmacniaczy ładunkowych często stosowaną konfiguracją jest układ zawiniętej kaskody, przedstawiony schematycznie na rys. 5. Na rysunku tym uwzględniono również źródła szumowe związane z poszczególnymi tranzystorami. Wzmocnienie napięciowe takiej konfiguracji w zakresie niskich częstotliwości dane jest w przybliżeniu formułą [4]

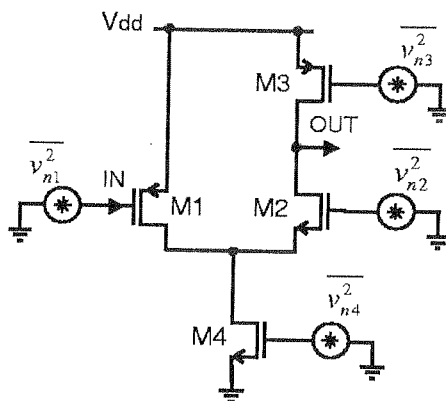
$$\frac{v_{out}}{v_{in}} \approx -\frac{g_{m1}}{g_{ds3}} \quad (7)$$

gdzie g_{m1} jest transkonduktancją tranzystora M1, a g_{ds3} konduktancją dren-źródło tranzystora M3. Szumy przeliczone na wejście układu można wyrazić formułą

$$\overline{\frac{dv_n^2}{df}} = \overline{\frac{dv_{n1}^2}{df}} + \frac{(g_{ds1} + g_{ds4})^2}{g_{m1}^2} \overline{\frac{dv_{n2}^2}{df}} + \frac{g_{m3}^2}{g_{m1}^2} \overline{\frac{dv_{n3}^2}{df}} + \frac{g_{m4}^2}{g_{m1}^2} \overline{\frac{dv_{n4}^2}{df}} \quad (8)$$

gdzie g_{m3} i g_{m4} są transkonduktancjami tranzystorów M3 i M4, a g_{ds1} i g_{ds4} są konduktancjami dren-źródło tranzystorów M1 i M4.

Aby oszacować wkład poszczególnych tranzystorów w ekwiwalentny szum napięciowy na wejściu wzmacniacza, należy zwrócić uwagę na możliwe wartości transkonduktancji i konduktancji dren-źródło pojawiających się w powyższym wyrażeniu.



Rys. 5. Układ zawiniętej kaskody z uwzględnieniem źródeł szumów

Fig. 5. Folded cascode configuration with noise sources

W układzie zawiniętej kaskody przedstawionej na rys. 5 dla uzyskania dużego wzmocnienia prąd płynący przez tranzystor M1 (a tym samym przez źródło polaryzujące M4) jest znacząco większy od prądów płynących przez tranzystory M2 i M3 (w przypadku układu RX64 stosunek tych prądów jest na poziomie 30). Ponadto tranzystor M1 charakteryzuje się dużym stosunkiem W_1/L_1 (gdzie W_1 i L_1 to odpowiednio szerokość i długość tranzystora M1), zatem g_{m1} jest znacząco większe od g_{m3} , g_{ds4} i oczywiście g_{ds1} . Przy powyższych założeniach formułę (8) można uprościć do postaci

$$\frac{\overline{dv_n^2}}{df} \approx \frac{\overline{dv_{n1}^2}}{df} + \frac{g_{m4}^2}{g_{m1}^2} \frac{\overline{dv_{n4}^2}}{df} \quad (9)$$

Ze wzoru (9) widać, że szумы tranzystora wejściowego M1 umiejscowione bezpośrednio na wejściu będą dominowały w układzie, podczas gdy składowa pochodząca od tranzystora M4 przy przeniesieniu na wejście układu jest redukowana ($g_{m4}/g_{m1} < 1$) z kwadratem stosunku odpowiednich transkonduktancji (tranzystor M1 posiada duży stosunek W_1/L_1 , podczas gdy tranzystor M4 pracuje jako źródło prądowe). Po ostatecznym ustawieniu wymiarów obu tych tranzystorów stosunek $g_{m4}/g_{m1} \approx 0.28$.

Na szумы wejściowe każdego z powyższych tranzystorów składają się zarówno szумы termiczne kanału, jak i szумы migotania. Szумы termiczne kanału tranzystora MOS pracującego w nasyceniu, opisuje formuła

$$\frac{\overline{dv_w^2}}{df} = 4kT\gamma \frac{g_m + g_{mb}}{g_m^2} = 4kT\eta\gamma \frac{1}{g_m} \quad (10)$$

gdzie $\eta = (g_m + g_{mb})/g_m$, g_m jest transkonduktancją tranzystora, g_{mb} transkonduktancją podłoża, γ to parametr zależny od punktu pracy tranzystora zmieniający się w zakresie od 1/2 do 2/3.

Szum migotania można wyrazić formułą

$$\frac{\overline{dv_f^2}}{df} = \frac{K_f}{C_{ox}^2} \frac{1}{WL} \frac{1}{f} \quad (11)$$

gdzie W i L są efektywną szerokością i długością tranzystora, C_{ox} jest pojemnością tlenku bramkowego liczoną na jednostkę powierzchni, a K_f współczynnikiem szumów migotania.

Porównując wzory (9)—(11) z ogólnym wyrażeniem (2), parametry A i B można zapisać jako

$$A = 4kT\eta_1\gamma_1 \frac{1}{g_{m1}} + \frac{g_{m4}^2}{g_{m1}^2} 4kT\eta_4\gamma_4 \frac{1}{g_{m4}} \quad (12)$$

$$B = \frac{K_{fp}}{C_{ox}^2} \frac{1}{W_1L_1} + \frac{g_{m4}^2}{g_{m1}^2} \frac{K_{fn}}{C_{ox}^2} \frac{1}{W_4L_4} \quad (13)$$

Zadanie minimalizacji szumów sprowadza się zatem do odpowiedniego doboru punktu pracy i wymiarów tranzystorów M1 i M4, tak aby uzyskać minimum dla wyrażań na ENC_w i ENC_f (wzory (4) i (5)). Patrząc na strukturę wyrażań (12) i (13), można przypuszczać, że dominujący wkład będzie pochodził od tranzystora wejściowego, dlatego też dobierzemy najpierw parametry tego elementu.

3.1.1. Dobór tranzystora wejściowego

Biorąc pod uwagę wkład do ENC tylko od tranzystora wejściowego M1, wyrażenia (4) i (5) można przepisać w postaci

$$ENC_{w1}^2 = \frac{0.847}{q^2} \frac{(C_d + C_1 + C_g)^2}{T_p} \times 4kT\eta_1 \frac{\gamma_1}{g_{m1}} \quad (14)$$

$$ENC_{f1}^2 = \frac{3.41}{q^2} (C_d + C_1 + C_g)^2 \times \frac{K_{fp}}{C_{ox}^2} \frac{1}{W_1L_1} \quad (15)$$

W układzie przedstawionym na rys. 5 tranzystorem wejściowym jest tranzystor PMOS. Można by zbudować konfigurację komplementarną, w której tranzystorem wejściowym byłby NMOS. Rozważmy jak dobór typu tranzystora wejściowego wpływa na szумы takiego stopnia. Z punktu widzenia szumów białych, przy ustalonym prądzie I_{DS1} korzystniej wypada tranzystor NMOS, który w zakresie silnej inwersji zapewnia większą transkonduktancję niż analogiczny tranzystor PMOS (o identycznych wymiarach i takim samym prądzie drenu). Na gruncie formuły (10) można porównać wkłady wnoszone przez analogiczne tranzystory NMOS i PMOS

$$\frac{(ENC_{w1})_{NMOS}}{(ENC_{w1})_{PMOS}} \approx \sqrt[4]{\frac{\mu_p}{\mu_n}} \quad (16)$$

gdzie μ_n , μ_p to odpowiednio ruchliwość dziur i elektronów w kanale. W przypadku stosowanej technologii powyższy stosunek jest na poziomie 0.8. Wkłady pochodzące od szumów migotania wnoszone przez odpowiedniego typu tranzystory mają się jak

$$\frac{(ENC_{f1})_{NMOS}}{(ENC_{f1})_{PMOS}} = \sqrt{\frac{K_{fn}}{K_{fp}}} \quad (17)$$

gdzie K_{fn} , K_{fp} to współczynniki dla szumów migotania odpowiednio dla tranzystorów NMOS i PMOS. W przypadku stosowanej technologii powyższy stosunek jest na poziomie 12 i znacznie korzystniej wypada tu tranzystor PMOS. Wydaje się zatem, że o wyborze typu tranzystora decydują wzajemne relacje pomiędzy wkładem wnoszonym do ekwiwalentnego ładunku szumów ENC poprzez ENC_w i ENC_f . Patrząc na strukturę wzorów (14) i (15) można powiedzieć, że przy dużym prądzie I_{DS1} (g_{m1} rośnie wraz ze wzrostem I_{DS1}) i długim czasie kształtowania T_p składowa ENC_{w1} maleje, podczas gdy ENC_{f1} pozostaje na stałym poziomie. Ale obok prądu I_{DS1} i czasu kształtowania T_p istotne są również wymiary tranzystorów i dopiero znając te parametry łącznie, można powiedzieć, który z omawianych składników dominuje.

Istnieją natomiast jeszcze inne przesłanki, dla których jako tranzystor wejściowy w technologii ze studnią typu n preferowany jest tranzystor PMOS. Chodzi mianowicie o przesłuchy w mieszanych układach analogowo-cyfrowych, dyskutowane szczegółowo w rozdziale czwartym. Dlatego w dalszej części pracy skoncentrujemy się na doborze prądu polaryzacji i optymalnych wymiarów tranzystora wejściowego, zakładając że jest to tranzystor PMOS.

Rozważmy prąd polaryzacji I_{DS1} tranzystora wejściowego, od którego zależy transkonduktancja g_{m1} pojawiająca się w wyrażeniu (14) na ENC_{w1} . W zakresie silnej inwersji g_{m1} dana jest formułą

$$g_{m1} = \sqrt{2\mu_p C_{ox} \frac{W_1}{L_1} I_{DS1}} \quad (18)$$

natomiast w zakresie słabej inwersji g_{m1} można wyrazić jako

$$g_{m1} = \frac{I_{DS1}}{n\varphi_T} \quad (19)$$

gdzie $\varphi_T = kT/q \approx 26$ mV w temperaturze 300 K, natomiast współczynnik n mieści się typowo w zakresie od 1 do 1.5 [15]. W obu przypadkach wraz ze wzrostem prądu I_{DS1} wzrasta transkonduktancja g_{m1} , zatem z punktu widzenia minimalizacji składowej ENC_{w1} korzystna jest praca z jak największym prądem, pod warunkiem że:

— nie spowoduje to znaczącego wzrostu temperatury układu, co wiązałoby się ze wzrostem poziomu szumów termicznych w układzie i szumów śrutowych detektora (w przypadku stosowanych w układzie detektorów krzemowych główna składowa prądu upływu pochodzi od prądu generacyjno-rekombinacyjnego, który w sposób eksponen-

— nie pojawiają się w tranzystorze MOS efekty związane z gorącymi nośnikami [15].

Jednak w omawianej klasie układów scalonych limity związane z maksymalną dopuszczalną mocą przypadającą na jeden kanał odczytowy są mocno ograniczone poprzez ilość tychże kanałów w obrębie pojedynczego układu scalonego. W przypadku układu RX64, aby uniknąć konieczności stosowania dodatkowego systemu chłodzącego, maksymalny dopuszczalny prąd I_{DS1} w tranzystorze wejściowym ustalono na poziomie $500 \mu A$, co przy nominalnym napięciu zasilania daje około 100 mW mocy rozpraszanej tylko w stopniach wejściowych wielokanałowego układu RX64.

Przejdźmy do doboru optymalnych wymiarów tranzystora M1 z punktu widzenia minimalizacji szumów. W literaturze przedmiotu można znaleźć wiele prac, które koncentrują się na tym zagadnieniu, jednak przy założeniu pracy tranzystora wejściowego w zakresie silnej inwersji [14,16]. W pracach tych przyjmuje się, że długość bramki tranzystora wejściowego ustawiona jest na wymiar minimalny L_{min} dopuszczalny w danej technologii i wtedy poszukuje się optymalnej szerokości tranzystora W_{opt} , od której zależy zarówno pojemność C_g , jak i transkonduktancja g_{m1} pojawiające się we wzorze (14) na ENC_{w1} .

Schemat wyznaczenia W_{opt} przedstawia się następująco. Przy założeniu pracy w zakresie silnej inwersji pojemność C_g można wyrazić jako [16]

$$C_g = \frac{2}{3} C_{ox} W_1 L_{min} \quad (20)$$

Wtedy transkonduktancję ze wzorów (18) i (20) można zapisać jako

$$g_{m1} = \sqrt{3\mu_p C_g \frac{1}{L_{min}^2} I_{DS1}} \quad (21)$$

a przyjmując $\gamma_1 = 2/3$, $\eta_1 = 1$, wyrażenie (14) na ENC_{w1} można zapisać w postaci

$$ENC_{w1}^2 = \frac{0,847 (C_d + C_1 + C_g)^2}{q^2 T_p} \times \frac{8}{3} kT \frac{1}{\sqrt{3\mu_p C_g \frac{1}{L_{min}^2} I_{DS1}}} \quad (22)$$

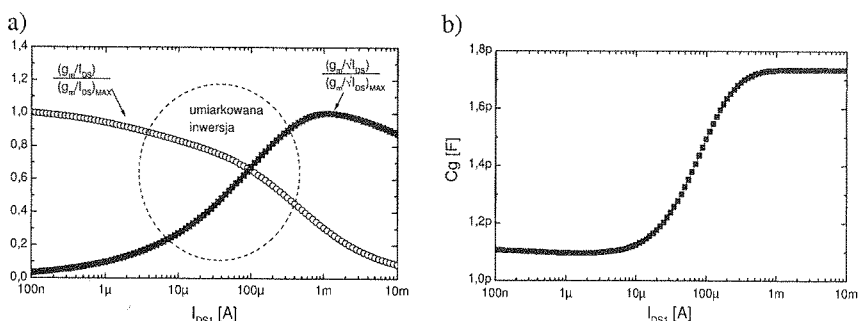
Łatwo wtedy pokazać, że ENC_{w1} przyjmuje wartość minimalną dla $C_g = (C_d + C_1)/3$ [16], a korzystając z równania (20) można znaleźć odpowiadającą tej pojemności optymalną szerokość W_{opt} daną wówczas formułą

$$W_{opt} = \frac{C_d + C_1}{2C_{ox} L_{min}} \quad (23)$$

Przyjęte w powyższym wyprowadzeniu założenia dotyczące formuł na g_{m1} , C_g i γ_1 nie zawsze są poprawne. A mianowicie, chcąc uzyskać niski poziom szumów, projektant stara się podnieść transkonduktancję tranzystora poprzez dobranie dużego stosunku

W_1/L_1 , przy jednocześnie istniejącym ograniczeniu na prąd I_{DS1} . Zatem efektywna gęstość prądu w kanale tranzystora jest nieduża i tranzystor zamiast w spodziewanej silnej inwersji pozostaje w zakresie umiarkowanej inwersji, gdzie formuły (18) i (20) przestają obowiązywać (zjawisko jest tym silniejsze im z nowszą technologią mamy do czynienia [17]). Ilustruje to rys. 6, który pokazuje zachowanie się transkonduktancji i pojemności bramki w zależności od prądu I_{DS1} dla tranzystora o szerokości kanału $W_1 = 1000 \mu\text{m}$ i długości $L_1 = 0.8 \mu\text{m}$ (technologia AMS0.8 μm).

Rysunki sporządzono w oparciu o wyniki symulacji z użyciem programu HSPICE i modeli tranzystorów BSIM3v3, przy założeniu, że $U_{GS} = U_{DS}$. Przedstawiając zachowanie się transkonduktancji, na rys. 6a pokazano stosunek $g_{m1}/\sqrt{I_{DS1}}$ (unormowany do wartości maksymalnej), który zgodnie z równaniem (18) powinien być stały w zakresie silnej inwersji, oraz stosunek g_{m1}/I_{DS1} (unormowany do wartości maksymalnej), który zgodnie z równaniem (19) powinien być stały w zakresie słabej inwersji. Zachowanie się pojemności (rys. 6b) wynika z klasycznej krzywej C-V dla struktury MOS [15], zgodnie z którą pojemność takiej struktury mierzona przy niskich częstotliwościach jest mniejsza w obszarze słabej i umiarkowanej inwersji w stosunku do pracy w zakresie silnej inwersji.



Rys. 6. Symulacje tranzystora PMOS o wymiarach $W_1/L_1 = 1000 \mu\text{m}/0.8 \mu\text{m}$ w funkcji prądu drenu I_{DS1} : a) zachowanie się transkonduktancji, b) pojemność widziana od strony bramki

Fig. 6. Simulations of PMOS transistor of $W_1/L_1 = 1000 \mu\text{m}/0.8 \mu\text{m}$ as a function of drain current I_{DS1} : a) behaviour of transconductance, b) gate capacitance

Jeżeli natomiast chodzi o wartość współczynnika γ we wzorze (10) to wiadomo, że w zakresie słabej inwersji współczynnik ten wynosi $1/2$, podczas gdy w zakresie silnej inwersji przyjmuje wartość $2/3$. Aby znaleźć zachowanie się tego współczynnika w zakresie umiarkowanej inwersji warto odwołać się do modelu EKV [18], który dla modelowania zakresu umiarkowanej inwersji używa odpowiedniej funkcji interpolacji. Zgodnie z tym modelem dla tranzystora pracującego w nasyceniu zachowanie się współczynnika γ można opisać funkcją [15]

$$\gamma = \frac{1}{2} + \frac{1}{6} \frac{u}{[f(u)]^2} \quad (24)$$

gdzie $f(u)$ dane jest formułą

$$f(u) = \frac{1}{2} (\sqrt{1 + 4u} + 1) \quad (25)$$

natomiast u definiowane jest jako $u = I_{DS}/I_S$, a I_S jest prądem charakterystycznym danym formułą

$$I_S = \frac{W}{L} \mu C_{ox} 2n \varphi_T^2 \quad (26)$$

Przykładowa wartość prądu nasycenia dla tranzystora, którego charakterystyki przedstawiono rys. 6, wynosi $77 \mu\text{A}$. Przyjmuje się, że dla $u = 1$ tranzystor pracuje w zakresie umiarkowanej inwersji, podczas gdy dla $u \ll 1$ pozostaje w strefie słabej inwersji, a dla $u \gg 1$ w zakresie silnej inwersji.

Reasumując można powiedzieć, że specyficzne wymiary tranzystora wejściowego i narzucone ograniczenia na pobór prądu powodują, że projektant powinien uwzględnić możliwość pracy zarówno w zakresie silnej, jak i umiarkowanej inwersji, a przy bardzo małym prądzie I_{DS1} nawet w zakresie słabej inwersji. Zatem według autora, aby znaleźć optymalne wymiary tranzystora, należy korzystać z wartości transkonduktancji g_{m1} i pojemności bramki C_g wyznaczonych dla danego punktu pracy w programie HSPICE, np. przy użyciu modeli BSIM3v3 [19]. Dla wyznaczenia współczynnika γ będziemy się posługiwać formułą (24), przy czym dla obliczenia prądu charakterystycznego wartość współczynnika n obliczamy z formuły (19) jako

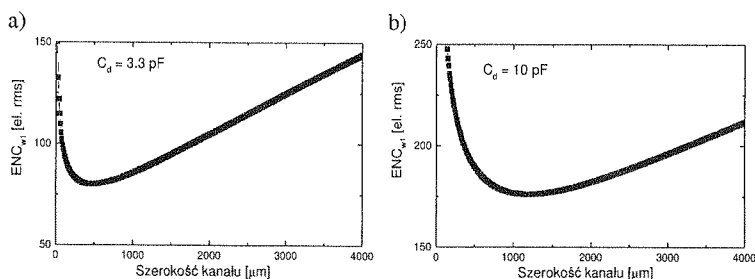
$$n = \frac{1}{\varphi_T (g_{m1}/I_{DS1})_{\max}} \quad (27)$$

gdzie $(g_{m1}/I_{DS1})_{\max}$ to maksymalna wartość stosunku transkonduktancji do prądu w zakresie słabej inwersji wyznaczona w programie HSPICE.

Rozpatrzmy problem doboru długości tranzystora L . Zagadnienie sprowadza się do odpowiedzi na pytanie, z jaką najmniejszą długością tranzystora można pracować, aby w układzie nie nastąpił wzrost szumów pochodzących od efektów krótkiego kanału. Okazuje się, że wybór minimalnej długości tranzystora niekoniecznie jest wyborem optymalnym i dotyczy to zarówno technologii starszych (Harris CMOS $1.2 \mu\text{m}$ [20]), jak i technologii najnowszych (IBM $0.18 \mu\text{m}$ [21]). Odpowiedź przekonywującą daje w tym przypadku jedynie pomiar i dlatego w projekcie omawianego układu wykonano zarówno wersję z długością kanału $1 \mu\text{m}$, jak i $1.5 \mu\text{m}$, tak dobierając szerokości tranzystorów, aby uzyskać ich jednakowe powierzchnie, a tym samym identyczny wkład od szumów $1/f$. Pomiarów dokonano przy takich samych prądach I_{DS1} i nie zauważono poprawy szumów w przypadku stosowania krótszych tranzystorów. Zdecydowano się zatem na wybór w układzie RX64 jako tranzystora wejściowego tranzystora o $L_1 = 1.5 \mu\text{m}$, wychodząc z założenia, że w takim tranzystorze przy identycznym napięciu U_{DS} natężenie pola wzdłuż kanału jest mniejsze, a po drugie zgodnie z pracą [22] wśród tranzystorów o jednakowej powierzchni, mniejszym rozrzutem charakteryzować

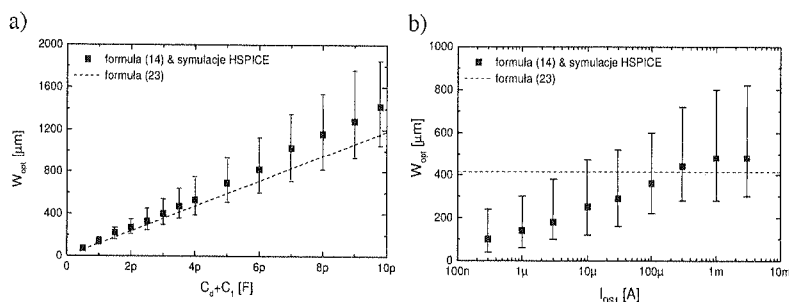
się będą tranzystory o dłuższym kanale, co jest istotne z punktu widzenia jednorodności parametrów układu wielokanałowego.

W zasadzie jedynym parametrem, jaki pozostaje do znalezienia, jest optymalna szerokość W_{opt} tranzystora M1, odpowiadająca minimalnej wartości ENC_{w1} . W oparciu o uzyskane z symulatora HSPICE wartości g_{m1} , g_{mb1} , C_g i obliczony współczynnik γ_1 można sporządzić wykres ENC_{w1} w zależności od szerokości bramki tranzystora. Przykłady takich zależności dla pojemności $C_d = 3.3$ pF i $C_d = 10$ pF przedstawiają odpowiednio rys. 7a i rys. 7b. Obliczenia wykonano dla tranzystora o długości kanału równej $1.5 \mu\text{m}$, $I_{DS1} = 500 \mu\text{A}$, $C_1 = 200$ fF oraz $T_p = 500$ ns. W obu przypadkach widać, że ENC_{w1} osiąga minimum, skąd określić można optymalną szerokość tranzystora W_{opt} z punktu widzenia szumów termicznych kanału. Porównując rys. 7a i 7b widać, że im większa wartość pojemności C_d , tym minimum jest bardziej płaskie.



Rys. 7. Wartość ENC_{w1} zgodnie ze wzorem (14) w funkcji szerokości kanału tranzystora wejściowego dla: a) $C_d = 3.3$ pF, b) $C_d = 10$ pF

Fig. 7. ENC_{w1} according to formula (14) as a function of transistor width for: a) $C_d = 3.3$ pF, b) $C_d = 10$ pF



Rys. 8. Optymalna szerokość tranzystora wejściowego w zależności od: a) sumy pojemności C_d i C_1 przy ustalonym prądzie $I_{DS1} = 500 \mu\text{A}$, b) w zależności od prądu I_{DS1} przy ustalonej wartości $C_d + C_1 = 3.5$ pF

Fig. 8. Optimal width of input transistor as a function of: a) sum of C_d and C_1 for $I_{DS1} = 500 \mu\text{A}$, b) drain current I_{DS1} for $C_d + C_1 = 3.5$ pF

Fakt ten uwzględniono na rys. 8, który przedstawia wyznaczoną zależność W_{opt} od sumy pojemności C_d i C_1 , z zaznaczeniem przedziału dla szerokości kanału tranzystora, dla której wartość ENC_{w1} jest wyższa jedynie o 3% od wartości minimalnej. Wykres wykonano dla $L_1 = 1.5 \mu m$, $C_1 = 200$ fF oraz $T_p = 500$ ns i $I_{DS1} = 500 \mu A$. Dla porównania ze stosowaną wcześniej metodą znajdowania W_{opt} [16], na rys. 8 linią przerywaną umieszczono również przewidywania tej teorii zgodnie ze wzorem (23).

Widać, że ze względu na optymalizację składowej ENC_{w1} preferowane są nieco większe szerokości kanału tranzystora niż wynikałoby to z formuły (23). Jednak wniosku tego nie można uogólnić dla dowolnej wartości prądu I_{DS1} , chociaż w formule (23) nie ma zależności od I_{DS1} . Spójrzmy na rys. 8b, który przedstawia zależność W_{opt} od prądu I_{DS1} przy ustalonej wartości $C_d + C_1 = 3.5$ pF. Dla małych wartości prądu preferowane są tranzystory o znacznie mniejszej szerokości od przewidywanych formuły (23), co związane jest z pracą tranzystora M1 w obszarze słabej inwersji. W obszarze słabej inwersji transkonduktancja g_{m1} nie zależy od wymiarów tranzystora zgodnie z formułą (19), natomiast zmniejszanie szerokości tranzystora obniża wartość pojemności C_g , co korzystnie wpływa na minimalizację wartości wyrażenia (14) na ENC_{w1} .

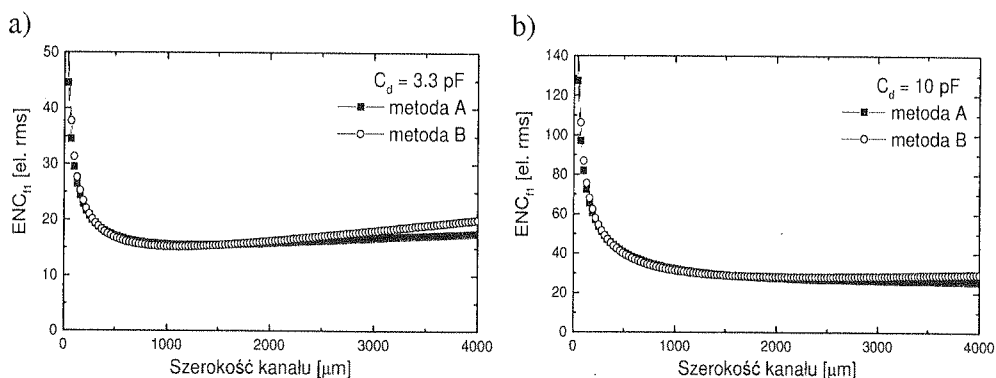
Przejdźmy do analizy szumów migotania. Do wyznaczenia gęstości widmowej tych szumów potrzebujemy wartości współczynnika K_f (patrz wzór (11) i (15)). Najlepszym rozwiązaniem byłby tu pomiar gęstości widmowych odpowiedniej klasy tranzystorów o wymiarach, które bierzemy pod uwagę [20,21]. W przypadku braku takich możliwości można postąpić w dwojaki sposób. Po pierwsze, przy pomocy HSPICE można wygenerować gęstości widmowe mocy szumów dla interesujących tranzystorów i stąd oszacować wartość współczynnika K_f potrzebnego we wzorze (15) — metoda A. Przykładowo, sporządzając takie charakterystyki gęstości widmowej mocy szumów w zakresie niskich częstotliwości dla tranzystorów z $L_1 = 1.5 \mu m$ przy prądzie $I_{DS1} = 500 \mu A$, otrzymano wartość $K_{fp} = 2.99 \times 10^{-28} C^2/m^2$ dla tranzystora o $W_1 = 200 \mu m$ oraz $K_{fp} = 2.78 \times 10^{-28} C^2/m^2$ dla tranzystora o $W_1 = 2000 \mu m$. Drugi sposób, zwany dalej metodą B, polega na odwołaniu się bezpośrednio do wzorów zastosowanych w programie HSPICE i odpowiedniej modyfikacji wyrażenia (15) na ENC_{f1} . Model szumów $1/f$ zaimplementowany w stosowanej technologii, zakłada gęstość widmową mocy szumów prądu drenu tranzystora MOS w postaci

$$\frac{\overline{di_f^2}}{df} = \frac{KF}{C_{ox}} \frac{I_{DS}^{AF}}{L^2} \frac{1}{f} \quad (28)$$

gdzie KF i AF są współczynnikami szumów migotania specyfikowanymi przez producenta. Dla znalezienia formuły na gęstość widmową szumów w bramce tranzystora należy powyższą formułę podzielić przez kwadrat transkonduktancji i wówczas poprzez analogię do równań (11) i (15) ENC_{f1} można zapisać w postaci

$$ENC_{f1}^2 = \frac{3.41}{q^2} (C_d + C_1 + C_g)^2 \times \frac{KF}{g_{m1}^2 C_{ox}} \frac{I_{DS1}^{AF}}{L_1^2} \quad (29)$$

Mozna zatem sporządzić zależność ekwiwalentnego ładunku szumów wnoszonego przez szumy migotania ENC_{fl} od szerokości tranzystora, używając albo metody A (wzór (15), przyjęto $K_{fp} = 2.9 \times 10^{-28} C^2/m^2$), albo metody B (wzór (29)). Sporządzone zależności dla $L_1 = 1.5 \mu m$, $I_{DS1} = 500 \mu A$, $C_1 = 200$ fF oraz $T_p = 500$ ns oraz dwóch różnych wartości C_d przedstawia rys. 9. Jak widać z rysunku różnice pomiędzy obu metodami są nieznaczne i pojawiają się jedynie w okolicach granicznych wartości szerokości kanału tranzystora uwzględnionych w symulacjach. Ponadto dla dużych szerokości kanału (powyżej $1000 \mu m$) wartość ENC_{fl} praktycznie nie ulega zmianie,



Rys. 9. Wartość ENC_{fl} w funkcji szerokości tranzystora M1 dla: a) $C_d = 3.3$ pF, b) $C_d = 10$ pF

Fig. 9. ENC_{fl} as a function of width of M1 transistor for: a) $C_d = 3.3$ pF, b) $C_d = 10$ pF

ponieważ spadek poziomu szumów migotania spowodowany wzrostem powierzchni tranzystora jest redukowany przez wzrost pojemności C_g pojawiającej się we wzorze (15) i (29).

Dla czasu kształtowania T_p w zakresie od 500 ns do $1 \mu s$ poziom ENC_{wl} jest znacząco większy od poziomu ENC_{fl} i dlatego szumy termiczne kanału tranzystora będą determinować wybór optymalnej szerokości tranzystora. Ostateczny optymalny wymiar tranzystora wyznaczono, biorąc pod uwagę pierwiastek z sumy kwadratów ENC_{wl} i ENC_{fl} . Dla układu RX64 wybrano tranzystor wejściowy o szerokości $L_1 = 1.5 \mu m$ i $W_1 = 400 \mu m$, pracujący przy prądzie $I_{DS1} = 500 \mu A$.

3.2. CAŁKOWITE SZUMY WZMACNIACZA

W zespole elementów wnoszących istotny wkład do szumów wzmacniacza należy uwzględnić jeszcze tranzystor M4 oraz pasożytniczą rezystancję R_{con} połączenia wzmacniacza z detektorem. O doborze wymiarów tranzystora M4 wynoszących $W_4/L_4 = 100 \mu m/10 \mu m$, pracującego w układzie zawiniętej kaskody jako źródło prądowe, zdecydował nie tylko jego wkład w szumy białe i migotania, ale również jego rezystancja wyjściowa i możliwości ustawienia właściwego punktu pracy. Rezystancja

R_{con} zdeterminowana jest maskami układu scalonego i samego detektora. Na jej wartość składają się: rezystancja bramki tranzystora wejściowego M1 wynosząca $14\ \Omega$, (obliczona na podstawie wzorów z pracy [23]), rezystancja połączenia bramki tranzystora M1 z wejściowym polem kontaktowym wynosząca średnio $10\ \Omega$, oraz rezystancja metalowego paska detektora krzemowego o wartości $15\ \Omega$ (wyznaczona na podstawie pomiarów).

Podsumowanie istotnych ustawień w układzie wzmacniacza ładunkowego determinujących całkowite szумы układu prezentuje tabela 1. Rysunek 10 pokazuje natomiast kontrybucję poszczególnych składowych do całkowitego ENC w zależności od ustawienia czasu kształtowania T_p , przy założeniu wartości pojemności $C_d + C_1 = 3.5\ \text{pF}$. Rozróżniono tam składowe pochodzące od szumów termicznych oraz szumów migotania, zgodnie z oznaczeniami z tabeli 1. Z zamieszczonego rysunku płyną następujące wnioski:

— minimum ENC występuje przy czasie kształtowania $T_{po} = 800 - 1000\ \text{ns}$ i wynosi 104 el. rms,

— przy krótszych czasach kształtowania, poniżej T_{po} , w układzie dominują szумы termiczne tranzystora M1,

— przy dłuższych czasach kształtowania, powyżej T_{po} , całkowita wartość ENC zdeterminowana jest przez efektywną wartość rezystancji w pętli sprzężenia zwrotnego wzmacniacza (składowa ENC_i),

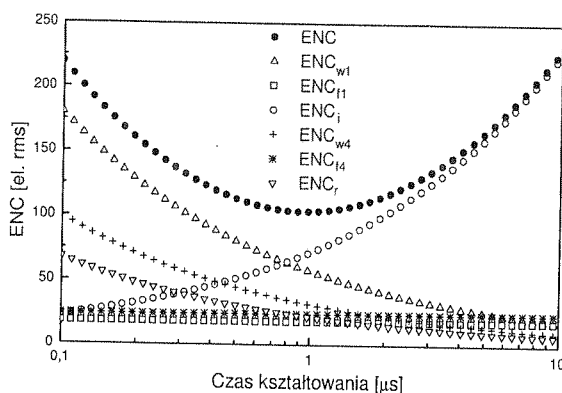
— w przypadku tej konkretnej konfiguracji układowej przy odpowiednim doborze wymiarów tranzystorów i parametrów kształtowania wkład szumów migotania jest pomijalnie mały.

Tabela 1

Istotne elementy składowe wzmacniacza ładunkowego wpływające na jego szумы

Important elements of charge amplifier which influence its noise performance

Element	Uwagi	Oznaczenie na rys. 10	
		Szумы techniczne	Szумы 1/f
Sprężenie zwrotne:			
– rezystencja R_{fed}	PMOS $W_5/L_5 = 2\mu\text{m}/384\mu\text{m}$	ENC_i	
Stopień wzmacniacza:			
– tranzystor M1	PMOS $W_1/L_1 = 400\mu\text{m}/1.5\mu\text{m}$	ENC_{w1}	ENC_{f1}
– tranzystor M4	NMOS $W_4/L_4 = 100\mu\text{m}/10\mu\text{m}$	ENC_{w4}	ENC_{f4}
Rezystancja połączenia wzmacniacza z detektorem	$R_{con} = 39\ \Omega$	ENC_r	



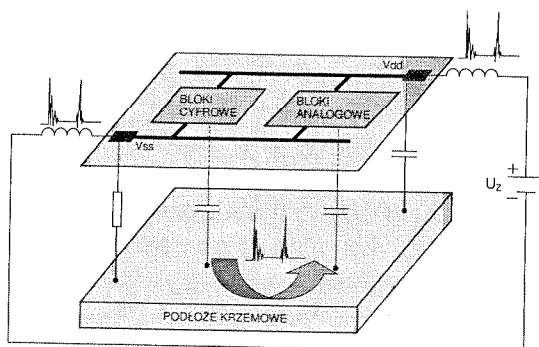
Rys. 10. Poszczególne składowe ENC w zależności od czasu kształtowania T_p

Fig. 10. Important components of ENC vs peaking time T_p

4. PRZESŁUCHY W MIESZANYCH UKŁADACH ANALOGOWO-CYFROWYCH

Pojawienie się na układzie scalonym obok niskoszumnych bloków analogowych szybkich układów cyfrowych wiąże się z koniecznością spojrzenia na projektowany układ scalony pod kątem efektów przesłuchu, bowiem pracujące bloki cyfrowe są źródłem sygnałów zakłócających. Jeżeli zakłócenia te pojawią się w torze użytecznego sygnału, a ich poziom przekroczy poziom szumów własnych bloków analogowych, to projekt układu kończy się fiaskiem. Zanim przejdziemy jednak do konkretnych rozwiązań minimalizujących szkodliwe efekty przesłuchu, przypomnijmy w skrócie genezę samego zjawiska.

Przełączanie inwerterów, bramek, przerzutników itp. zawsze wiąże się z ładowaniem i rozładowywaniem pewnych pojemności w krótkim okresie czasu, co powoduje generowanie w układzie krótkich impulsów prądowych. Im większa pojemność



Rys. 11. Transfer zakłóceń pomiędzy blokami cyfrowymi i analogowymi

Fig. 11. Transfer of disturbances between digital and analogue blocks

i krótszy czas jej przeładowania oraz większa różnica napięć, tym większa amplituda zakłócenia. Liczba przełączanych elementów przekłada się bezpośrednio na liczbę generowanych impulsów zakłócających. Zakłócenia pomiędzy blokami analogowymi i cyfrowymi mogą przenosić się na dwa sposoby (patrz rys. 11): poprzez wspólne linie zasilania albo poprzez podłoże krzemowe, na którym wykonane są układy (napięcie progowe MOS silnie zależy od U_{SB} , w modelu małosygnałowym efekt ten jest uwzględniony przez transkonduktancję podłoża g_{mb}).

W procesie projektowania mieszanego analogowo-cyfrowego układu scalonego można wyróżnić dwa podstawowe etapy, które zmierzają do minimalizacji efektów zakłóceń pochodzących od bloków cyfrowych. W pierwszym z tych etapów należy wziąć pod uwagę rozwiązania układowe pod kątem:

- podniesienia odporności bloków analogowych na sygnały zakłócające,
- redukcji ilości i poziomu zakłóceń generowanych w blokach cyfrowych,
- możliwości rozdzielania w czasie istotnych funkcji analogowych i cyfrowych,
- możliwości i sposobu filtracji sygnałów zakłócających.

Drugi etap koncentruje się na planie masek układu scalonego ze szczególnym zwróceniem uwagi na:

- odpowiednie rozmieszczenie poszczególnych bloków,
- rozprowadzenie linii zasilania i pól kontaktowych,
- rozmieszczenie i połączenie pierścieni ochronnych oraz kontaktów do podłoża,
- stosowanie warstw ekranujących.

Rozwiązania układowe powinny na etapie symulacji uwzględniać pasożytnicze rezystancje, pojemności i indukcyjności wynikające z planu masek układu scalonego, a w pewnych szczególnych przypadkach brać pod uwagę nawet rezystancyjny model podłoża.

4.1. MINIMALIZACJA EFEKTÓW PRZESŁUCHU W UKŁADZIE RX64

Rozważając układ RX64 pod kątem minimalizacji przesłuchów, ograniczymy się do przypomnienia najistotniejszych z tego punktu widzenia rozwiązań układowych, koncentrując się na planie masek układu scalonego.

Z punktu widzenia odporności bloków analogowych na zakłócenia szczególnie korzystnie wypadają struktury różnicowe, jednak stosowanie tego typu rozwiązań w pierwszych stopniach analogowego toru odczytowego jest wykluczone ze względu na optymalizację szumową. Starano się natomiast, tam gdzie to możliwe, tak dobrać architekturę pierwszych stopni, aby w torze sygnału stosować głównie tranzystory PMOS. I tak przykładowo we wzmacniaczu ładunkowym i układzie kształtującym główny tranzystor wzmacniający jest tranzystorem typu PMOS (patrz konfiguracja zawiniętej kaskody — rys. 5), natomiast studnię n (w której wykonany jest tranzystor PMOS) wykorzystano jako warstwę ekranującą od zakłóceń przenoszonych przez podłoże. Dla końcowego stopnia, jakim jest dyskryminator, zastosowano architekturę

różnicową. Dla tak zaprojektowanego toru elektroniki odczytu przeprowadzono serię symulacji z uwzględnieniem:

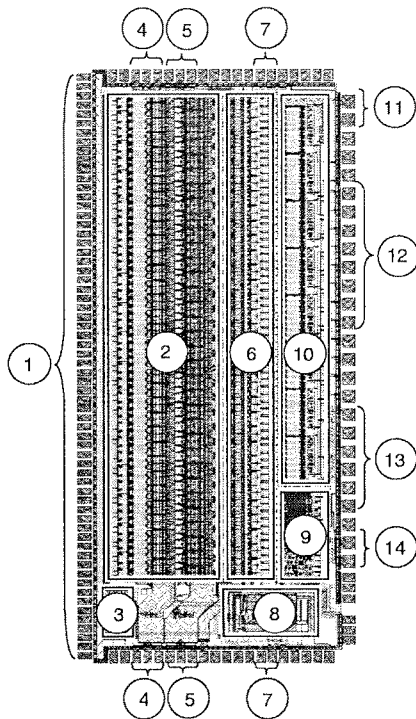
- pasożytniczej rezystancji linii zasilania,
- indukcyjności połączeń układu scalonego z płytą drukowaną,
- obliczonych na podstawie planu masek układu scalonego pojemności pasożytniczych do dodatniej linii zasilania i masy.

Przeprowadzone symulacje pokazały, że stopień wzmacniacza ładunkowego jest wrażliwy na zakłócenia pochodzące od dyskryminatora, które przenoszone są linią dodatniego zasilania. Dlatego zweryfikowano plan masek układu scalonego i zdecydowano się na poprowadzenie dwóch oddzielnych linii dodatniego zasilania: jednej dla wzmacniacza i układu kształtującego, a drugiej dla stopnia dyskryminatora. Dodatkowo w układzie dyskryminatora wprowadzono histerezę, która zapobiega kilkukrotnej odpowiedzi tego stopnia na obciążony szumami (lub zakłóceniami) stopień wejściowy.

W blokach cyfrowych ograniczono ilość przełączanych elementów poprzez oparcie struktury liczników na architekturze generatora pseudolosowego [24]. Do komunikacji układu scalonego ze światem zewnętrznym wykorzystano standard LVDS [25], gdzie sygnały cyfrowe o amplitudzie 400 mV są przesyłane różnicowo. Przebiegi czasowe odnoszące się do funkcjonowania poszczególnych bloków zostały tak zorganizowane, aby w trakcie obróbki sygnałów w kanałach analogowych jedynymi blokami cyfrowymi, w których następuje przełączanie, były układy liczników oraz dekodery komend zamykający i otwierający bramki transmisyjne dla zliczania impulsów.

4.1.1. Plan masek układu scalonego

Rysunek 12 przedstawia plan masek układu scalonego RX64, którego całkowita powierzchnia wynosi $2800 \times 6500 \mu\text{m}^2$. Ze względu na geometrię sensora 64 wejściowe pola kontaktowe układu scalonego znajdują się po jednej stronie. Zdeterminowało to rozmieszczenie poszczególnych bloków funkcjonalnych wewnątrz samego układu scalonego, jednak w miarę możliwości starano się przestrzegać zasady, aby wrażliwe na zakłócenia bloki analogowe (wzmacniacze ładunkowe i układy kształtujące) znajdowały się jak najdalej od bloków cyfrowych (liczniki i dekodery komend). Zaznaczone na rys. 12 bloki wzmacniaczy ładunkowych i układów kształtujących oraz dyskryminatorów obejmują w sumie 64 kanały użytkowe oraz 2 kanały testowe (porównaj rys. 3a). Każdy z kanałów zajmuje na układzie scalonym obszar w kształcie paska o szerokości $80 \mu\text{m}$ i długości $1500 \mu\text{m}$. Linie zasilania w pojedynczym kanale są rozprowadzone wzdłuż kanału warstwą metalu 1 o szerokości $10 \mu\text{m}$, co daje ich rezystancje pasożytnicze na poziomie 1Ω . Wewnątrz kanału rozdzielono zasilanie sekcji wzmacniacza ładunkowego i układu kształtującego od sekcji dyskryminatora. Prostopadle do kanałów poprowadzono warstwą metalu 2 zbiorcze linie zasilania. I tak w sekcji wzmacniaczy i układów kształtujących zbiorcze linie do dodatniego zasilania i masy są długości $6500 \mu\text{m}$ i szerokości $250 \mu\text{m}$ każda. Daje to rezystancję pasożytniczą każdej takiej linii około 1Ω . Napięcia zasilania z płytki drukowanej doprowadzane są

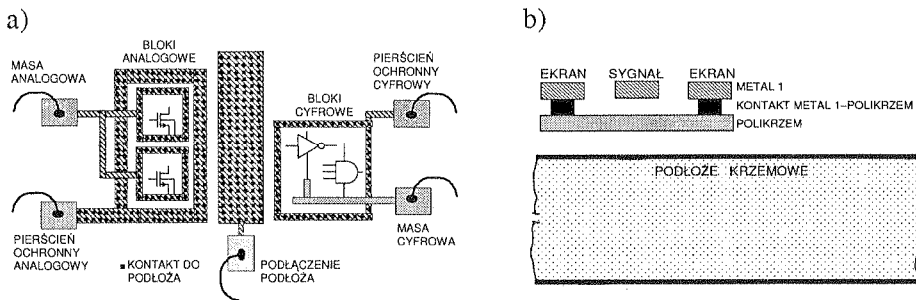


Legenda:

1. Wejściowe pola kontaktowe
2. Wzmacniacze ładunkowe i układy kształtujące
3. Układ kalibracyjny
4. Pola kontaktowe do zasilania wzmacniaczy ładunkowych i układów kształtujących
5. Pola kontaktowe do masy części analogowej i podłoża
6. Dyskryminatory
7. Pola kontaktowe zasilania dyskryminatorów
8. Przetworniki C/A
9. Dekoder komend
10. Liczniki
11. Pola kontaktowe do zasilania części cyfrowej
12. Pola kontaktowe szyny danych
13. Pola kontaktowe sygnałów sterujących
14. Pola kontaktowe do masy części cyfrowej

Rys. 12. Plan masek układu RX64

Fig. 12. Layout of RX64 chip



Rys. 13. Techniki izolacji i ekranowania: a) pierścienie ochronne (widok z góry), b) warstwy ekranujące dla linii sygnałowych (przekrój)

Fig. 13. Techniques of isolation and shielding: a) guard rings, b) shielding layers for signal lines

do każdej takiej linii przez 6 pól kontaktowych, rozmieszczonych symetrycznie po obu stronach układu scalonego. Komparatory z histerezą, stanowiące ostatni człon kanału analogowego, posiadają swoje odrębne zbiorcze linie zasilania prowadzone podobnie jak opisano wyżej, lecz o szerokości po 100 μm każda, podłączone poprzez dwa pola kontaktowe z każdej strony. Linie do zasilania liczników i pozostałych bloków cyfrowych są szerokości 80 μm i są złożeniem prowadzonych jedna nad drugą warstw metalu 1 i metalu 2 i również posiadają po 2 pola kontaktowe z każdej strony.

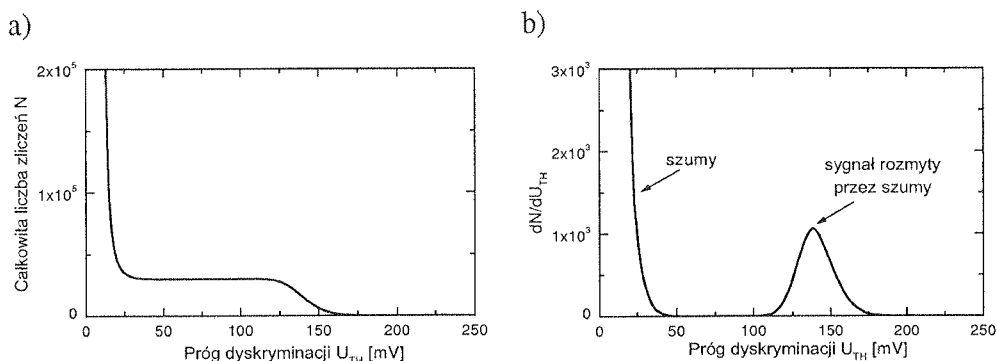
Układ RX64 wykonany jest na niskorezystywnym podłożu z cienką warstwą epitaksjalną grubości kilku μm . Przy tego typu podłożu dla dużych odległości (większych od czterokrotnej grubości warstwy epitaksjalnej) sygnały zakłócające pochodzące od bloków cyfrowych rozchodzą się w niskorezystancyjnej warstwie dolnej podłoża. Dlatego zwrócono szczególną uwagę na rozmieszczenie i połączenie kontaktów do podłoża, które z jednej strony dzięki ich dużej ilości miały zapewnić niską (na poziomie dziesiątek $\text{m}\Omega$) rezystancję połączenia podłoża z masą, z drugiej strony odpowiednio uformowane tworzyły pierścienie ochronne czy to bloków analogowych, czy cyfrowych, tak jak ilustruje to rys. 13a. Układ scalony montowany jest na płycie drukowanej z użyciem kleju przewodzącego, natomiast linie masy i pierścienie ochronne uformowane z kontaktów do podłoża są łączone poprzez pola kontaktowe układu scalonego z masą płytki drukowanej. W blokach cyfrowych dla linii zegara zastosowano ekranowanie, tak jak przedstawia to schematycznie rys. 13b, natomiast wszelkie wolne przestrzenie w układzie cyfrowym wypełniano kondensatorami zapiętymi pomiędzy liniami zasilania.

5. POMIARY SZUMÓW W UKŁADZIE RX64

Pomiar szumów w układzie z architekturą binarną, w której analogowy tor odczytu zakończony jest dyskryminatorem, można przeprowadzić albo poprzez pomiar tzw. rozmycia szumowego sygnału, albo poprzez pomiar ilości zliczeń na wyjściu dyskryminatora generowanych wyłącznie przez szumy. Każda z tych metod wiąże się z pewnymi trudnościami pomiarowymi oraz posiada swoje ograniczenia, które teraz zasygnalizujemy.

Gdy dokonujemy pomiaru rozmycia szumowego zazwyczaj na wejściu układu podajemy ciąg monoamplitudowych impulsów, natomiast na wyjściu dyskryminatora zliczamy impulsy w określonym przedziale czasu i przy pewnym progu dyskryminacji. Jeżeli pomiar powtórzymy dla różnych progów dyskryminacji, to w efekcie otrzymamy ilość zliczeń w funkcji progów dyskryminacji tak to jak przedstawia rys. 14a. Dla niskich progów dyskryminacji obserwujemy gwałtowny wzrost ilości zliczeń. Za ten wzrost ilości zliczeń odpowiedzialne są szumy na wejściu dyskryminatora. Różniczkując całkowitą ilość zliczeń w funkcji progów dyskryminacji, otrzymujemy rozkład amplitudowy impulsów wejściowych (patrz rys. 14b), w którym zarówno same szumy, jak i sygnał rozmyty przez szumy, są dobrze widoczne. Jeżeli rozkład wartości chwilowych szumów ma charakter gaussowski, to obserwowane na rozkładzie amplitudowym rozmycie

szumowe też ma charakter gaussowski. Dokonując pomiaru rozmycia szumowego i znając wzmacnienie, można wyznaczyć interesujący nas w tym przypadku ekwiwalentny ładunek szumów ENC.



Rys. 14. Zliczenia na wyjściu dyskryminatora w zależności od progu dyskryminacji: a) całkowita ilość zliczeń, b) rozkład amplitudowy

Fig. 14. Counts at the discriminator output vs threshold level: a) total number of counts, b) differential pulse high spectrum

Pomiarów szumów na podstawie rozmycia szumowego sygnału dokonano zarówno z użyciem impulsów generowanych przez wewnętrzny układ kalibracyjny, jak i z użyciem źródeł promieniotwórczych. Pomiar z użyciem wewnętrznego źródła kalibracyjnego wydaje się być pomiarem stosunkowo łatwym do przeprowadzenia. Podczas pomiaru na wejście układu podawany jest skok napięcia o amplitudzie V_i (na poziomie pojedynczych mV) poprzez niewielkie kondensatory testowe o wartości około $C_i = 70$ fF rozmieszczone na układzie scalonym na wejściu każdego kanału. Odpowiada to wstrzykiwaniu do wejścia wzmacniacza impulsów prądowych niosących ładunek $Q_{in} = C_i \times V_i$. Należy tu jednak zwrócić uwagę na dwa efekty. Sam układ kalibracyjny (czy też zewnętrzny generator) może być źródłem szumów, tym bardziej że chodzi o wstrzykiwanie impulsów o bardzo małych amplitudach. Po drugie, aby szumy przeliczyć na wejście układu potrzebujemy dokładnie znać wzmacnienie w każdym z 64 kanałów układu scalonego. Ze względu na możliwe efekty niedopasowania dla niewielkich kondensatorów testowych znacznie lepiej jest odwołać się do pomiaru wzmacnienia z wykorzystaniem fizycznych sygnałów generowanych w sensorze przez promieniowanie X.

W testach z wykorzystaniem fizycznych sygnałów generowanych w sensorze używano źródeł promieniowania X o dobrze określonych energiach począwszy od 6.4 keV do 20.1 keV. Tego typu źródła promieniotwórcze w detektorze krzemowym generują impulsy prądowe niosące dobrze określony ładunek, wstrzykiwany do wejść układu scalonego. Dla współczynnika konwersji w krzemie wynoszącego 3.67 eV na generację pary elektron-dziura stosowane źródła promieniotwórcze generowały ładunek z zakre-

su od 1744 do 5477 elektronów. Szczegółowy opis prowadzenia tego typu pomiarów można znaleźć w pracy [26]. O ile pomiar wzmocnienia jest rzeczywiście bardzo dokładny, o tyle przy pomiarze szumów trzeba rozważyć dodatkowe efekty związane z mechanizmami, które mają źródło w geometrii samego detektora. Z punktu widzenia pomiaru szumów najistotniejszy jest efekt podziału ładunku [27]. Efekt ten ma miejsce w przypadku, gdy wpadający kwant promieniowania X generuje ładunek w okolicach środka pomiędzy dwoma paskami detektora. Wówczas, w zależności od rozkładu natężenia pola elektrycznego, część ładunku może zostać zebrana przez jeden pasek detektora, a część przez pasek sąsiedni. W efekcie mierzone rozmycie sygnałów może być nieco większe.

Drugim podejściem do pomiarów szumów jest pomiar ilości zliczeń na wyjściu dyskryminatora generowanych wyłącznie przez same szumy. Zgodnie z formułą Rice'a [28], przy gaussowskim rozkładzie wartości chwilowych szumów, częstość impulsów na wyjściu dyskryminatora pochodzących od szumów można wyrazić następująco

$$f_n = f_0 \exp\left(-\frac{U_{TH0}^2}{2U_N^2}\right) \quad (30)$$

gdzie f_0 jest częstością sygnałów szumowych dla zerowego progu dyskryminacji (zależną od sposobu filtracji), U_{TH0} to wartość progu dyskryminacji (w naszym przypadku skorygowana o zmierzone wcześniej napięcie niezrównoważenia), a U_N jest wartością średniokwadratową szumów na wejściu dyskryminatora.

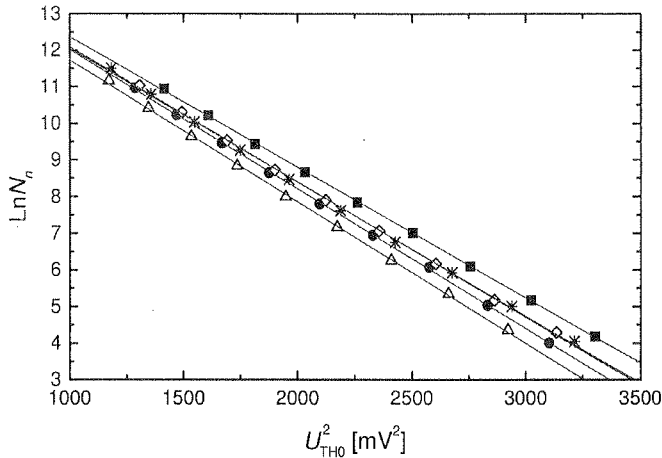
W przypadku układu RX64, bez wstrzykiwania jakiegokolwiek ładunku na wejście, dokonano pomiaru ilości zliczeń N_n przy niskich progach dyskryminacji i przy długich czasach akwizycji (po 20 minut na każdą ustawioną wartość progu dyskryminacji). Wyniki uzyskanych pomiarów przedstawia rys. 15, gdzie zamieszczono wykres logarytmu naturalnego z liczby zliczeń N_n w funkcji kwadratu progu dyskryminacji $(U_{TH0})^2$ dla pięciu reprezentatywnych kanałów układu RX64 (aby uniknąć zaciemniania wykresu). Otrzymane dane leżą na prostych, co oznacza, że rozkłady wartości chwilowych impulsów na wejściach dyskryminatorów mają charakter gaussowski. Tego typu rozkładem charakteryzują się impulsy pochodzące od szumów fizycznych układu, a nie impulsy pochodzące od sygnałów zakłócających mających swe źródło w przełączaniu bloków cyfrowych [29,30].

Ponadto, aby upewnić się czy układ jest wolny od efektów przesłuchu, dokonano pomiaru szumów w dwóch przypadkach, a mianowicie:

— w przypadku gdy w układzie pracują jednocześnie wszystkie 64 kanały — na ich wejścia podawane są sygnały,

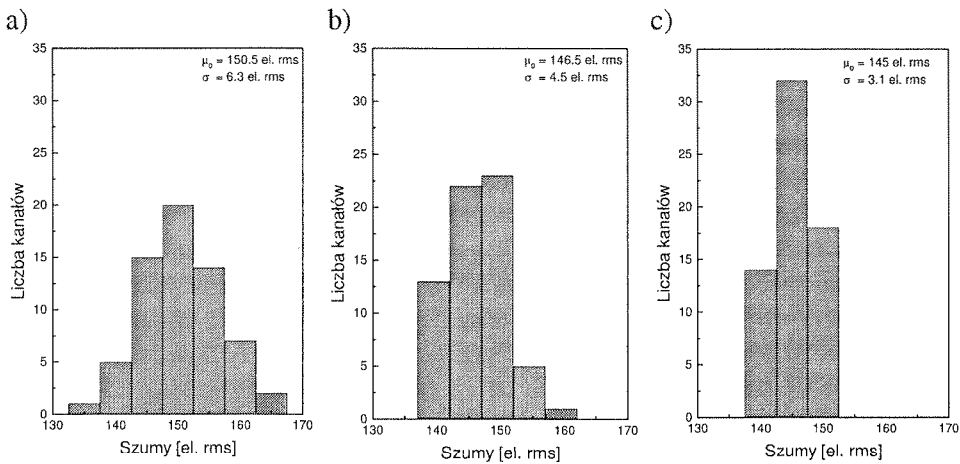
— w przypadku gdy sygnały wejściowe podawane są tylko do 3 kanałów, natomiast do wejść pozostałych kanałów nie podaje się impulsów.

Ponieważ ilość zakłóceń generowanych przez układy cyfrowe jest proporcjonalna do ilości aktywnych przełączających układów cyfrowych, a w układzie RX64 nie zaobserwowano różnicy w mierzonym poziomie szumów w obu powyższych przypadkach stąd wniosek, że efekty przesłuchów udało się sprowadzić do poziomu zaniedbywalnego.



Rys. 15. Logarytm naturalny z liczby zliczeń w funkcji kwadratu napięcia progowego dla pięciu reprezentatywnych kanałów układu RX64

Fig. 15. Natural logarithm of number of counts as a function of square of threshold voltage for five representative channels of RX64



Rys. 16. Histogramy szumów zmierzone w 64 kanałach układu RX64: a) przy użyciu wewnętrznego układu kalibracyjnego, b) przy użyciu źródła promieniotwórczego Pu-238, c) korzystając ze wzorów Rice'a

Fig. 16. Distributions of noise in 64 channels of RX64 chip measured by: a) using internal calibration generator, b) using radioactive source Pu-238, c) using the Rice formula

Podsumowanie wyników pomiarowych dotyczących szumów przeprowadzonych dla czasu kształtowania $T_p = 800$ ns zestawiono na histogramach na rys. 16, przedstawiających wyniki ze wszystkich 64 kanałów dla pomiarów z użyciem wewnętrznego

układu kalibracyjnego (rys. 16a), źródła promieniotwórczego (rys. 16b) oraz z bezpośredniego zliczania impulsów szumowych (rys. 16c). Wszystkie otrzymane wartości są do siebie zbliżone, przy czym rozrzut zmierzonych wartości szumów jest najmniejszy dla histogramów na rys. 16b i 16c, co preferuje te metody pomiarowe.

Zmierzone wartości szumów są wyższe w stosunku do przewidywań teoretycznych przedstawionych w rozdziale 3. Powodem tej różnicy nie są efekty przesłuchu pochodzące od układów cyfrowych, jako że zmierzony szum ma charakter gaussowski w szerokim zakresie amplitud (por. rys. 15). Prawdopodobnie obserwowana różnica ma swoje źródło w efektach gorących nośników w tranzystorze wejściowym wzmacniacza ładunkowego, które pomimo zwiększenia długości kanału tranzystora ponad dwukrotnie w stosunku do dopuszczalnej wartości minimalnej są nadal obecne.

W praktyce aby dać przekonującą odpowiedź, należałoby wcześniej przed rozpoczęciem projektu wykonać serię pomiarów szumów na strukturach testowych tranzystorów dla stosowanej technologii. Tego typu struktury testowe i pomiary wykonywane są zawsze, kiedy problem szumów jest krytyczny, a w perspektywie przewiduje się produkcję specyfikowanych układów scalonych na większą skalę.

6. PODSUMOWANIE

Przedstawione podejście do optymalizacji szumów stopnia wejściowego pokazuje w jaki sposób, pomimo niedoskonałości modeli szumowych, można zidentyfikować główne źródła szumów własnych w układzie elektronicznym i jak dobrać wymiary i polaryzacje poszczególnych elementów, aby przy narzuconych ograniczeniach uzyskać minimalną wartość szumów. Przedstawiona metoda obejmuje zakres pracy tranzystorów w silnej, umiarkowanej i słabej inwersji, co wydaje się być użyteczne, zwłaszcza w przypadku nowych technologii submikronowych. Interesujące są rezultaty dotyczące doboru szerokości tranzystora wejściowego w zależności od pojemności sensora oraz od prądu w tranzystorze wejściowym wzmacniacza ładunkowego. Ponadto wskazano na inne elementy w układzie, których kontrybucja do szumów wzmacniacza ładunkowych jest zazwyczaj pomijana. Określono również wkład poszczególnych składowych w całkowity ekwiwalentny ładunek szumów w zależności od doboru czasu kształtowania.

W związku z mieszanym analogowo-cyfrowym charakterem układu scalonego omówiono zastosowane metody minimalizacji efektów przesłuchu. Wskazano na istotne aspekty związane z wyborem architektury układu i rysowaniem planu masek.

Dokonano dokładnych pomiarów szumów dla wszystkich kanałów w układzie RX64 trzema metodami, wskazując na niedoskonałości metod związanych z pomiarem rozmycia szumowego impulsów. Przeprowadzone wyniki pomiarów potwierdzają, że efekty przesłuchów w układzie mieszanym udało się sprowadzić do poziomu zanieczyszczenia.

7. Bibliografia

1. P. M. Solomon (ed): *Scaling CMOS to the limit*. IBM Journal of Research and Development, vol. 46, no. 2/3, 2002.
2. B. Hilt, P. Fessler, G. Prevot: *New quantum detection system for very low dose X-ray radiology*. Nuclear Instruments and Methods, vol. A442, 2000, pp. 38–44.
3. E. Beuville, et al.: *An application specific integrated circuit and data acquisition system for digital X-ray imaging*. Nuclear Instruments and Methods, vol. A406, 1998, pp. 337–342.
4. P. Gryboś: *Low Noise Multichannel Integrated Circuits in CMOS Technology for Physics and Biology Applications*. Monografie 117, AGH Uczelniane Wydawnictwa Naukowo-Dydaktyczne, Kraków 2002.
5. W. Dąbrowski, P. Gryboś, K. Świentek, P. Wiącek: *Integrated readout of silicon strip detectors for position sensitive measurement of X-rays*. Nuclear Instruments and Methods, vol. 512, 2003, pp. 213–219.
6. G. Baldazzi, et al.: *X-ray imaging with a silicon microstrip detector coupled to the RX64 ASIC*. Nuclear Instruments and Methods, vol. A509, 2003, pp. 315–320.
7. P. Ratio Mendes, et al.: *Silicon strip detectors for two-dimensional soft X-ray imaging at normal incidence*. Nuclear Instruments and Methods, vol. A509, 2003, pp. 333–339.
8. M. Feuerstack-Raible: *Overview of microstrip readout chips*. Nuclear Instruments and Methods, vol. A447, 2000, pp. 35–43.
9. G. Hall: *LHC front-end electronics*. Nuclear Instruments and Methods, vol. A453, 2000, pp. 353–364.
10. P. Gryboś, W. Dąbrowski, P. Hottowy: *NEURO64 — design for a good matching performance in multichannel ASIC*. Proceedings of the 16th European Conference on Circuit Theory and Design ECCTD'2003 Kraków, Poland, September 2003, vol. 2, pp. 1–4.
11. P. Gryboś, W. Dąbrowski, T. Fiutowski, K. Słowikowski: *Design for low noise and good matching in multichannel amplifier for recording neuronal signals in modern neuroscience experiments*. Proceedings of 9th International Conference on Mixed Design of Integrated Circuits and Systems – MIXDES'2002 Wrocław, Poland, June 2002, pp. 255–260.
12. K. Korbel: *Układy elektroniki front-end*. AGH Uczelniane Wydawnictwa Naukowo-Dydaktyczne, Kraków 2000.
13. E. Gatti, P. F. Manfredi: *Processing the signal from solid-state detectors in elementary particle physics*. La Revista del Nuovo Cimento, vol. 9, no. 1, 1986.
14. W. Sansen, Z. Y. Chang: *Limits of low noise performance of detector readout front ends in CMOS technology*. IEEE Transactions on Circuits and Systems, vol. 37, no. 11, 1990, pp. 1375–1382.
15. Y. Tsividis: *Operation and modeling of the MOS Transistor*. New York, WCB McGraw-Hill, 1999.
16. E. Nygard, et al.: *CMOS low noise amplifier for microstrip readout. Design and results*. Nuclear Instruments and Methods, vol. A301, 1991, pp. 506–516.
17. D. Foty: *Modernizing the framework for deep submicron design – for CMOS and beyond*. Proceedings of 10th International Conference on Mixed Design of Integrated Circuits and Systems MIXDES'2003 Łódź, Poland, June 2003, pp. 19–24.
18. C. Enz, F. Krummenacher, E. A. Vittoz: *An analytical MOS transistor model valid in all regions of operation and dedicated to low-voltage and low-current application*. Analog Integrated Circuits and Signal Processing, vol. 8, 1995, pp. 83–114.
19. BSIM3v3 manual. Department of Electrical Engineering and Computer Sciences, University of California Berkeley, CA 94720, 1996.
20. M. Raymond, et al.: *Radiation hard electronics for LHC*. Nuclear Instruments and Methods, vol. A360, 1995, pp. 162–165.

21. M. Manghisoni, et al.: *Submicron CMOS technologies for low-noise analog front-end circuits*. IEEE Transactions on Nuclear Science, vol. 49, no. 4, 2002, pp. 1783–1790.
22. J.S. Lovett, M. Welten, A. Mathewson, B. Mason: *Optimizing MOS transistor mismatch*. IEEE Journal of Solid-State Circuits, vol. SC-33, no. 1, 1998, pp. 47–150.
23. V. Re, et al.: *Experimental study and modeling of the white noise sources in submicron P-and N-MOSFETs*. IEEE Transactions on Nuclear Science, vol. 48, no. 4, 2001, pp. 1577–1586.
24. P. Gryboś, W. Dąbrowski, P. Hottowy, R. Szczygieł, K. Świentek, P. Wiącek: *Multichannel mixed-mode IC for digital readout of silicon strip detectors*. Microelectronics Reliability, vol. 42, 2002, pp. 427–436.
25. IEEE Std 1596.3-1996, IEEE Standard for Low-Voltage Differential signals (LVDS) for Scalable Coherent Interface (SCI).
26. P. Gryboś, W. Dąbrowski: *Development of a fully integrated readout system for high count rate position-sensitive measurements of X-rays using silicon strip detectors*. IEEE Transactions on Nuclear Science, vol. 48, no. 3, 2001, pp. 466–472.
27. W. Dąbrowski, P. Gryboś: *Position sensitive semiconductor strip detectors.*, [in:] K. Tsuji, J. Injuk, R. Van Grieken (eds), *X-ray Spectrometry: Recent Technological Advances*. John Wiley & Sons, Ltd, 2004, pp. 247–275, to be printed.
28. S. O. Rice: *Mathematical analysis of random noise*. Bell System Tech. J., 24, 1945.
29. M. Heijningen, et al.: *Analysis and experimental verification of digital substrate noise generation for epi-type substrates*. IEEE Journal of Solid-State Circuits, vol. SC-35, no. 7, 2000, pp. 1002–1008.
30. T. Schmerbeck: *Noise coupling in mixed-mode ICs: Mechanism/Simulation/Measurement, Design Strategy/H W Example*. Lausanne, Switzerland, EPFL Electronics Laboratories Advanced Engineering Course CMOS & BiCMOS IC Design '99: Practical Aspects in Analog and Mix-Mode ICs 1999.

P. GRYBOŚ

DESIGN FOR A LOW NOISE OF MULTICHANNEL INTEGRATED CIRCUITS ON EXAMPLE OF RX64 CHIP

Summary

This paper describes designing of multichannel mixed-mode Application Specific Integrated Circuits in CMOS technology. The author discusses the problems of noise optimisation and crosstalk in mixed-mode integrated circuit on example of 64-channel chip RX64. This chip has been designed for the digital readout of silicon strip detectors used for position-sensitive X-ray imaging. Various requirements and constraints implied by this particular application have been taken into account in the design stage.

The RX64 chip consists of low noise analogue front-end electronic and digital blocks for data storage, bias control and communication via serial link. An architecture of integrated circuit has been described with a special attention paid to noise optimisation. A method of noise minimisation of front-end electronic has been developed taking into account capacitance of X-ray sensor and different value of bias current in the input transistor of charge preamplifier. The second order effects in noise calculation have also been investigated.

The minimisation of crosstalk in mixed-mode RX64 chip has been considered. As the chip contains analogue and digital blocks placed on common epi-type substrate, particular attention has been paid to the layout. The floorplan, power distribution and guardring placement are described.

The results of the design are verified by noise measurements using internal calibration generator, radioactive source with X-ray sensor and considering the noise counts according to the Rice formula. The equivalent input noise measured at room temperature for a sensor capacitance of 2.5 pF and peaking time of 0.8 μ s is only 145 el. rms.

Keywords: multichannel integrated circuits, ASIC, noise, mixed-mode circuits

Inexpensive Characteristics Scanner of Photovoltaic Modules

DARIUSZ MAKOWSKI, WITOLD MARAÑDA, RAFAŁ PIJARSKI

*Technical University of Łódź
Department of Microelectronics and Computer Science
Al. Politechniki 11, 93-590 Łódź, Poland
Email: dmakow@dmcs.p.lodz.pl, maranda@dmcs.p.lodz.pl*

*Otrzymano 2003.07.09
Autoryzowano 2003.10.03*

Detailed performance evaluation of any photovoltaic module requires the knowledge of its I-V characteristics [5]. While photovoltaic elements are constantly getting cheaper and more commonly used, specialized measurement equipment is generally out of reach for most maintainers of small and simple PV-systems. This paper presents the construction of the stand-alone characteristics measurement system. The adopted approach, based on a typical microcontroller with the reduction of other costly elements, makes the device fast and inexpensive, yet fully functional, thus targeted towards the use in small PV systems.

Keywords: PV measurements, PV performance evaluation, I-V curve

1. INTRODUCTION

The energy of solar radiation that reaches the Earth amounts to 10^{18} kWh every year. Photovoltaic cells are able to convert this energy directly into electricity. The graph presented on fig. 1 represents the behaviour of a typical solar cell at particular intensity of solar radiation.

The I-V characteristics of the photovoltaic (PV) module provides us with thorough information about the element performance. The knowledge of such characteristics is necessary for detailed performance evaluation of any PV-system. Performance evaluation of older systems, where ageing processes and possible failures in its internal structure are inevitable and influence the characteristics heavily, is of fundamental importance [3]. The measurement of the I-V curve of PV-module requires a specialised meter, which can drive the working point along the whole curve, taking measurements of current and voltage. The accuracy depends on the concentration of measurement

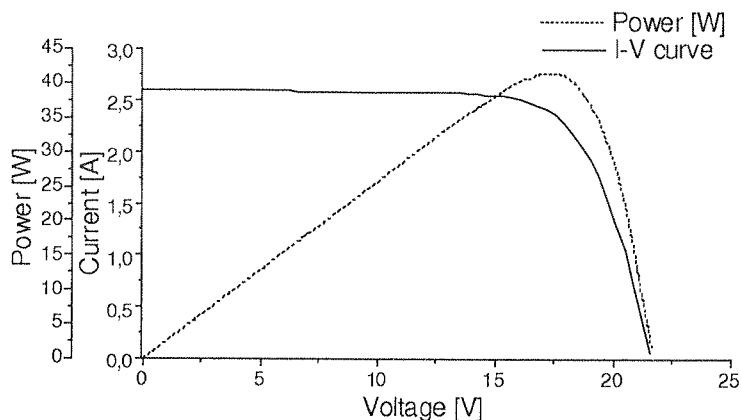


Fig. 1. Main characteristics of a typical 50Wp photovoltaic module

Rys. 1. Charakterystyka typowego modułu fotowoltaicznego o mocy 50Wp

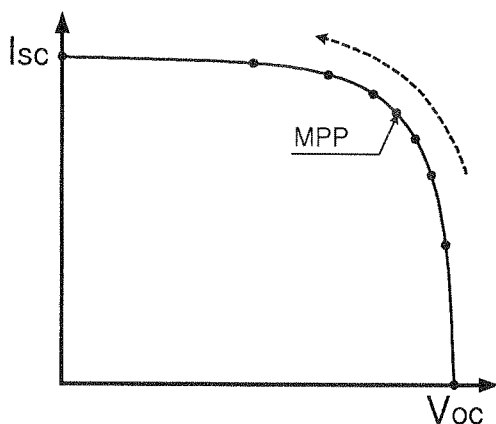


Fig. 2. Measurement point path during the measurement process

Rys. 2. Tor ruchu punktu pomiarowego podczas pomiaru

points, especially in the region of Maximum Power Point (MPP), whereas near V_{OC} and I_{SC} fewer number of measurements is needed and the characteristics can be easily approximated (see fig. 2). Therefore the measurement algorithm can have a significant influence on the speed and accuracy of the whole process [1].

Moreover, the measurement of modules in existing PV-system sets the limit on the total measurement time due to varying insolation conditions. We can make an assumption that it should be less than a second to measure all points on the I-V curve, which may require fast circuitry, contradicting the meter simplicity and low costs [1].

The meter described in this paper is dedicated to work with PV-modules up to 240 Wp.

2. TYPICAL MEASUREMENT SYSTEM

The construction of a typical measurement system is shown in fig. 3 [2]. The major points of such a system are: variable load, D/A and A/D converters for controlling the load and taking the measurements of voltages and currents. The current measurement is realised with the use of a shunt resistance, whereas the voltage can be measured directly. Both signals with the proper range adjustment are passed through the multiplexer to the A/D converter.

The measurement of the I-V curve of a PV-module consists in driving the operating point along the whole curve with the use of a variable load resistance. Such a load is usually a high speed MOS transistor with low channel resistance, in order to minimise the power dissipation and increase the accuracy near I_{SC} point (see fig. 2).

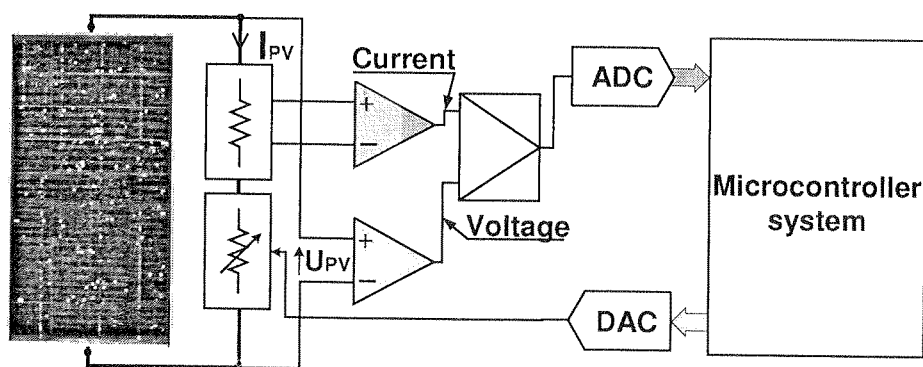


Fig. 3. Structure of a typical I-V characteristics meter

Rys. 3. Schemat blokowy typowego miernika charakterystyki I-V

In order to obtain fine control of the load, the gate voltage must be precisely adjusted with accuracy well below 1mV. This requires the use of 14-bit or at least 12-bit D/A converters, which unfortunately implies the relatively long conversion time [2]. Also, the cost of this element may be significant for the simple system.

3. SYSTEM WITH CHARGE INTEGRATION LOAD CONTROL

The proposed solution gets rid of the expenses caused of A/C and D/A converters usage (see fig.4). In order to minimise the cost, microcontroller equipped with A/D converter and the method of charge integration were used. This solution assures fine control of the voltage at the MOS gate. The voltage on the capacitor C can be risen in small steps by short impulses from the switch S_1 , controlled by the microcontroller (fig. 4) [1].

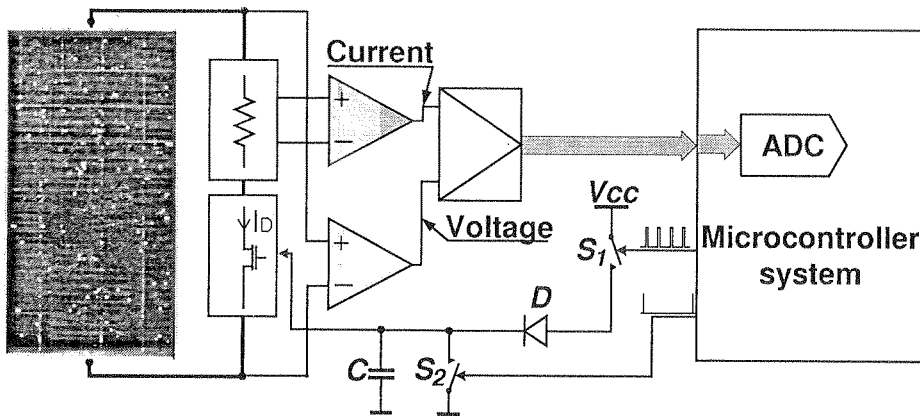


Fig. 4. System with charge integration load control

Rys. 4. Schemat blokowy systemu pomiarowego wykorzystującego metodę integracji ładunku

With this approach it is not possible to set the MOS gate voltage to a known value, but the fine voltage control can be obtained by changing the impulse width. After the full measurement cycle, the capacitor is discharged by the switch S_2 . Full one-cycle measurement process is shown in fig. 5.

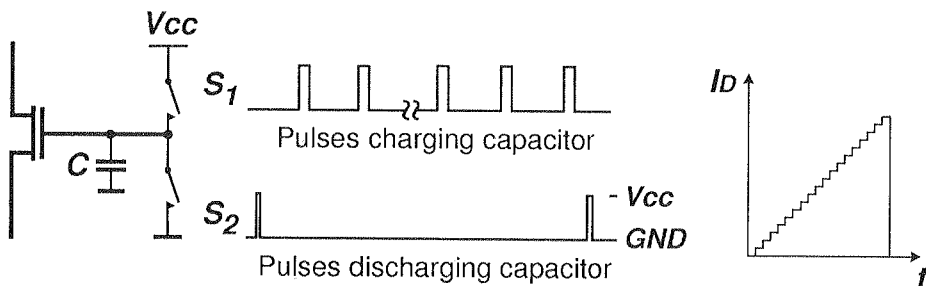


Fig. 5. Measurement cycle

Rys. 5. Cykl pomiarowy charakterystyki I-V

The measurement method actually does not require the knowledge of the voltage on the MOS-gate, but the rising PV-current (or falling PV-voltage) is checked to find out when the measurement is to be stored in memory [1].

However, the charge integration method requires more sophisticated microcontroller software to compensate the non-linear voltage growth on capacitor C . It is due to the parallel capacitance of the MOS gate and leakage currents, since the C capacitance must be small enough (in range of nF) to allow fast operation.

4. REALIZATION OF THE METER

The presented device was designed and built as a stand-alone characteristic meter. The built hardware is shown in fig. 6.

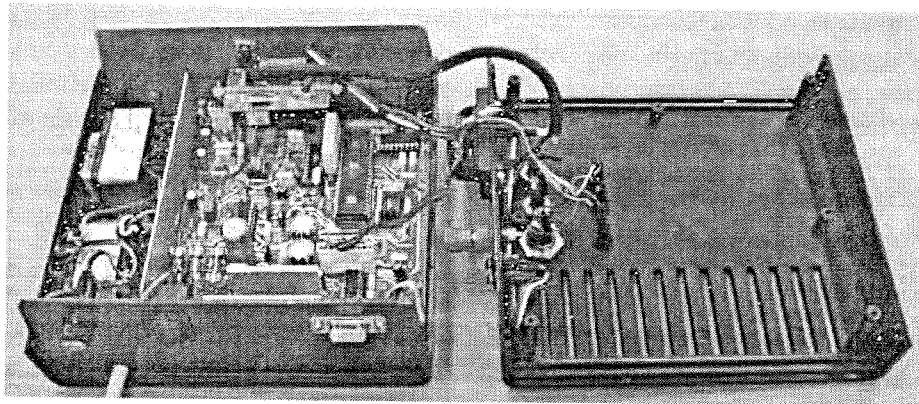


Fig. 6. I-V characteristics scanner

Rys. 6. Miernik charakterystyk I-V

The construction of the measurement system is based on the microcontroller AT-MEL ATmega 163 with built in A/D converter. As a variable load, IRF1010E NMOS transistor was used. This transistor has channel resistance small enough for this purpose. In order to use the characteristics scanner with typical photovoltaic modules, maximum measured current was limited to 8A and the voltage delivered from PV panel must be lower than 30V (see table 1). The microcontroller is responsible for generating control signals and collecting current and voltage samples. Measured data are sent to the main computer via RS232 interface. The device can work with two voltage and current ranges, which are automatically selected by checking the VOC and ISC before the measurement process (see fig. 2). Key parameters of the presented scanner are shown in table 1.

Table 1

Key parameters of the presented scanner

Parametry miernika

Current range	4A	8A
Voltage range	15V	30V
Current and Voltage resolution (points per range)	256	
Full characteristic measurement time	< 1s	

5. MEASUREMENT RESULTS

The meter was tested with the 1kWp PV-system of the Department of Microelectronics and Computer Science, at the Technical University of Łódź. The installation comprises of 20 modules SF52A from Solar Fabrik of 50Wp each (see fig. 7). In order to collect all the panels characteristics, it is necessary to use very fast meter and scan the panels one by one. Another alternative is to integrate each panel with separate scanner. The measurements should be done simultaneously to collect the characteristics at the same time. Typical clear-sky characteristics measured by the device

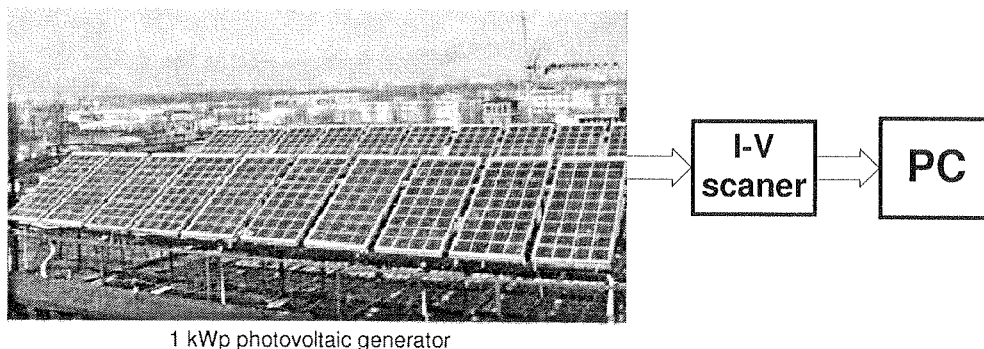


Fig. 7. I-V curve measurement system

Rys. 7. System pomiarowy charakterystyk I-V

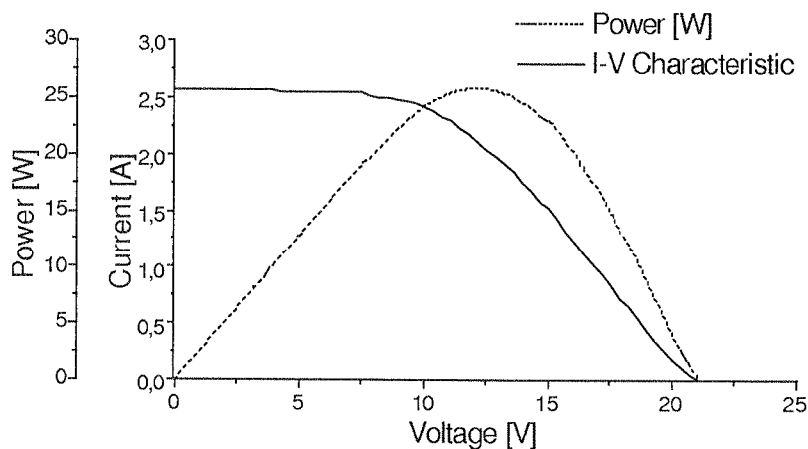


Fig. 8. I-V curve of the PV-module

Rys. 8. Przykładowa charakterystyka I-V modułu fotowoltaicznego

is shown fig. 1. This characteristics was obtained when the solar radiation was near 1000W/m^2 . The maximum power is equal to 42Wp . On the other hand, fig.8 shows the I-V characteristics of the same module under weak solar radiation. The power in the maximum point is much lower and equals to 26Wp .

More sophisticated characteristics can be seen in fig. 9. These curves were acquired for the same solar radiation as shown in fig. 1. The maximum power in this case is near 22Wp . This experiment was done for obscured PV cell in the right-upper corner of the PV module. Due to the presence of the bypass diodes, the module is still able to deliver some power, although its characteristics is distorted [4].

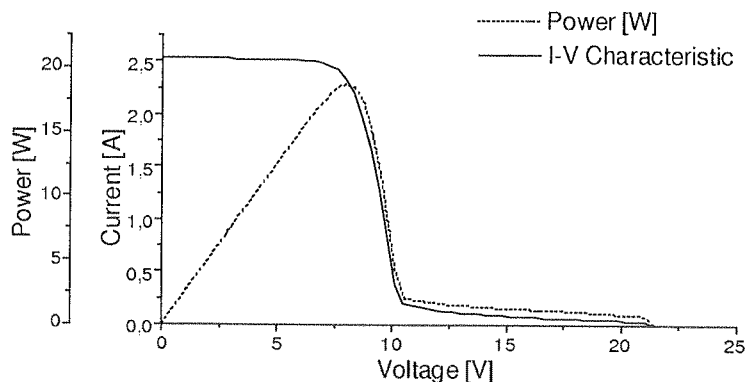


Fig. 9. I-V curve for partly shaded PV-module

Rys. 9. Przykładowa charakterystyka I-V dla częściowo zasłoniętego modułu

Fig. 10 represents similar situation. In this case, the panel produce less than 2Wp . The curves are as well deformed because of bypass diodes effect.

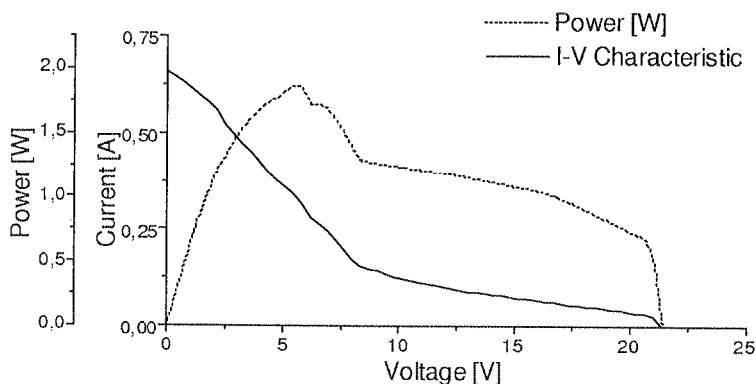


Fig. 10. I-V curve with shaded certain cell of PV-module

Rys. 10. Charakterystyki I-V modułu PV z zakrytymi ogniwami fotowoltaicznymi

6. CONCLUSIONS

In order to draw constructive conclusions there is a need to monitor the photovoltaic system for a long time. The I-V characteristics should be collected from all the panels a few times per hour [4]. Moreover, the scanning process of all the PV panels should be done simultaneously. The measurement process generates enormous amounts of data, which should be collected in a database. In our installation the presented scanner cooperates with software written under Linux. This software consists of three main modules: acquisition, database and visualisation with www user interface.

The cost-effective construction allows duplicating the scanner. In future it may be possible to integrate the device with each photovoltaic panel and collect the characteristics simultaneously.

The main disadvantage of the scanner is the resolution. In order to build as cheap device as possible, a microcontroller with integrated A/D converter was used. This fact imposes limitations for measurement accuracy. This problem is particularly bothersome for small current (represented in fig. 10). To obtain better accuracy, external A/D converter should be used, although in some cases the measurement cycle could be lengthened [1]. This situation may force to use more intelligent measurement algorithms e.g. adaptive algorithm.

7. REFERENCES

1. R. Pijarski: *System pomiarowy charakterystyk ogniw słonecznych*. M.Sc. thesis, 2002;
2. T. Żdanowicz, H. Roguszcza: *Automated Outdoor Data Acquisition System for Prolonged Testing of PV Modules*. 13th E. PV Solar Energy Conf., Nice, 1995, p. 2322
3. Edson L. Meyer, E. Ernest van Dyk: *The behaviour of photovoltaic modules under reduced light levels*. 17th E.PV Solar Energy Conf., Munich, 2001, p. 528
4. W. Maranda: *Parallel operation of non-uniformly oriented PV-arrays*. 17th E.PV Solar Energy Conf., Munich, 2001, p. 2665;
5. T. Żdanowicz: *Pomiary ogniw i modułów fotowoltaicznych*. Kazimierz Dolny, 1997, p. 159;

D. MAKOWSKI, W. MARAÑDA, R. PIJARSKI

NIEDROGI MIERNIK CHARAKTERYSTYK MODUŁÓW FOTOWOLTAICZNYCH

Streszczenie

Dokładna analiza pracy systemów fotowoltaicznych (PV) wymaga znajomości prądowo-napięciowych charakterystyk poszczególnych paneli. Koszty instalacji fotowoltaicznych maleją, jednakże specjalizowany sprzęt pomiarowy jest nadal bardzo kosztowny i pozostaje poza zasięgiem większości użytkowników instalacji PV. W artykule przedstawiono budowę przenośnego miernika charakterystyk I-V. Kluczowym elementem urządzenia jest sterowane obciążenie oraz układ pomiaru prądu i napięcia. W rozwiązaniu duży nacisk położono na wyeliminowanie kosztownych elementów takich jak szybki przetwornik cyfrowo-analogowy, co znacząco obniżyło koszty wykonania urządzenia, jednocześnie pozwalając na zachowanie

odpowiedniej szybkości pracy oraz dokładności. Zastosowane rozwiązanie opiera się na wykorzystaniu mikrokontrolera z wbudowanym przetwornikiem A/C do pomiaru prądu płynącego przez obciążenie oraz napięcia. Sterowanie zmiennym obciążeniem (tranzystor MOS), polegające na zmianie wartości napięcia na bramce tranzystora, zrealizowano metodą integracji ładunku w pojemności. Zmianę ładunku uzyskuje się impulsami generowanymi z portu mikrokontrolera. Metoda pomiaru polega na zwiększaniu napięcia na bramce tranzystora i jednoczesnym pomiarze wartości prądu i napięcia modułu fotowoltaicznego. Dane pomiarowe dla konkretnego punktu charakterystyki są rejestrowane, dopiero wtedy, gdy różnią się od poprzednich o narzuconą wartość kroku. Podejście takie nie wymaga dokładnej znajomości napięcia sterującego obciążeniem, a jedynie możliwość jego precyzyjnego zwiększania.

W celu zwiększenia szybkości działania urządzenia, po zakończeniu całego cyklu pomiarowego, pojemność tę rozładowuje się poprzez zwarcie klucza elektronicznego. Najdłuższy czas pojedynczego pomiaru nie przekracza 1s.

Miernik służy do analizy typowych modułów fotowoltaicznych o mocy do 240Wp, ograniczonej ze względu na dopuszczalną moc traconą w obciążeniu. Urządzenie wykorzystuje port szeregowy do transmisji danych do komputera PC. Długość ramki danych zmienia się w zależności od liczby pomierzonych punktów charakterystyki.

W artykule zamieszczono przykładowe wyniki pomiarów charakterystyk modułu SF52A Solar Fabrik o mocy 50Wp w różnych warunkach oświetlenia i częściowego zacienienia.

Słowa kluczowe: pomiary modułów fotowoltaicznych, charakterystyka I-V

to c
has
foll

Stat
nera
~ 1

High-power Yb-doped silica fiber laser

JACEK ŚWIDERSKI, ANDRZEJ ZAJĄC, MAREK SKÓRCZAKOWSKI, PIOTR KONIECZNY

*Instytut Optoelektroniki, Wojskowa Akademia Techniczna
ul. S. Kaliskiego 2, 00 - 908 Warszawa
E-mail: jswiders@poczta.onet.pl*

*Otrzymano 2003.06.04
Autoryzowano 2003.11.07*

In this paper high power Yb-doped double-clad fibre laser is presented. A Yb³⁺-silica fiber laser has been cladding pumped at 937 nm by a InGaAs semiconductor laser diode and operated with the slope efficiency of $73 \pm 3\%$ with respect to the incident pump power. We obtained 4 W cw-output power when working with 20-m long fibre with laser core in 7 μm diameter (cut-off wavelength $\sim 1 \mu\text{m}$) and pump core in $120 \times 120 \mu\text{m}$ (rectangle-shaped, NA = 0.6). Yb-doping concentration has been at the level $3 - 10 \times 10^{19} \text{cm}^{-3}$ and attenuation in laser core has been 8–12 dB/km. The fiber laser operated at 1084 nm.

Keywords: double-clad fiber laser, diode pumping, high-power lasers

1. INTRODUCTION

Fibre lasers belong to the family of waveguide lasers and they can operate due to diode pump. The lately observed progress in the scope of fibre lasers technology has caused them to be treated as a separate group of high-power lasers [1-3]. The following ions are usually used as active dopants of high power fibre lasers:

- neodymium (Nd³⁺) $\lambda_g \sim 1060 - 1120 \text{ nm}$,
- ytterbium (Yb³⁺) $\lambda_g \sim 1020 - 1180 \text{ nm}$,
- erbium (Er³⁺) $\lambda_g \sim 1530 - 1565 \text{ nm}$.

Fibre lasers have specific features in relation to conventional Diode Pumped Solid State Lasers (DPSSLs): high amplification, high efficiency and low threshold of generation due to very low resonator losses ($\sim 5 \text{ dB/km}$), diffraction limited beam ($M^2 \sim 1$), easy heat exchange with environment, high-power density in relation to pump

power density. Efficiency of cooling can be estimated as a ratio of a surface through which an active element is cooled (S) to its volume (V):

$$\chi = \frac{S}{V} \quad (1)$$

The ratio form (Eq. 1) depends on geometry of active materials and the way of their cooling. For the rods cooled by a side surface (this group includes also fibres) $\chi_p = 2/R$, where R is the radius of an active rod (fibre) [4].

For slabs, discs, and mirrors the active parameter of (Eq. 1) depends on quantity of the cooled surfaces of discs and slabs and it is equal to $\chi_d = t^{-1}$ for one cooled surface and $\chi_d = 2t^{-1}$ for two cooled surfaces where t is a slab thickness. Considering the fact that fibre diameter is many times lower than other dimensions of DPSSL active elements ($R_{\text{fibre}} \ll t < R_{\text{rod}}$) it is clear why fibre lasers have significantly better heat exchange with environment (10-100 times) than the lasers of typical construction [4].

A specific structure of active optical fibres and methods of its excitation are also very important issue. Fibre amplifiers and the first lasers using this geometry of active medium were excited by introducing pump energy to the core from the front face or by means of directional couplers. This type of excitation meets some difficulties when single-mode fibres have to be excited with large powers (two dimensional diode array). These difficulties can be overcome using double cladding of active fibres [5,6]. Pump radiation is introduced into the fibre in the form of cladding modes. For pump radiation, the fibre behaves as multimode fibre of $NA = 0.4 - 0.5$ [7]. Insertion of pump energy to the fibre is easy, even from laser diodes of high power, directly through the fibre front or by means of fibres bundle from many diodes with fibre outputs.

High-power fiber lasers based on double-clad rare-earth-doped silica-glasses became of strong interest during the last few years. They have found a variety of applications in industry and science as a result of their high efficiency, reliability, low cost and relatively small dimensions. Moreover, novel fiber designs like double clad fibers, M-profile fibers or large core fibers enable the generation of high output powers and diffraction-limited beam quality simultaneously [4]. In order to increase the absorption of pump light, several novel fiber geometries (offset, rectangle, D-shape, star-shape) have been developed [5,6]. This effort has resulted in successively more impressive reports of rapidly increasing fiber laser output powers, from the original tens of milliwatts to the current record of 150 W in a Yb-doped fiber laser [8]. The high-power Yb-doped fiber lasers can be used for many applications, including thermal printing, marking and other materials processing applications.

2. YTTERBIUM DOPANT

Ytterbium is one of the most versatile laser ions in silica-baser hosts. It offers very attractive futures, in particular an unusually broad absorption band that stretches from below 850 nm to ca. 980 nm and from ca. 1010 nm to above 1070 nm (Fig. 1a).

Yb^{3+} ions in silica have just one broadband laser transition, between 0.97 and 1.2 μm . Therefore, they can generate many wavelengths of general interest, e.g. for spectroscopy or for pumping other fiber lasers. The main transition of Yb^{3+} in silica is 1083 nm. It is a very useful wavelength because it can be used for optical pumping of He [9] and as a pump source for optical parametric oscillators to generate wavelengths in the 3.8 – 4.3 μm range for applications in gas detection [10].

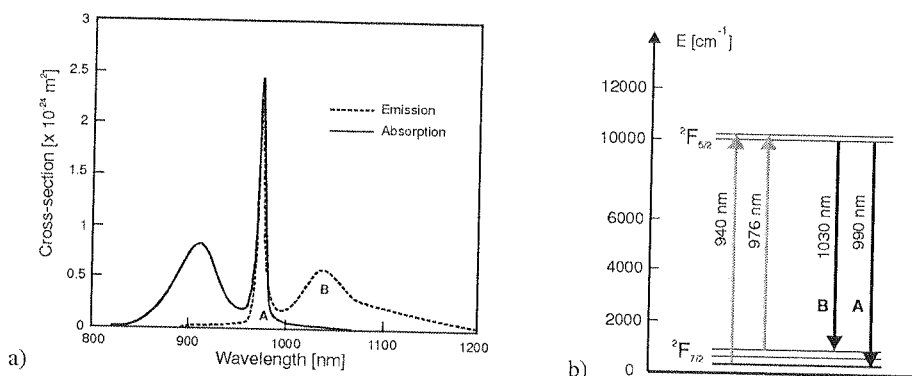


Fig. 1. The absorption and emission cross section of Yb^{3+} in silica (a) and energy level diagram of Yb^{3+} in silica (b)

Rys. 1. Absorpcyjny i emisyjny przekrój czynny jonów Yb^{3+} w szkle (a) oraz schemat poziomów energetycznych Yb^{3+} w szkle (b)

Ytterbium-doped fiber lasers can be pumped by AlGaAs ($\sim 800 - 850 \text{ nm}$) or InGaAs ($\sim 980 \text{ nm}$) semiconductor laser diodes, by Nd:YLF (1047 nm) and Nd:YAG (1064 nm) lasers, which are generally available on the industrial market. Ytterbium ions are characterized by simplicity of its energy level diagram (Fig 1b). In the Yb structure of energy level exhibits only a ground state ($^2F_{7/2}$) and a metastable state ($^2F_{5/2}$) spaced by approximately 10000 cm^{-1} . All other levels lay in the ultraviolet region. The absence of higher energy levels reduce the incidence of multi-photon relaxation and excited state absorption phenomenon (ESA). For a Yb^{3+} – doped germanosilicate glasses the measured absorption cross section at $\sim 910 \text{ nm}$ is $800 \times 10^{-27} \text{ m}^2$, at 975 nm it is $2600 \times 10^{-27} \text{ m}^2$ and at 1064 nm it is $5 \times 10^{-27} \text{ m}^2$. The Yb-doped fiber laser is typically pumped into the higher sublevels of the $^2F_{5/2}$ manifold. At wavelengths below about 990 nm , it behaves as a true three-level system (transition A in Fig. 1a), whereas at longer wavelengths, from ~ 1000 to $\sim 1200 \text{ nm}$, (transition B in Fig. 1a), it behaves as a quasi-four-level system.

Because of absence of higher energy levels, high concentration of Yb^{3+} are possible in silica-based hosts, often up to several thousand ppm. Nowadays Yb^{3+} dopant is regarded as the main dopant of high-power fiber lasers.

3. MEASUREMENT SET-UP

The aim of the experimental research was to measure output power versus absorbed power, as well as to estimate the level of slope efficiency of the ytterbium-doped fiber laser. To reach this goal we built a measurement set-up, that was presented in Fig. 2. We used a ytterbium-doped silica double-clad optical fiber as an active medium. The inner cladding had a square shape with $120\mu \times 120\mu$ to match the dimension of pump laser diode. The square geometry of inner cladding was designed to prevent whispering modes where light spirals within the cladding without entering the core, thus promoting more rapid and efficient pump light transfer into the core. Background losses in the inner cladding (for pump wavelength) was 20 – 40 dB/km. Numerical aperture for pump radiation equalled $NA = 0.6$. The laser core was doped with $3 - 10 \times 10^{19} \text{ cm}^{-3}$ ytterbium dopant and had a diameter of 6.5 – 7.5 μm (cut-off wavelength – 0.9 – 1.0 μm), resulting in transversal single-mode wave guiding for wavelength above 1 μm . Background losses in the active core (for laser radiation) was 8 – 12 dB/km. The fiber length was 20 m.

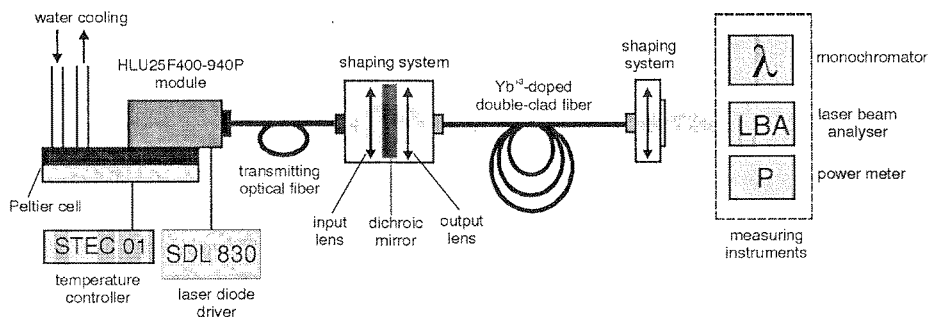


Fig. 2. Experimental set-up of Yb-doped fiber laser

Rys. 2. Układ eksperymentalny włóknowego lasera iterbowego

As a pump source for the fiber a high-power diode laser module (with additional transmitting optical fiber) generating at a wavelength of 937 nm with a special diffractive optic was used. This optic transformed the beam to a spot size of less than 400 μm to couple the pump light into the transmitting fiber. The transmitting fiber (1.5 m length) was characterised by numerical aperture $NA = 0.22$. The laser diode module (HLU25F400-940P) was produced by LIMO Micro-Optics & Laser Systems (Jena, Germany) and delivered 25 W continuous output power. Pump source was cooled by water cooling system. To control the working temperature we used a Peltier cell. Pump radiation coming out of transmitting fiber, thanks to shaping system (two aspheric lenses, AR-coated) was inserted into the active optical fiber. To protect laser diode module from destruction by laser radiation we used a dichroic mirror, which was highly transmittive for wavelength between 915 and 980 nm and highly reflective for wavelengths between 1000 and 1100 nm. Pump light transmission of the shaping

system equalled 84 %. The used aspheric lenses set wasn't optimum, mainly due to financial constrains. Nevertheless, the shaping system guaranteed ca. 30 % launching of pump light to the active fiber. In this way, a fiber Fabry-Perot resonator with cavity reflectivity of ~ 4 % (Fresnel reflection) and more than 99 % was formed. The output power generated from the single mode core of the fiber was measured after passing through the output objective lens.

4. EXPERIMENTAL RESULTS

The fiber laser output power as a function of the absorbed pump power is shown in Fig. 3. We obtained 4 W cw output power for 20 m ytterbium-doped optical fiber. For this laser system configuration the slope efficiency was 72 ± 3 %.

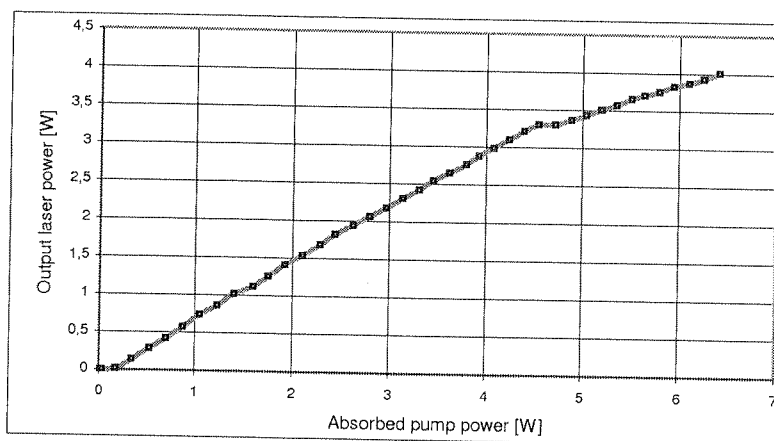


Fig. 3. Measured output power from Yb-doped silica fiber laser as a function of absorbed pumped power

Rys. 3. Charakterystyka zmierzonej mocy wyjściowej włóknowego lasera iterbowego w funkcji zaabsorbowanej mocy pobudzającej

As it is shown in Fig. 3 along with the increase of absorbed pump power (till ca. 4.5 W) the output laser power increases proportionally linear. After crossing 4.5 W of absorbed pump power we can observe the decrease of increment of laser power. This phenomenon was caused by limited ability of cooling system. Moreover, the shaping system guaranteed ca. 30 % launching of pump light to the active fiber. Theoretically we calculated that for this system configuration it is possible to couple more than 70 % of pump light. In the near future, using the suitably designed shaping system we are planning to obtain more than 10 W output laser power.

The measurement of laser emission spectrum width was presented in Fig. 4. The laser worked at 1084 nm wavelength and the spectral width of its emission was 7 nm.

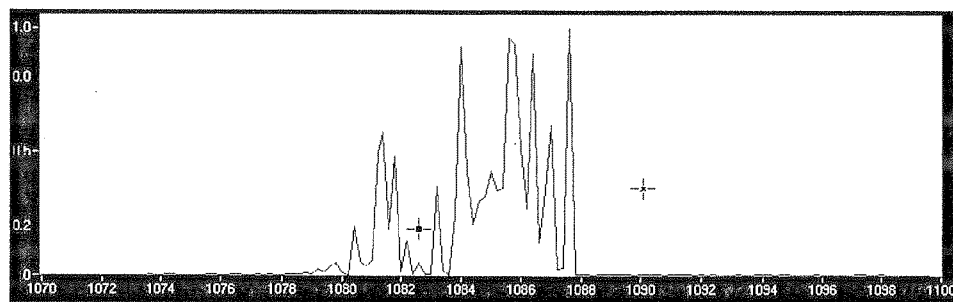


Fig. 4. Measured laser emission spectrum width of Yb-doped fiber laser

Rys. 4. Pomiar widma emisji włóknowego lasera iterbowego

5. CONCLUSIONS

In conclusion, we have demonstrated a Yb^{3+} – doped double-clad fiber laser generated at 1084 nm. The output laser power was 4 W with $72 \pm 3\%$ slope efficiency. We are planning to increase a level of output power as well as slope efficiency of examined laser system mainly by means of improving excitation efficiency and by increasing the number of pumping LEDs.

By their nature double-clad fiber lasers offer advantages over conventional laser systems due to their ability to incorporate a laser medium and a laser power delivery system into a single element. In the future, new fiber laser designs and pumping configurations will provide users with low divergence, high brightness, high power laser devices that overcome the serious drawbacks, limitations and deficiencies of conventional laser system. However, double clad fiber lasers, especially ytterbium doped fiber lasers will allow to construct devices delivering kilowatt output power.

6. ACKNOWLEDGEMENTS

This work was partially supported by Polish State Committee for Scientific Research under the project 4T11B07722. The authors would like to thank Jan Karol Jabczynski and Waldemar Zendzian for their constructive comments.

7. REFERENCES

1. J. Zautier, E. Peik, H. Wather, A. Y. Nevsky, M. N. Srivortsow, S. N. Bagaev: *An ultra frequency stable diode pumped Nd:YAG laser at 946 nm*. Technical Digest: Novel Lasers and Devices - Basic Aspects, Munich, 1999, paper LTuD5-1, pp. 115–117.
2. Y. Akiyama, T. Takase, M. Sasaki, A. Takada, H. Yuasa, A. Ono: *High power all solid state laser technology: rod type laser*. Proc. of The 3rd Symposium on Advanced Photon Processing and Measurement Technologies, Tokyo, 1999, pp. 18–19.

3. S. G. Anderson: *Review and forecast of laser markets — part 1*. Laser Focus World, vol. 34, 1999, pp. 80–100.
4. Z. Jankiewicz: *Diode-pumped solid state lasers*. Opto- Electronics Review, vol. 9, no. 1, 2001, pp. 27.
5. N. S. Kim, T. Hamada, M. Prabhu, C. Li, J. Song, K. Ueda, A. Liu, H. J. Kong: *Numerical analysis and experimental results of output performance for Nd-doped double-clad fiber lasers*. Optics Communications 180, 2000, pp. 329–337.
6. L. Philippe, V. Doya, R. Philippe, P. Dominique, M. Fabrice, L. Olivier: *Experimental study of pump power absorption along rare-earth doped double clad optical fibers*. Optics Communications 218, 2003, pp. 249– 254.
7. S. J. Matthews: *Bright light from glass wires*. Laser Focus World, vol. 38, no. 1, 2002, 2002, pp. 131–136.
8. J. Limpert, A. Liem, S. Hofer, H. Zellmer, A. Tennermann: *150 W Nd/Yb codoped fiber laser at 1.1 μm* . Proc. CLEO'2002, paper CThX1, p. 590.
9. J. Hamel, A. Cassini, H. Abu-Safia, M. Leduc: Optics Communication 63, 1987, pp. 114.
10. L. Goldberg, J. P. Koplov, R. P. Moeller, D. A. V. Kliner: *High power superfluorescent source with a side-pumped Yb-doped double-cladding fiber*. Optics Letters, 23, 1998, pp. 1037–1039.

J. ŚWIDERSKI, A. ZAJĄC, M. SKÓRCZAKOWSKI, P. KONIECZNY

ITERBOWY LASER WŁÓKNOWY DUŻEJ MOCY

Streszczenie

Artykuł porusza zagadnienie praktycznej realizacji układu włóknowego lasera dużej mocy. Zagadnienia związane z badaniami, opracowaniem i konstrukcją tego typu układów stanowią w ostatnim czasie jeden z istotniejszych obszarów badań stosowanych w technice laserowej. Układy laserowe zbudowane w oparciu o włókna aktywne typu dwupłaszczowego charakteryzują się m. in. wysoką sprawnością generacji, rozbieżnością wiązki generowanej zbliżoną do rozbieżności dyfrakcyjnej, wysokim poziomem mocy wyjściowej, brakiem konieczności wymuszonego chłodzenia cieczonego.

W artykule zaprezentowany został praktyczny układ lasera włóknowego dużej mocy zbudowanego w oparciu o włókno typu „double-clad” domieszkowane jonami iterbu. W prezentowanym układzie (rys. 2) jako źródło promieniowania pompującego użyto lasera półprzewodnikowego produkcji niemieckiej (LIMO — Lissotschenko Mikrooptok GmbH) emitującego maksymalnie 25 W mocy ciągłej na długości fali 937nm. Dioda laserowa zasilana była z układu zasilającego, umożliwiającego pracę ciągłą oraz impulsową (modulacja wewnętrzna — układ sterowania typu SDL830). Wiązka światła generowanego wprowadzana była do włókna aktywnego poprzez dodatkowy światłowód transmisyjny. Dzięki temu nie posiadała ona, typowej dla diod laserowych, asymetrii. Dioda LIMO chłodzona była chłodziarką Peltiera oddającą ciepło do chłodnicy wodnej. Regulację temperatury lasera pompującego realizowano za pomocą chłodziarki termoelektrycznej (STEC 01).

Na wyjściu układu lasera uzyskano 4W mocy ciągłej przy sprawności różniczkowej na poziomie $73 \pm 3\%$. Zastosowane włókno aktywne charakteryzowało się rdzeniem o średnicy $7\mu\text{m}$, płaszczem wewnętrznym w kształcie kwadratu ($120 \times 120\mu\text{m}$) oraz koncentracją domieszki czynnej na poziomie $3 - 10 \times 10^{19} \text{ cm}^{-3}$. Laser generował promieniowanie na długości fali 1084nm.

Słowa kluczowe: laser włóknowy typu „double-clad”, pompowanie diodowe, lasery dużej mocy

INFORMACJE DLA AUTORÓW

Redakcja przyjmuje do publikowania prace oryginalne, przeglądowe i monograficzne wchodzące w zakres szeroko pojętej elektroniki. Ponieważ KWARTALNIK ELEKTRONIKI I TELEKOMUNIKACJI jest czasopismem Komitetu Elektroniki i Telekomunikacji Polskiej Akademii Nauk, w związku z tym na jego łamach znajdują się prace naukowe dotyczące podstaw teoretycznych i zastosowań z zakresu elektroniki, telekomunikacji, mikroelektroniki, optoelektroniki, radiotechniki i elektroniki medycznej.

Artykuły powinny charakteryzować oryginalne ujęcie zagadnienia, własna klasyfikacja, krytyczna ocena (teorii lub metod), omówienie aktualnego stanu, lub postępu danej gałęzi techniki oraz omówienie perspektyw rozwojowych. Artykuły publikowane w innych czasopismach nie mogą być kierowane do druku w Kwartalniku Elektroniki i Telekomunikacji w drugiej kolejności zgłoszenia.

Objętość artykułu nie powinna przekraczać 30 stron po około 1800 znaków na stronie, w tym rysunki i tabele.

Wymagania podstawowe.

Artykuły należy nadsyłać na wyraźnym, jednostronnym, czarno-białym wydruku komputerowym. Wydruk w formacie A4 powinien mieć znormalizowaną liczbę wierszy i znaków w wierszu (30 wierszy po 60 znaków w wierszu), w dwóch egzemplarzach, w języku polskim lub angielskim wybranym przez autora. Do wydruku powinna być dołączona dyskietka z elektronicznym tekstem artykułu. Preferowane edytory to WORD 6 lub 8. Układ artykułu (w wersji podstawowej) musi być następujący:

- Tytuł.
- Autor (imię i nazwisko autora/ów).
- Miejsce pracy (nazwa instytucji, miejscowość, adres. + ew. adres elektroniczny (e-mail)).
- Zwięzłe streszczenie powinno być w języku takim, w jakim jest pisany artykuł (wraz ze słowami kluczowymi).
- Tekst podstawowy powinien mieć następujący układ:

1. WPROWADZENIE

2. np. TEORIA

3. np. WYNIKI NUMERYCZNE

3.1.

3.2.

4.

5.

6. PODSUMOWANIE

7. ew. PODZIĘKOWANIA

8. BIBLIOGRAFIA

- Układ streszczenia w dodatkowej wersji językowej powinien być następujący:

AUTOR (inicjał imienia i nazwisko).

TYTUŁ (w języku angielskim – o ile artykuł pisany jest w języku polskim i na odrwót).

Obszerne do 3600 znaków streszczenia (wraz z słowami kluczowymi) w języku:

a. angielskim, gdy artykuł pisany jest w języku polskim.

b. polskim, gdy artykuł pisany jest w języku angielskim.

Streszczenie to powinno pozwolić czytelnikowi na uzyskanie istotnych informacji zawartych w pracy. Z tego względu w streszczeniu tym mogą być cytowane numery istotnych wzorów, rysunków i tabel zawartych w podstawowej wersji językowej.

- Wszystkie strony muszą mieć numerację ciągłą.

Sposób pisania tekstu.

Tekst powinien być pisany bez używania wyróżnień, a w szczególności nie dopuszcza się spacjiowania, podkreślania i pisania tekstu dużymi literami z wyjątkiem wyrazów, które umownie pisze się dużymi literami (np. FORTRAN). Proponowane wyróżnienia Autor może zaznaczyć w maszynopisie zwykłym ołówkiem za pomocą przyjętych znaków adjustacyjnych, np. podkreślenie linią przerywaną oznacza spacjiowanie (rozstrzelenie), podkreślenie linią ciągłą – pogrubienie, podkreślenie wężykiem — kursywa.

Tekst powinien być napisany z podwójnym odstępem między wierszami, tytuły i podtytuły małymi literami. Marginesy z każdej strony powinny mieć około 35 mm. Wielkość czcionki wydruku powinna być zbliżona co najmniej co wielkości czcionki maszyny do pisania (minimum 12 punktów). Przy podziale pracy na rozdziały i podrozdziały cyfrowe ich oznaczenia nie powinny być większe niż II stopnia (np. 4.1.1.).

Sposób pisania tabel.

Tabele powinny być pisane na oddzielnych stronach. Tytuły rubryk pionowych i poziomych powinny być napisane małymi literami z podwójnym odstępem między wierszami. Przypisy (notki) dotyczące tabel należy pisać bezpośrednio pod tabelami. Tabele należy numerować kolejno liczbami arabskimi, u góry każdej tabeli podać tytuł dwujęzyczny. W pierwszej kolejności w podstawowej wersji językowej, a później w dodatkowej wersji językowej. Tabele umieścić na końcu maszynopisu. Przyjmowane są tabele algorytmów i programy na wydrukach komputerowych. W tym przypadku zachowany jest ich oryginalny układ. Tabele powinny być cytowane w tekście.

Sposób pisania wzorów matematycznych.

Rozmieszczenie znaków, cyfr, liter i odstępów powinno być zbliżone do rozmieszczenia elementów druku. Wskaźniki i wykładniki potęg powinny być napisane wyraźnie i być prawidłowo obniżone lub podwyższone w stosunku do linii wiersza podstawowego. Znaki nad literami i cyframi, całkami i in. symbolami (strzałki, linie, kropki, daszki) powinny być umieszczone dokładnie nad tymi elementami, do których się odnoszą. Numery wzorów cyframi arabskimi powinny być kolejne i umieszczone w nawiasach okrągłych z prawej strony. Nazwy jednostek, symbole literowe i graficzne powinny być zgodne z wytycznymi IEE (International Electrotechnical Commission) oraz ISO (International Organization of Standardization).

Powołania.

Powołania na publikacje powinny być umieszczone na ostatnich stronach tekstu pod tytułem „Bibliografia”, opatrzone numeracją kolejną bez nawiasów. Numeracja ta powinna być zgodna z odnośnikami w tekście artykułu. Przykłady opisu publikacji:

- periodycznej 1. F. Valdoni: A new milimetre wave satelite. E.T.T. 1990, vol. 2, no 5, pp. 141–148
- nieperiodycznej 2. K. Andersen: A resource allocation framework. XVI International Symposium, Stockholm (Sweden), may 1991, paper A 2,4
- książki 3. Y.P. Tvidis: Operation and modeling of the MOS transistors. New York, McGraw-Hill, 1987, p. 553

Materiały ilustracyjne.

Rysunki powinny być wykonane wyraźnie, na papierze gładkim lub milimetrowym w formacie nie mniejszym niż 9 × 12 cm. Mogą być także w postaci wydruku komputerowego (preferowany edytor Corel Draw). Fotografie lub diapozytywy przyjmowane są raczej czarno-białe w formacie nie przekraczającym 10 × 15 cm. Na marginesie każdego rysunku i na odwrocie fotografii powinno być napisane ołówkiem imię i nazwisko Autora oraz skrót tytułu artykułu, do którego są przeznaczone oraz numer rysunku. Spis podpisów pod rysunki i fotografie powinien być umieszczony na oddzielnej stronie. Podpisy pod rysunkami (fotografiami) powinny być dwujęzyczne: w pierwszej kolejności w podstawowej wersji językowej, a później w dodatkowej wersji językowej. Rysunki powinny być cytowane w tekście.

Uwagi końcowe.

Na odrębnej stronie powinny być podane następujące informacje:

- adres do korespondencji z kodem pocztowym (domowy lub do miejsca pracy),
- telefon domowy i/lub do miejsca pracy,
- adres e-mailowy (jeśli autor posiada).

Autorowi przysługuje bezpłatnie 20 odbitek artykułu. Dodatkowe egzemplarze odbitek, lub cały zeszyt Autor może zamówić u wydawcy na własny koszt.

Autora obowiązują korekta autorska, którą powinien wykonać w ciągu 3 dni od daty otrzymania tekstu z Redakcji oraz zwrócić osobiście lub listownie pod adres Redakcji. Korekta powinna być naniesiona na przekazanych Autorowi szpaltach na marginesach ew. na osobnym arkuszu w przypadku uzupełnień tekstu większych niż dwa wiersze. W przypadku nie zwrócenia korekty w terminie, korektę przeprowadza Redakcja Techniczna Wydawcy. Redakcja prosi Autorów o powiadomienie ją o zmianie miejsca pracy i adresu prywatnego.

INFORMATION FOR AUTHORS OF K.E.T.

The editorial staff will accept for publishing only original monographic and survey papers concerning widely understood electronics. Because of the fact that KWARTALNIK ELEKTRONIKI I TELEKOMUNIKACJI is a journal of the Committee for Electronics and Telecommunications of Polish Academy of Science, it presents scientific works concerning theoretical bases and applications from the field of electronics, telecommunications, microelectronics, optoelectronics, radioelectronics and medical electronics.

Articles should be characterised by original depiction of a problem, its own classification, critical opinion (concerning theories or methods), discussion of an actual state or a progress of a given branch of a technique and discussion of development perspectives.

An article published in other magazines can not be submitted for publishing in K.E.T. The size of an article can not exceed 30 pages, 1800 character each, including figures and tables.

Basic requirements

The article should be submitted to the editorial staff as a one side, clear, black and white computer printout in two copies. The article should be prepared in English or Polish. Floppy disc with an electronic version of the article should be enclosed. Preferred wordprocessors: WORD 6 or 8.

Layout of the article.

- Title.
- Author (first name and surname of author/authors).
- Workplace (institution, adress and e-mail).
- Concise summary in a language article is prepared in (with keywords).
- Main text with following layout:
 - Introduction
 - Theory (if applicable)
 - Numerical results (if applicable)
 - Paragraph 1
 - Paragraph 2
 -
 -
 - Conclusions
 - Acknowledgements (if applicable)
 - References
- Summary in additional language:
 - Author (firs name initials and surname)
 - Title (in Polish, if article was prepared in English and vice versa)
 - Extensive summary, hawever not exceeching 3600 characters (along with keywords) in Polish, if artide was prepared in English and vice versa). The summary should be prepared in a way allowing a reader to obtaoin essential information contained in the artide. For that reason in the summary author can place numbers of essential formulas, figures and tables from the article.

Pages should have continues numbering.

Main text

Main text can not contain formatting such as spacing, underlining, words written in capital letters (except words that are commonly written in capital letters). Author can mark suggested formatting with pencil on the margin of the article using commonly accepted adjusting marks.

Text should be written with double line spacing with 35 mro left and right margin. Titles and subtitles should be written with small letters. Titles and subtitles should be numberd using no mor than 3 levels (i.e. 4.1.1.).

Tables

Tables with their titles should be places on separate page at the end of the article. Titles of rows and columns should be written in small letters with double line spacing. Annotations concerning tables

should be placed directly below the table. Tables should be numbered with Arabic numbers on the top of each table. Table can consist algorithm and program listings. In such cases original layout of the table will be preserved. Table should be cited in the text.

Mathematical formulas

Characters, numbers, letters and spacing of the formula should be adequate to layout of main text. Indexes should be properly lowered or raised above the basic line and clearly written. Special characters such as lines, arrows, dots should be placed exactly over symbols which they are attributed to. Formulas should be numbered with Arabic numbers placed in brackets on the right side of the page. Units of measure, letter and graphic symbols should be printed according to requirements of IEE (International Electrotechnical Commission) and ISO (International Organisation of Standardisation).

References

References should be placed at the end of the main text with the subtitle „References“. References should be numbered (without brackets) adequately to references placed in the text. Examples of periodical [1], non-periodical [2] and book [3] references:

1. F. Valdoni: A new millimetre wave satellite. E.T.T. 1990, vol. 2, no 5, pp. 141–148
2. K. Anderson: A resource allocation framework. XVI International Symposium (Sweden). May 1991. paper A 2.4
3. Y.P. Tividis: Operation and modeling of the MOS transistors. New York. McGraw-Hill. 1987. p. 553

Figures

Figures should be clearly drawn on plain or millimetre paper in the format not smaller than 9×12 cm. Figures can be also printed (preferred editor – CorelDRAW). Photos or diapositives will be accepted in black and white format not greater than 10×15 cm. On the margin of each drawing and on the back side of each photo author's name and abbreviation of the title of article should be placed. Figure's captions should be given in two languages (first in the language the article is written in and then in additional language). Figure's captions should be also listed on separate page. Figures should be cited in the text.

Additional information

On the separate page following information should be placed:

- mailing address (home or office),
- phone (home or/and office),
- e-mail.

Author is entitled to free of charge 20 copies of article. Additional copies or the whole magazine can be ordered at publisher at the author's expense.

Author is obliged to perform the author's correction, which should be accomplished within 3 days starting from the date of receiving of the text from the editorial staff. Corrected text should be returned to the editorial staff personally or by mail. Correction marks should be placed on the margin of copies received from the editorial staff or if needed on separate pages. In the case when the correction is not returned in said time limit, correction will be performed by technical editorial staff of the publisher.

In case of changing of workplace or home address Authors are asked to inform the editorial staff.