

Departamento de Señales, Sistemas y Radiocomunicaciones

Escuela Técnica Superior de Ingenieros de Telecomunicación

Tesis Doctoral

RF Front-End Non-Linear Coupling Cancellation

Autor: **Igor Arambasic, Ingeniero de Telecomunicación**

Director: **Francisco Javier Casajús de Quirós**

Madrid, 2008

TESIS DOCTORAL
RF FRONT-END NON-LINEAR COUPLING CANCELLATION
CANCELACIÓN DE ACOPLO NO LINEAL EN FRONTALES DE RF

AUTOR: Igor Arambasic

DIRECTOR: Francisco Javier Casajús de Quirós

El tribunal nombrado por el Mgfco. y Excmo. Sr. Rector de la Universidad Politécnica de Madrid, el día dede 2008.

PRESIDENTE:

VOCAL:

VOCAL:

VOCAL:

SECRETARIO:

Realizado el acto de defensa y lectura de la Tesis el día de.....de 2008,
en Madrid.

Calificación:

EL PRESIDENTE

LOS VOCALES

EL VOCAL SECRETARIO

Acknowledgments

I must admit I was going to write just one sentence of acknowledgments. However, I had to give up that idea because, even if I forget mentioning someone here I'll do my best to name most of the persons that I am thankful and that helped on professional or personal level to bring this work to the end. It took six years to finish it, hence it is hard to mention individually all the persons that influenced me one way or another. Still, I must name Anita, Mate, Jon, Galo, Mariano, Santiago, Iván and Narciso who made my life in Spain and my stay at UPM a lot more enjoyable. Mentioning joy I cannot forget Aleksandra, a friend I could always count on. Our discussions were mostly on personal level but often included topics about professional future which were always useful and important to me. One other important thing... now that I finished the thesis..... I hope you won't stop bringing me *pljeskavice*:-).

I would also like to express my most sincere gratitude to Filip Vucak who was kind enough to help a stranger in need, and gave me the opportunity to work on the thesis while serving the civil service. Besides Mr. Vucak, I must thank to Maja, Branka and Tamara who transformed not so delightful serving duty into a pleasant working place. Without their support and understanding this PhD. student would probably given up midway the road.

I am thankful to my mentor Javier not only for his thesis guidance but, even more because it was not only limited to professional side but he also generously offered useful advices concerning personal life situations.

I cannot avoid mentioning two closest friends of mine, Davor and Iris, who influenced me immensely on a personal level. Even though we were talking about this thesis many times, I am sure none of you have the slightest idea what it really is about. Actually, it doesn't matter, thank you for showing interest, for listening and backing me up in many situations throughout the years.

I should also mention Vida and Jurica who helped me tremendously, especially in the last year when the thesis was coming to its end, and the stress was increasing. Thank you for adjusting your life according to my needs. Sorry for occupying your home computer and not letting you play *kuglice* in your own house while my simulations were running.

Naturally, without the love and encouragement of my parents, this thesis would never be finished. Hence.. Dragi roditelji, oduvijek ste bili moj uzor i inspiracija. Hvala vam na bezu-

vjetnoj ljubavi koju ste mi pruzili. Hvala sto ste uvijek bili uz mene. Hvala na razumijevanju puta koji sam odabral i na podrscu koju mi dajete kroz zivot.

Finally, I am deeply grateful to my family for their support. Ivana, you are the person who played a decisive role in keeping the thesis alive. Thank you for your enthusiastic support, for your love, patience and your understanding. Goran, although you probably won't remember much of this, thank you for putting up with many lost weekends and odd working hours which were hard to understand. I'll do my best to make it up to you and I promise it won't happen again.

Contents

1 Resumen (Summary in Spanish)	5
2 Motivation and background	27
3 RF Front-end description	36
3.1 Traditional RF Front-End Architecture	38
3.2 Future Radio System	40
3.2.1 Ideal Software Radio	42
3.2.2 Feasible SWR Receiver	48
3.3 Typical single antenna RF Front-end Imperfections	50
3.3.1 DC offset	50
3.3.2 Flicker noise	54
3.3.3 Second Order Intermodulation	56
3.3.4 I/Q Imbalance	57
3.4 Multi-antenna RF Front-end	60
4 Direct Coupling Model	64
4.1 Nonlinear Coupling Description	64
4.2 Multi-antenna RF front-end coupling model	67
4.3 Influence of RF front-end nonlinear coupling on the output of I/Q downconverter	69
4.4 The power of the coupling interference inside RF front-end	73

5 Inverse coupling function	78
5.1 Analytical solution	78
5.2 Surface fitting approximations	87
5.3 System description	100
6 Decoupling Module Search Algorithms	107
6.1 Simulated Annealing	108
6.1.1 Adopting SA to Nonlinear Decoupling Problem	110
6.1.2 Fine tuning of SA parameters	124
6.2 Improved Fast Simulate Annealing (IFSA)	142
6.2.1 Fine tuning of IFSA parameters	149
6.3 Guided Multi-level (GML) search	158
6.3.1 Describing GML search approach	159
6.3.2 Adopting GML search to nonlinear decoupling	167
7 Noiseless Calibration	178
7.1 Decoupling Degrees of Freedom	178
7.2 Calibration size	190
8 Calibration with noise	197
8.1 Adaptation of surface form and calibration size	197
8.2 Algorithm Comparison	203
9 Conclusions	209
References	218

Abstract

In order to extend the communication range, improve reliability and achieve higher bit rates recent trends in communication industry include the use of multiple antennas in transmitter and at the receiver. At the same time the dimensions of mobile terminals are getting smaller incrementing the level of electronic interferences in their interior. The closeness of data metallic wires and many electronic circuits creates rather hostile environment which, among other distortions, stimulates the appearance of undesirable energy transfers between physical mediums. This signal distortion, also known as multiantenna radio-frequency (RF) front-end coupling is analyzed throughout the thesis, and as a result three coupling cancellation procedures are developed and later tested in simulated coupling environments.

The coupling phenomena depends on many variables, starting with the selection of electronic elements and RF front-end design, up to the transmission system working conditions. Since the complexity of coupling distortion discards simple linear models, nonlinear modelling based on polynomials, limited to third order, is implemented. The coupling model proposed in this thesis is consistent with the analog electronic elements and follows the nonlinear behaviour experienced in many RF front end imperfections.

The coupling cancellation approach is based on nonlinear software decoupling module located at the output of analog-to-digital (AD) converter. The advantage of this location is that it deals directly with sampled physical data avoiding any dependency on specific signal type. Hence the decoupling module can be easily implemented in wide range of transmission systems. Inside the module, signal enhancement is achieved with the approximation of inverse nonlinear coupling function based on the sum of two independent polynomials modelling the inverse function of coupling between two antennas. The adequate decoupling surface is obtained with three different search methods: Simulated Annealing (SA), Improved Fast Simulated Annealing (IFSA) and Guided Multi-Level (GML) search.

The thesis is divided into eight chapters. The first two correspond to technological background and RF front end description, respectively. In chapter three, nonlinear coupling model is developed with emphasis on receiver's RF front-end based on I/Q downconverter. Chapter four introduces the inverse coupling function and develops the adequate approximation surface model used for coupling cancellation. The three approaches used for the corresponding surface search are than analyzed in chapter five. The following two chapters first simplify the surface

form in noiseless conditions, and afterwards the performance of adopted search methods is analyzed in environment with AWGN noise during the calibration. The last chapter corresponds to conclusion and possible future applications of decoupling module.

RF Front-End Non-Linear Coupling Cancellation

1 Resumen (Summary in Spanish)

Para lograr coberturas más amplias, mejorar la fiabilidad y alcanzar tasas de transmisión altas, la tendencia reciente en la industria de telecomunicaciones es la del uso de las múltiples antenas tanto en el transmisor como en el receptor. Otra tendencia clara que se observa en la actualidad es la miniaturización de los terminales móviles. En la etapa de radio frecuencia (RF), dentro del frontal de RF, estas dos tendencias juntas resultan en las distorsiones de las señales recibidas por diferentes antenas dando lugar al acoplamiento no lineal entre las antenas. El acoplamiento aparece por la transmisión no deseada de la energía entre las dos antenas. El fenómeno de acoplamiento depende de muchos elementos, empezando por la selección de los elementos electrónicos y el diseño del frontal de RF, hasta llegar a las condiciones del entorno en el que funciona el sistema de comunicaciones. La distorsión por acoplamiento entre antenas es, por la naturaleza de los elementos analógicos que componen el frontal de RF, de una complejidad importante y su aproximación con modelos lineales no es apropiada. En esta tesis se ha desarrollado un modelo polinomial que describe el acoplamiento entre dos antenas. Se ha utilizado el modelo que llega hasta la tercera potencia, ya que se ha mostrado bastante bueno para modelar las medidas de acoplamiento disponibles.

El trabajo fundamental de esta tesis es la cancelación de la distorsión no lineal entre las antenas. La atención se ha centrado en el caso cuando el receptor es el terminal móvil, y no se puede disponer de procesadores con cargas computacionales muy altas. En esta situación, el cancelador de acoplamiento tiene que ser un modulo sencillo desde el punto de vista computacional, pero que es capaz de cancelar o minimizar el acoplamiento no lineal de la señal recibida. La cancelación se lleva a cabo en un modulo que se encuentra después del conversor analógico digital (AD) en la cadena receptora. Esta posición permite que se procesen las muestras de la señal sin tener en cuenta el tipo específico de la señal recibida

(si es de tipo OFDM, AM, etc). Esta independencia del tipo de la señal habilita el uso de este módulo en varios tipos de sistemas de comunicaciones. El proceso de cancelación de acoplamiento se lleva a cabo basándose en el modelo de dos polinomios independientes con los que se aproxima el inverso de la función de acoplamiento. La superficie de la función de cancelación de acoplamiento se obtiene con tres diferentes métodos de búsqueda: Simulated Annealing (SA), Improved Fast Simulated Annealing (IFSA) y Guided Multi-Level (GML) search.

La tesis está dividida en 8 secciones aparte de esta. En las primeras dos se revisan el historial tecnológico, se identifican las tendencias en la arquitectura de los terminales móviles, y se describe el frontal de RF. En la tercera sección se introduce el modelo de acople con el especial énfasis en el sistema con el conversor I/Q. La cuarta sección se ha dedicado a la introducción y el análisis del modelo de cancelador de acople, mientras que en la quinta sección se analizan en detalle los tres algoritmos propuestos para la búsqueda de la superficie adecuada. En las últimas dos secciones se simplifican las formas de las superficies cuando el entrenamiento se realiza con y sin ruido gaussiano aditivo. La última sección se ha dedicado a las conclusiones y a las posibles líneas de investigación futuras sobre la cancelación de acople no lineal. A continuación se presentará un breve resumen por cada sección.

Sección 2

Al principio del vigésimo primer siglo, los teléfonos móviles se encuentran por todas partes. Actualmente, alrededor de 80% de la población mundial dispone de la cobertura para los servicios móviles, y se espera que la cobertura llegará al 90% en el año 2010. Su presencia en la vida cotidiana ya es tan corriente que hay varios países europeos con mayor número de móviles que de residentes. El número de móviles que se utilizan en el mundo es una incógnita, pero, según los pronósticos de un analista de telecomunicaciones, Paul Budde, se alcanzó el número de 3 mil millones de móviles en septiembre de 2007 [Bud07]. Este dato de número de móviles sigue cambiando a ritmo de 15% al año, y la industria de telecomunicaciones mantiene este incremento a base de ofrecer constantemente nuevas aplicaciones y servicios y las mejoras en las prestaciones. Algunas de las aplicaciones que se ofrecen actualmente (radio FM, cámara y cámara de vídeo, MP3, juegos de tipo PlayStation, los programas de navegación por GPS, TV en directo, etc) se pueden ver en la Figura 1. Este incremento en el número de usuarios de los móviles muestra la razón del gran interés en sus futuras mejoras. En esta tesis se



Figure 1: Los posibles servicios integrados en un terminal móvil

ha añadido un grano de arena para hacer que esta tendencia siga cumpliéndose, intentando resolver uno de los problemas con los que se encontrarán los terminales móviles avanzados.

En la sección introductoria se ha revisado la historia de evolución de los terminales móviles, llegando al momento presente en el cual dentro de los terminales existe un entorno muy propenso a las interferencias entre señales en el interior del móvil, lo que favorece acoplamiento. La apariencia de acoplamiento es debida a la miniaturización de los terminales, al incremento de la potencia de procesado y a la introducción de múltiples antenas y muchos otros componentes que no están relacionados directamente con el proceso de transmisión y recepción de señal. A la vez, hay que tener en mente que el mercado es competitivo, por lo que el precio de los componentes utilizados debe ser lo más reducido. Esto lleva a que el interior del terminal esté lleno de los componentes electrónicos baratos que tienden a disipar energía, lo que resulta en un entorno hostil. Además, los componentes digitales están conectados con un número elevado de los cables metálicos y diferentes elementos analógicos que crean los campos magnéticos, incrementan la temperatura y favorecen la aparición de diferentes distorsiones de la señal. Estas características de los terminales móviles son precisamente las que resultan en la aparición de acoplamiento: por la falta de espacio la protección que minimiza la disipación de

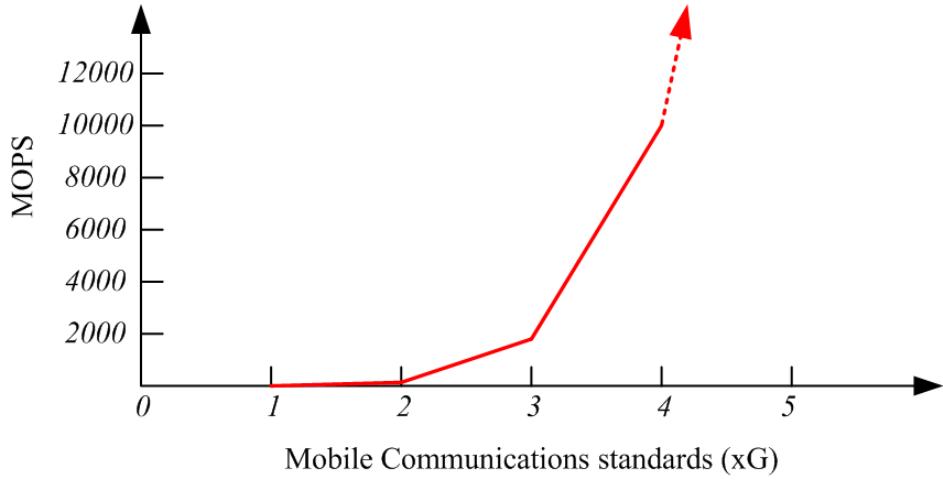


Figure 2: Evolución de la potencia de computo de diferentes generaciones de los sistemas móviles (xG)

energía no es la adecuada y el uso de los componentes baratos con las características inferiores. Sin embargo, la evolución de los terminales móviles también ha hecho posible la reducción de estos problemas con la disponibilidad de mayor potencia para el procesado de la señal. Con la evolución de las aplicaciones y las velocidades de transmisión más altas se ha requerido para cada nueva generación de los terminales móviles mayor potencia de cálculo. Este fenómeno se observa en la Figura 2 donde el incremento de las velocidades de transmisión ha sido seguido por el incremento exponencial de la complejidad algorítmica. Los requerimientos de la potencia de cálculo se presentan en Millones de Operaciones Por Segundo (MOPS). Este incremento en la potencia de cálculo permite el uso de algoritmos avanzados para reducir el acoplamiento que ha aparecido.

La cancelación de acoplamiento se ha analizado solo en la cadena de recepción de terminal, aunque el acoplamiento aparece tanto en la parte transmisora como receptora. Sin embargo, aunque se disponga de mayor capacidad de procesado de la señal en los terminales, es más recomendable hacer la mayor parte del procesado en las estaciones base, para minimizar el consumo de las baterías de los terminales. Por esta razón, en la tesis se ha analizado solo la minimización de acoplamiento en la cadena receptora del terminal, suponiendo que la cancelación de la cadena transmisora se consigue en la estación base.

Sección 3

En la tercera sección se describen las características principales de un frontal de RF y se describen las distorsiones que aparecen en los sistemas con una antena. A día de hoy, los frontales de RF siguen siendo sobre todo compuestos de los elementos analógicos para la recepción de la señal, la digitalización se hace en la etapa de IF (frecuencia intermedia) y el procesado de la señal es basado en la tecnología de procesado discreto de señales. Los sistemas más comunes tienen las arquitecturas sin la fase de IF (zero-IF) o con IF en las frecuencias bajas (low-IF). Con las mejoras de las características de los procesadores digitales, DSP, sobre todo, la conversión analógico-digital (A/D) se está acercando más a la antena. El resultado son los sistemas conocidos bajo el nombre de software defined radio (SWR, la radio definida por software). Sin embargo, este proceso de discretización tiene sus límites, ya que el sistema SWR no puede prescindir de los osciladores y amplificadores analógicos. Esto indica que algún tipo de distorsión no lineal, debido al comportamiento inherentemente no lineal de los sistemas analógicos, aparecerá en los sistemas futuros, por lo que los resultados obtenidos en esta tesis van a seguir siendo útiles. Las distorsiones analizadas en esta sección incluyen el offset de DC, el ruido de parpadeo (flicker noise), la intermodulación del segundo orden y el desequilibrio de las ramas I/Q (I/Q abalance). Todas estas distorsiones son no lineales por naturaleza. Sin embargo, los científicos que los han analizado han buscado las aproximaciones lineales lo suficientemente fiables para poder desarrollar los algoritmos para minimizar los efectos de las distorsiones. Es interesante observar que todas estas distorsiones pueden minimizarse tanto en el dominio analógico como digital siendo este último el más utilizado, por las velocidades de procesado de DSP ya comentadas.

Sección 4

En la cuarta sección se han analizado las distorsiones de un frontal de RF de un sistema con una sola antena y se ha visto que los autores están *linearizando* estas distorsiones, para simplificar su análisis y encontrar los modelos más sencillos para minimizar sus efectos en el sistema. Sin embargo, a veces es mejor utilizar el modelo no lineal, ya que se adapta mejor a la realidad y permite que los algoritmos para su minimización sean más precisos y consiguen mejores resultados. Los modelos no lineales se suelen evitar por la complejidad de la búsqueda de la función inversa que cancelaría la distorsión no lineal. Para modelar el acoplamiento entre múltiples antenas, todas las distorsiones del sistema con una antena analizadas en la

sección anterior se suman, además de las distorsiones adicionales, propias de los sistemas multiantena, que todavía no han sido estudiadas o analizadas a fondo. Por estas razones, el acoplamiento no se puede simplificar utilizando los modelos lineales sino que es necesario utilizar los modelos no lineales aunque sean más complejos. En la tesis se propone el modelo polinomial de acoplamiento no lineal, limitando el grado del polinomio a la tercera potencia. El modelo simplificado de acoplamiento se muestra en la Figura 3 donde la señal recibida por 2 antenas es denominada x_1 y x_2 y la señal con acoplamiento es y_1 e y_2 . Las funciones g_1 y g_2 describen la rama directa de propagación de la señal de cada antena, mientras que las funciones f_1 y f_2 modelan el acoplamiento.

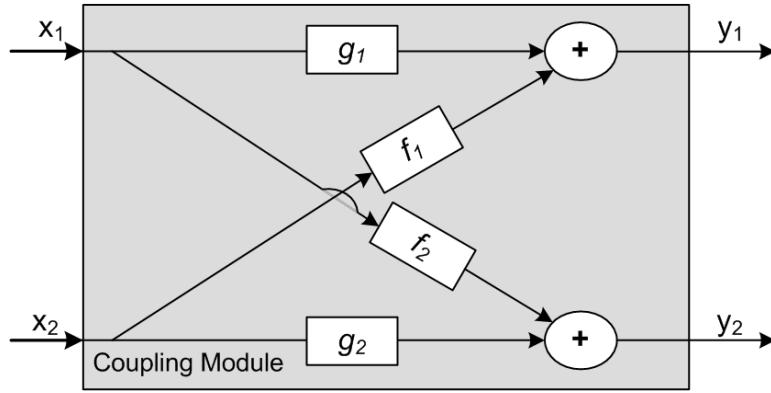


Figure 3: El modelo simplificado de acoplamiento

Matematicamente, estas funciones se expresan como:

$$\begin{aligned} y_1 &= g_1(x_1, Dist_1) + f_1(x_2, Coup_{21}) \\ y_2 &= g_2(x_2, Dist_2) + f_2(x_1, Coup_{12}) \end{aligned} \quad (1)$$

donde las funciones f_i se definen en esta tesis como

$$\begin{aligned} f_1(x_2, Coup_{21}) &= c_{11}x_2 + c_{12}x_2^2 + c_{13}x_2^3 \\ f_2(x_1, Coup_{12}) &= c_{21}x_1 + c_{22}x_1^2 + c_{23}x_1^3 \end{aligned} \quad (2)$$

Los coeficientes de acoplamiento c_{ij} se expresan en decibelios a lo largo de la tesis, denominándolos C_{ij} , y se definen como el logaritmo del resultado de la división de la potencia de la

señal del interés y la potencia de la señal interferente:

$$C_{ij} = 10 \log \frac{P_j}{P_i} [\text{dB}] \quad (3)$$

donde P_i corresponde a la potencia de la señal recibida por la antena i , mientras que P_j es la potencia de la interferencia del orden j -ésimo.

Con este modelo el acoplamiento se puede presentar como una *caja negra* con sus valores de entrada y los correspondientes de salida. Este modelo no requiere la descripción y análisis de todas las distorsiones que resultan en el acoplamiento, y simplifica el modelo no lineal. Este enfoque también permite que los algoritmos desarrollados en esta tesis se puedan aplicar a otros sistemas no lineales que se pueden aproximar con el modelo propuesto.

En esta sección, después de la descripción detallada del modelo de acoplamiento propuesto por el autor, se aplica este modelo al sistema que se analizará a continuación de la tesis: el del receptor con la conversión de bajada I/Q. Este sistema tiene sus peculiaridades, ya que por el filtro paso-bajo el efecto de acoplamiento cuadrático no se percibe y no influye en las prestaciones del sistema.

Sección 5

En la quinta sección se ha llevado a cabo el análisis matemático de la función inversa de acoplamiento. La solución analítica es muy compleja y no tiene una forma única ya que cambia en función de características de acoplamiento y de conjunto de señales transmitidas. También es posible que la solución inversa ni siquiera exista o que la solución proporcionada no sea aceptable. Por estas razones, se ha procedido a la comprobación de la existencia de una solución única y a la posterior búsqueda de la solución con el método de Neptuno. Este procesado requiere elevada potencia de procesado por lo que se introducen los algoritmos para la búsqueda de la función inversa aproximada. Estos algoritmos tienen al principio una fase de entrenamiento (calibración). En la fase de calibración se define la superficie de acoplamiento, lo que representa la función que traslada los símbolos transmitidos a los recibidos. El inconveniente de estos algoritmos, denominados heurísticos, es que no ofrecen ningún tipo de garantías de que van a proporcionar la solución óptima. Por el otro lado, su ventaja está en que no son ambiguos y una vez que se defina bien la superficie inversa, solo

se tiene que ejecutar una pequeña parte del algoritmo para obtener la solución. Utilizando los algoritmos heurísticos se evita la búsqueda analítica de la solución y se consigue una aproximación de la superficie.

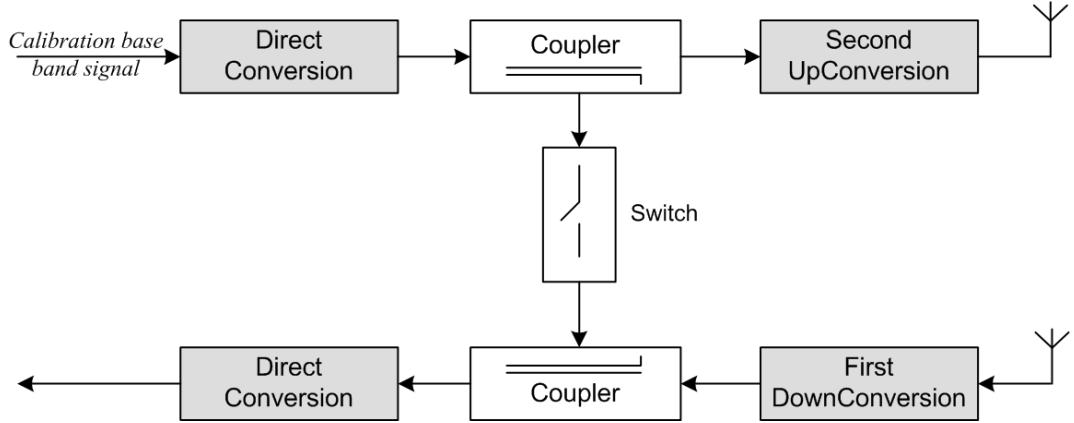


Figure 4: Diagrama del bucle de calibración dentro del terminal

Desde el punto de vista tecnológico la calibración se puede realizar de tres maneras:

- con un bucle en el terminal móvil, pasando por su cadena de transmisión y recepción, sin llegar a transmitirse por las antenas. esto se muestra en la Figura 4.
- con un bucle entre la estación base y el terminal móvil
- en la sala reverberante, a nivel industrial.

Utilizando los algoritmos heurísticos y con la introducción de la superficie de transmisión y de su superficie inversa, la mayor complejidad computacional se encuentra en la búsqueda de la superficie inversa aproximada. Esta búsqueda se lleva a cabo en el modulo de desacoplamiento que se encuentra a la salida del conversor A/D del receptor, como se muestra en la Figura 5. Se ha elegido esta posición porque permite utilizar el procesado digital de señal y es la menos dependiente del tipo de señal que transmite, lo que permite su aplicación en diferentes sistemas de comunicación. El modelo para la superficie inversa de acoplamiento que se ha utilizado es el modelo polinomial del noveno grado. El noveno grado se ha elegido porque el análisis analítico se paró cuando apareció el sistema de dos ecuaciones con noveno grado de dos incógnitas. Además, la superficie construida de esta forma tiene varios mínimos y máximos y se puede adaptar adecuadamente a diferentes tipos de acoplamiento.

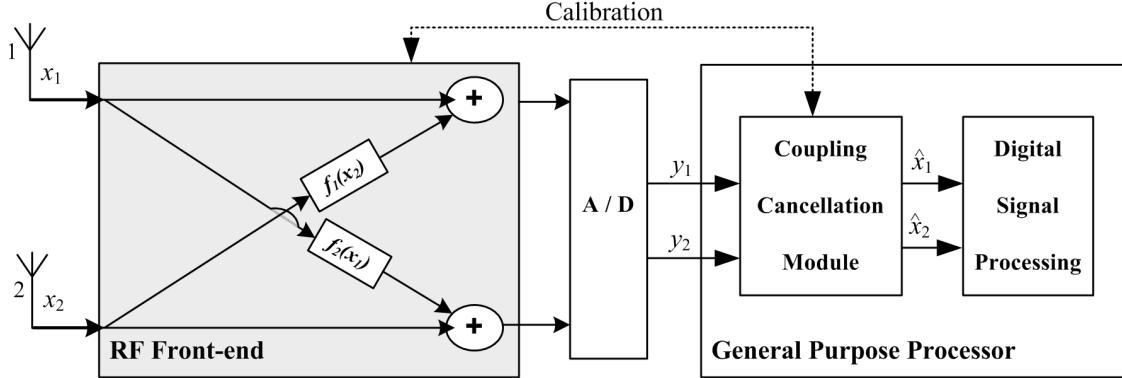


Figure 5: El esquema de posicionamiento del modulo de desacoplamiento en la cadena receptora

En esta sección se introduce la necesidad de una medida del parecido de la superficie inversa aproximada a la real y que caracteriza el comportamiento del algoritmo de desacoplamiento. Es muy importante que estas funciones, que se minimizan con los algoritmos propuestos, reflejen lo mejor posible las prestaciones del sistema con desacoplamiento en términos de BER. Se analizan varias medidas entre los que se han escogido las siguientes dos, que se han analizado en detalle con el algoritmo de SA:

$$Q_2 = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_{1i} - \hat{x}_{1i})^2} \quad Q_3 = \sum_{i=1}^N |x_{1i} - \hat{x}_{1i}| \quad (4)$$

donde N representa al número de símbolos de calibración, mientras que x_{1i} y \hat{x}_{1i} son los valores verdaderos y estimados del i^{simo} símbolo de calibración. La primera medida es la del error cuadrático medio, y la segunda es la suma de errores absolutos. La primera medida es más común para la caracterización, pero la segunda es menos costosa computacionalmente.

Para simular y analizar las prestaciones del modulo de desacoplamiento, se ha utilizado el modelo de sistema de comunicaciones que se muestra en la Figura 6. El sistema consiste de 2 antenas en el transmisor y 2 en el receptor con el procesado de conversión de bajada I/Q. Diferentes constelaciones se han utilizado para ver el efecto de acoplamiento en las constelaciones muy densas como 64QAM o poco densas, como QPSK.

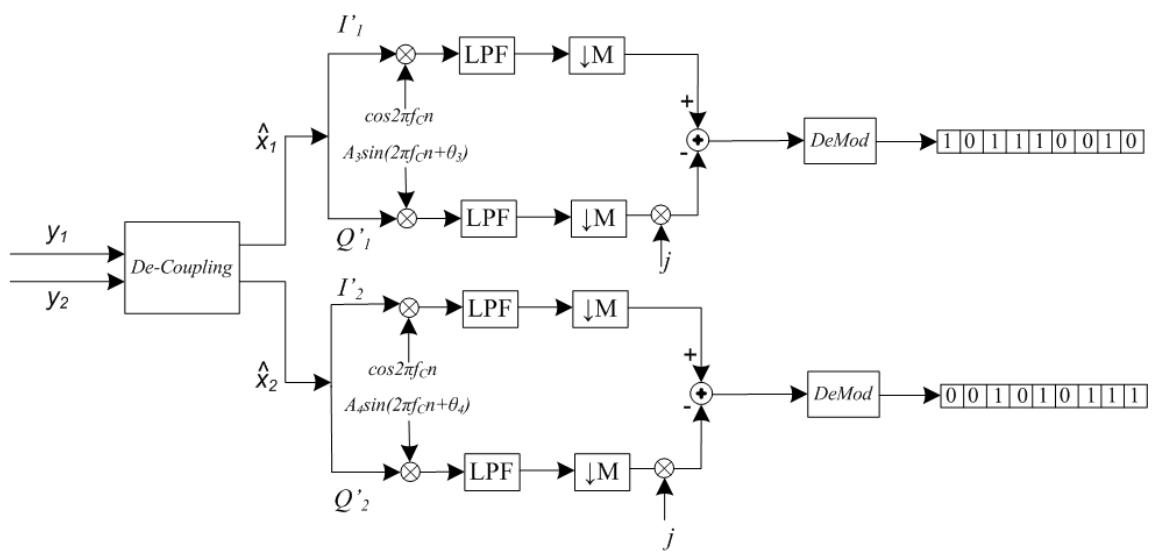
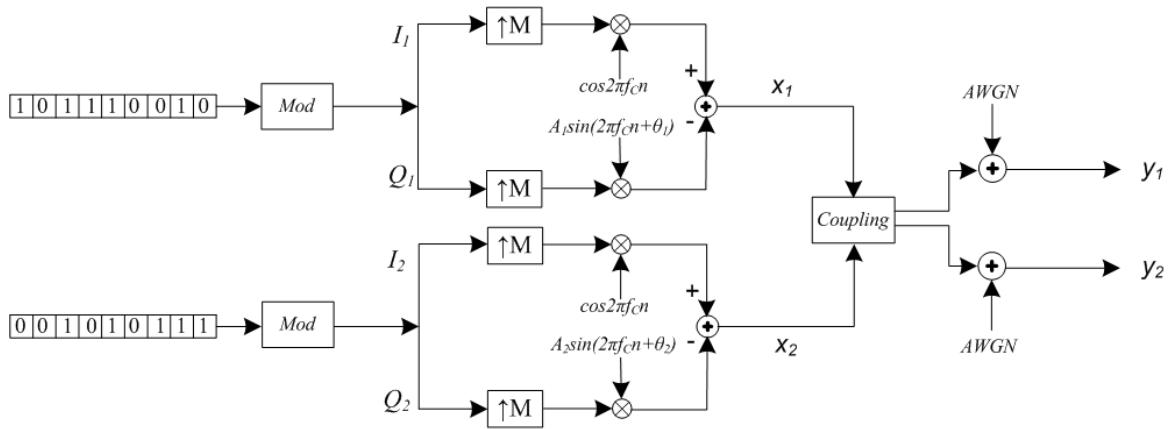


Figure 6: La estructura del sistema simulado en MatLab

Sección 6

En la sexta sección se desarrollan y analizan en profundidad tres algoritmos para la búsqueda de la superficie inversa y se les aplica en un sistema de comunicación con diferentes valores de acoplamiento. Los tres algoritmos son el Recocido Simulado (Simulated Annealing, SA), el Recocido Simulado Rápido Mejorado (Improved Fast Simulated Annealing, IFSA) y la Búsqueda Multi-Nivel Guiada (Guided Multi-Level, GML). El algoritmo de SA se ha analizado porque es el que más frecuente se utiliza en las aplicaciones en las que se tienen muchos grados de libertad y gran número de posibles soluciones. Este algoritmo, como su nombre indica, está inspirado en un proceso físico de la metalurgia, el recocido (annealing) de metal. La técnica imita la naturaleza de metal que se calienta y luego enfria lentamente para llegar al estado de mínima energía. El algoritmo SA se utiliza frecuentemente porque puede garantizar que se encontrará la solución óptima, pero no se puede estimar el numero de iteraciones necesarias para encontrarlo. Este algoritmo es conocido por su lenta convergencia, por lo que el número de iteraciones suele ser alto, por lo que la búsqueda puede llegar a ser inaceptablemente exhaustiva. La solución a este problema computacional de SA es encontrar un punto medio entre la precisión garantizada y la velocidad. Por esta razón el algoritmo que se ha utilizado en la tesis es una versión mejorada (mejor adaptada al problema de acoplamiento) cuya velocidad de convergencia está mejorada. Sin embargo, se ha dado gran prioridad a la precisión, manteniéndola casi inalterada respecto al algoritmo original, por lo que los grados de libertad para mejorar la velocidad de convergencia del algoritmo no han sido muchos, y el tiempo de procesado está muy por encima del necesario para funcionar en tiempo real. Dándole prioridad a la precisión frente a la velocidad se ha querido alcanzar una cota de las prestaciones de los algoritmos de desacoplamiento. Este algoritmo se ha aprovechado para encontrar el tipo de la función de coste más adecuado, el punto de partida de los coeficientes y para comprobar la selección del modelo de la superficie con sus prestaciones en término de la probabilidad de error.

A la hora de aplicar SA a un problema específico, se tiene que obtener un juego de parámetros que se indican a continuación:

- Para el espacio del estado o la resolución de los coeficientes
 - la precisión del símbolo (Res_{symb})
 - la precisión de los coeficientes (Res_{coef})

- Los valores iniciales de los coeficientes o el punto de arranque ($Coeff_0$)
- El método de selección del vecino ($New_Candidate$)
 - El máximo desplazamiento del coeficiente (R_0)
 - El número de candidatos visitado por cada nivel de temperatura ($NVCPL$)
- La forma de la función de coste ($Cost_Func$)
- La función de probabilidad de transición ($Trans_Prob$)
 - la probabilidad inicial de aceptación del peor movimiento (p_{max})
- La función de programación de recocido ($Ann_Schedule$)
 - el factor de reducción de tiempo (α)
- El número máximo permitido de iteraciones (I_{MAX})
 - la mínima probabilidad del peor movimiento durante la búsqueda (p_{min})
 - iteraciones correspondientes a la búsqueda propiamente dicha de SA (I_{SA})
 - el número de las iteraciones codiciosas cuesta abajo (I_{DH})

En esta sección se han descrito en detalle y posteriormente optimizado todos estos parámetros de SA para adaptar el algoritmo lo máximo posible a las necesidades del problema de desacoplamiento. A continuación se explicarán dos de estos parámetros, α y $NVCPL$ y R_0 , a modo de ejemplo, y se mostrarán algunos resultados de las simulaciones para ajustarlos al problema de desacoplamiento.

Por analogía al proceso físico de recocido, la función que se optimiza, la función de coste, se corresponde a la energía del fluido, mientras que las variables del diseño se corresponden a las posiciones de las moléculas. Lo que se intenta conseguir es traer al sistema, desde un punto arbitrario de partida, al punto que se corresponde a la mínima energía (función de coste en este caso). Esto se consigue con las perturbaciones aleatorias de los coeficientes de diseño. Antes de que la nueva perturbación se acepte, el algoritmo calcula el cambio en la función de coste en los nuevos estados candidatos. Si esta función disminuye, la perturbación se acepta. Incluso si ocurre al contrario, y el nuevo nivel de energía (función de coste) es mayor, es posible que

el nuevo estado se acepte, dependiendo del valor de la función de probabilidad, que depende del nivel de temperatura actual, $p = f(T)$. El parámetro T , la temperatura disminuye con el progreso de las iteraciones dependiendo del parámetro α :

$$T_{k+1} = \alpha T_k \quad (5)$$

La función de probabilidad se elige para que se cumplan las siguientes dos requerimientos:

- Cuando T es grande debe ser poco probable que el algoritmo se atasque en un mínimo local, por lo que los movimientos hacia arriba (uphill, hacia el valor de función de coste más alto) se aceptan con mayor probabilidad
- Mientras la temperatura se vaya acercando al cero (enfriamiento), no se deben permitir los cambios bruscos en el valor de la función de coste y se aceptan sobre todo los que llevan a los estados de energía más bajos

El número de nuevos puntos, que se denominan candidatos y cuya función de coste se calcula en cada nivel de temperatura, se define con el parámetro N_{VCPL} . Este valor es directamente relacionado con el problema específico para el que se utiliza el algoritmo. Si este valor es grande se comprueban muchos candidatos por cada nivel y la probabilidad de encontrar la solución óptima es alta. Sin embargo, un número elevado de búsquedas hace que el algoritmo sea lento, ya que se pierde mucho tiempo en las búsquedas innecesarias.

Para encontrar los valores óptimos, es decir, los que reunen tanto las propiedades de robustez, precisión y baja carga computacional, se ha simulado el sistema con varios niveles de acoplamiento y se han sacado los valores de la función de coste en función de factor de reducción de temperatura (α) y el número de candidatos visitados en un nivel (N_{VCPL}). En la Figura 7 se muestra el resultado de simulación para el acoplamiento de $[-10, -12, -12]$ dB con el valor máximo de desplazamiento permitido fijado a $R_0 = 0.05$. las simulaciones muestran que para el valor de $\alpha = 0.3$, el parámetro de N_{VCPL} no influye en las prestaciones, y no se pueden mejorar respecto al punto de partida ($Q = 0.069$ en este caso) aunque se elija el número de candidatos visitados muy alto. Esto ocurre porque el paso de disminución de temperatura es muy alto y la temperatura baja demasiado rápido. Sin embargo, cuando α se va acercando a la unidad, el mínimo de la función de coste disminuye para todos los valores de N_{VCPL} excepto para $N_{VCPL} = 2$ en cuyo caso no se tienen suficientes puntos en la búsqueda.

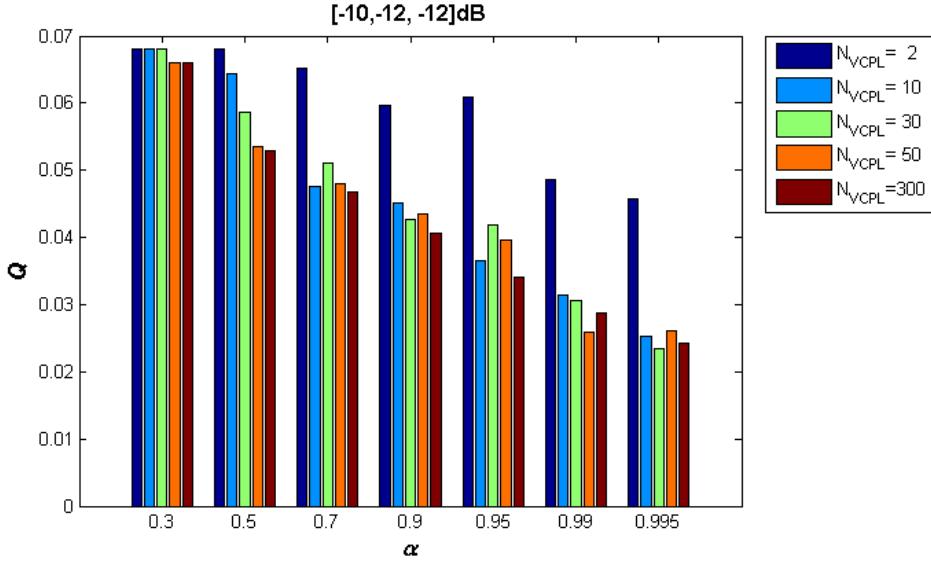


Figure 7: Q en función de α y N_{VCPL} para el acoplamiento de $[-10, -12, -12]\text{dB}$

El número máximo de candidatos (puntos) por cada nivel de temperatura es N_{VCPL} , y el máximo desplazamiento de cada punto es R_0 , por lo que el área máxima por cada nivel está restringida al producto de estos dos valores, $R_0 \times N_{VCPL}$. Para encontrar al R_0 óptimo, se han hecho simulaciones del sistema con diferentes niveles de acoplamiento y se han obtenido las graficas de la función de coste Q en función de R_0 . Una de estas gráficas se muestra en la Figura 8. Se puede observar la forma de U de esta función, y la existencia de un mínimo claro

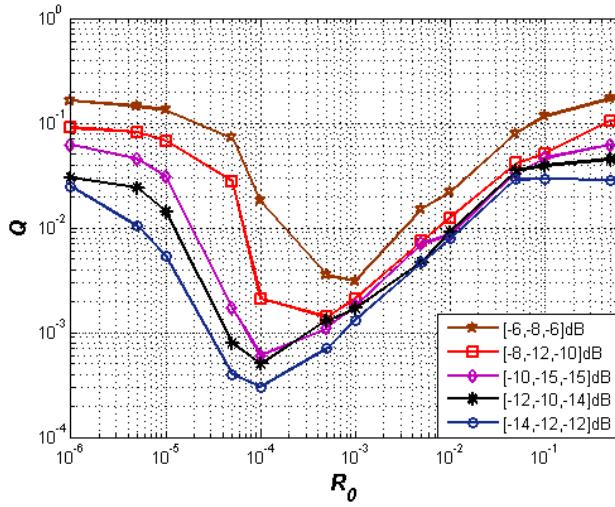


Figure 8: Q en función de R_0 en diferentes condiciones de acoplamiento

de la función de coste. Esto es debido a que los valores pequeños de R_0 limitan demasiado el área de búsqueda y el mínimo no se puede encontrar, mientras que los valores grandes de R_0 hacen que la precisión de búsqueda sea muy mala.

También se han analizado las dos funciones de coste propuestas en la sección anterior. Se han hecho simulaciones con diferentes niveles de acoplamiento, y se ha requerido que el comportamiento de la función de coste esté de acuerdo con el comportamiento de la BER: si disminuye la función de coste, esto se debería reflejar en la BER. Sin embargo, se ha observado que este comportamiento no se cumple en todos los casos cuando la función de coste es el error absoluto, tal como se muestra en la Figura 9, por lo que se ha elegido como función de coste el de error cuadrático medio cuyo comportamiento ofrecía mayor robustez.

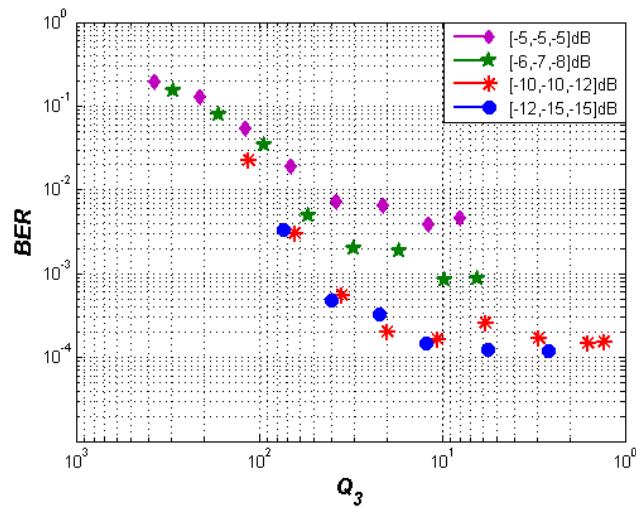


Figure 9: : BER en función de Q_3 con diferentes niveles de acoplamiento

El segundo algoritmo de desacoplamiento, denominado el SA Rápido Mejorado (Improved Fast SA) se ha introducido y desarrollado en la tesis para acelerar la búsqueda. Se ha desarrollado por el autor específicamente para el problema de acoplamiento no lineal y se pretende que sea útil en las aplicaciones reales. Dentro de IFSA, se mantienen los métodos heurísticos, y la mejora en la velocidad se obtiene a costa de la precisión de la superficie aproximada. Aunque se ha conseguido la mejora de alrededor de 10 en la velocidad respecto al algoritmo de SA, las aplicaciones de este algoritmo siguen siendo inviables para utilizaciones en tiempo real. Otro inconveniente de IFSA es que el final del algoritmo está determinado con el máximo permitido tiempo de procesado. Por esta razón, la precisión de la solución no se puede garantizar, y a

veces hay que volver a ejecutar el algoritmo para mejorar la precisión. Teniendo en cuenta que el camino de búsqueda de la superficie inversa aproximada se determina de modo estadístico, nueva ejecución del algoritmo proporcionará nueva solución.

Finalmente, para evitar la influencia de factor probabilístico en la búsqueda de la solución y la indecisión que este conlleva, se ha desarrollado el algoritmo determinista iterativo denominado la Búsqueda Multi-Nivel Guiada (Guided Multi-Level Search, GML search). Con este algoritmo se reduce la carga computacional media requerida para el cálculo de la función de coste, se simplifica el método de búsqueda, y, como consecuencia, se acelera el proceso de búsqueda de superficie inversa aproximada. Aunque este algoritmo se haya desarrollado específicamente para el problema del acoplamiento no lineal, se puede utilizar en otros problemas con muchos grados de libertad. La ventaja del algoritmo propuesto es la posibilidad de tener el control de modo sencillo de la velocidad y precisión, la reducción de los requerimientos de procesado, la implementación sencilla y la eliminación del factor probabilístico de la búsqueda. La idea básica es la de construir una superficie de desacoplamiento con una base sólida, empezando con dos niveles, e introduciendo gradualmente más niveles de libertad. Este método asegura la solución determinista al problema de desacoplamiento con una precisión excelente, buena velocidad de convergencia y sin requerimientos de procesado elevado.

En esta sección los tres algoritmos de han analizado y aplicado para la cancelación de acoplamiento no lineal. El análisis se ha llevado a cabo en el escenario sin ruido aditivo gaussiano durante el proceso de calibración y con un número elevado de símbolos de calibración para evitar la influencia de estos parámetros en los resultados de los análisis. Este escenario se correspondería al calibración a nivel industrial, en la sala reverberante. Los resultados confirman que el algoritmo de SA es el más preciso, como era de esperar, bajo todas las condiciones de acoplamiento. El SA va seguido por GML cuya precisión está un poco inferior, mientras que la precisión de IFSA es muy inferior a la de estos dos algoritmos. En cuanto a la velocidad de procesado, medido en la velocidad de convergencia, los algoritmos de IFSA y GML requieren un número de iteraciones significativamente inferior al de SA. Sin embargo, la complejidad computacional de GML está considerablemente por debajo de la de IFSA. Más concretamente, GML requiere 10% de los requerimientos de procesado de IFSA, mientras que IFSA, a su vez, necesita 10% de procesado del algoritmo SA original. Los tres algoritmos ofrecen de modo sistemático las soluciones adecuadas en este escenario, por lo que se pueden adoptar para el problema de desacoplamiento no lineal.

Sección 7

En la séptima sección se ha analizado más a fondo el caso de desacoplamiento cuando no hay ruido en el proceso de calibración, y se han optimizado los algoritmos propuestos. La primera simplificación consiste en la reducción del número de potencias en los polinomios de la superficie inversa aproximada. Esta simplificación reduce la carga computacional, pero degrada las prestaciones de los algoritmos; el nivel de degradación depende de los parámetros de acoplamiento. Un compromiso entre las prestaciones de sistema simplificado y la reducción de la carga computacional sería limitar las potencias hasta la quinta, con lo que el número de grados de libertad sería 12.

En el caso del receptor basado en la conversión I/Q de bajada este número de coeficientes se puede reducir aún más casi sin perdida de prestaciones. Esto es válido para los tres algoritmos estudiados. La reducción es posible gracias a la característica de este receptor de que las potencias pares se rechacen en los dos filtros paso-bajo del conversor de bajada I/Q. Por lo tanto, el acoplamiento correspondiente al cuadrado no va a afectar a la señal recibida y los requerimientos para el algoritmo de desacoplamiento se pueden relajar.

Como ejemplo, en la Figura 10 se muestra el resultado de una de las simulaciones del sistema con el acoplamiento de $[-9, -8, -6]$ dB y con el algoritmo de SA . Se puede observar que la degradación por el paso de (9,9) a (7,7) potencias es pequeño en términos de BER. Este

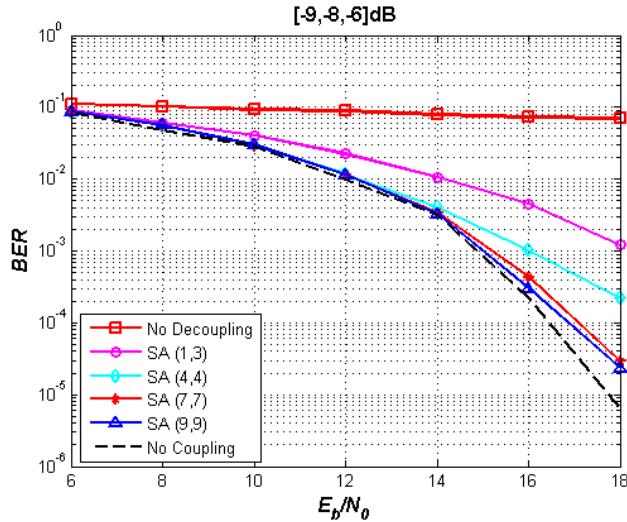


Figure 10: BER con diferente número de coeficientes de la superficie inversa aproximada

paso supone que en lugar de 20 coeficientes (9 coeficientes correspondientes a las potencias y la constante por cada antena) en el sistema (9, 9) se necesitan 16 para el sistema (7, 7). Sin embargo, en el caso de acoplamiento de $[-9, -8, -6]$ dB, que se corresponde a acoplamiento fuerte del factor cúbico, ni siquiera con 20 grados de libertad se pueden conseguir la eliminación completa de acoplamiento y las prestaciones del sistema iguales a las de un sistema 64QAM sin acoplamiento (la curva negra en la Figura 10, denominada “No Coupling”).

En el caso de forzar que los coeficientes de las potencias pares sean iguales a 0, ya que el receptor con el conversor I/Q de bajada no les deja pasar, se han hecho varios estudios comparando las prestaciones de este sistema con el que utiliza todos los coeficientes para encontrar la superficie de acoplamiento inversa. En el primer caso, se tendrá un total de $D = 11$ coeficientes, y en el segundo $D = 20$. Los resultados para el acoplamiento de $[-8, -7, -8]$ dB se muestran en la Figura 11. Se observa que las prestaciones son muy parecidas, como ya se ha comentado, pero, al tener menos coeficientes en el caso de $D = 11$, la carga computacional de cálculo es menor y es más recomendable usar este algoritmo.

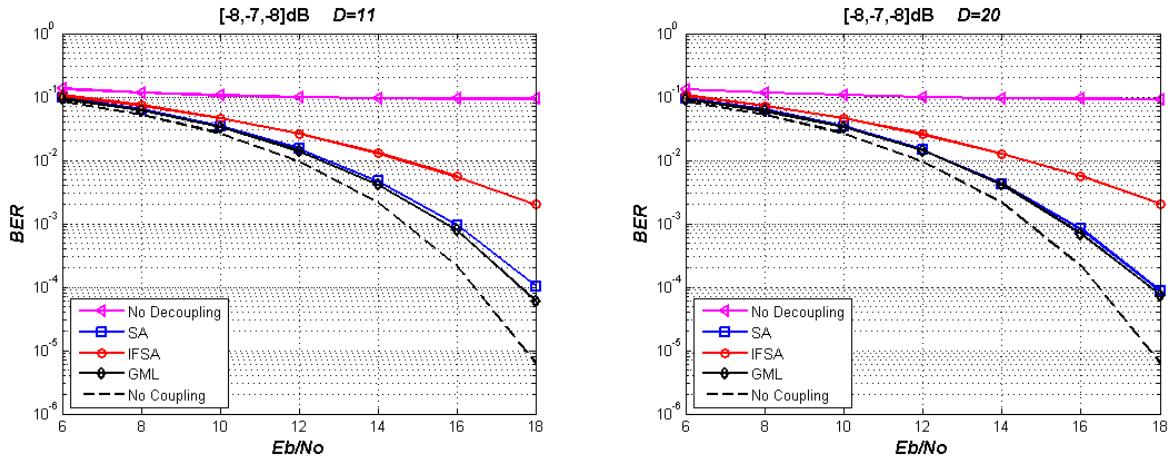


Figure 11: Prestaciones de SA, IFSA y GML con acoplamiento de $[-8, -7, -8]$ dB y diferentes grados de libertad

En esta sección también se ha estudiado el número de puntos de calibración cuando el número de coeficientes que definen a la superficie inversa se reduce, según lo resumido en el párrafo anterior. La robustez es la característica a la que se ha prestado más atención a la hora de determinar el tamaño de la secuencia de calibración. Según las prestaciones de los tres algoritmos estudiados, se ha elegido el tamaño de la trama de calibración de 512 muestras. Se muestran a continuación, en la Figura 12, las prestaciones del algoritmo de GML con diferente

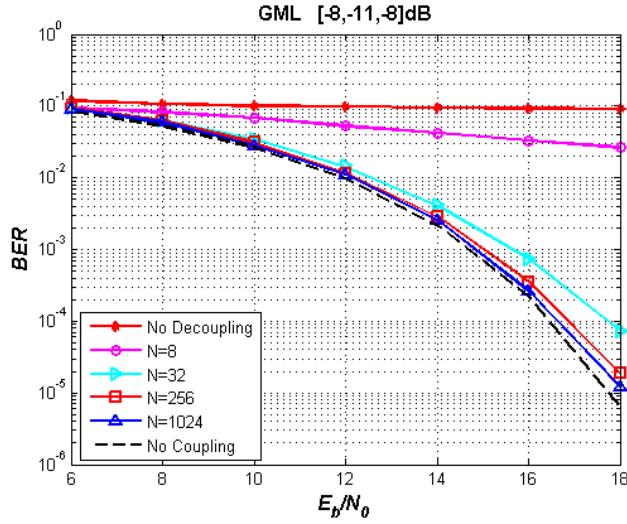


Figure 12: Prestaciones de GML con diferentes tamaños de la trama de calibración

número de símbolos en la trama de calibración. Se puede observar que con 1024 símbolos las prestaciones se acercan mucho a las de un sistema sin acoplamiento, mientras que para el tamaño de 32 símbolos se tiene una pérdida de algo más de 1dB para la BER de 10^{-4} . Teniendo en cuenta las prestaciones de otros algoritmos, SA e IFSA, se ha elegido el tamaño de 512 símbolos como el más adecuado para los tres algoritmos cuando no haya ruido en proceso de calibración.

Sección 8

En la última sección se analizan las prestaciones de los algoritmos cuando existe el ruido durante el proceso de calibración. Este escenario se corresponde al caso en el que la calibración está implementada dentro del terminal móvil o entre la estación base y el terminal móvil. A diferencia del caso anterior, sin ruido, la superficie de acoplamiento ya no se construye en las condiciones ideales, sino que, por los efectos del ruido, puede haber errores en la estimación de esta superficie puesto que la estimación tanto de superficie directa como inversa se lleva a cabo en las mismas condiciones de ruido. También se ha analizado el tamaño de la secuencia de calibración cuando el algoritmo de desacoplamiento tiene el número reducido de 11 coeficientes. Como era de esperar se ha llegado a la conclusión que en este caso se necesitan más muestras que en el caso sin ruido en la fase de calibración, como era de esperar. El valor elegido de número de muestras para la calibración es de 1024.

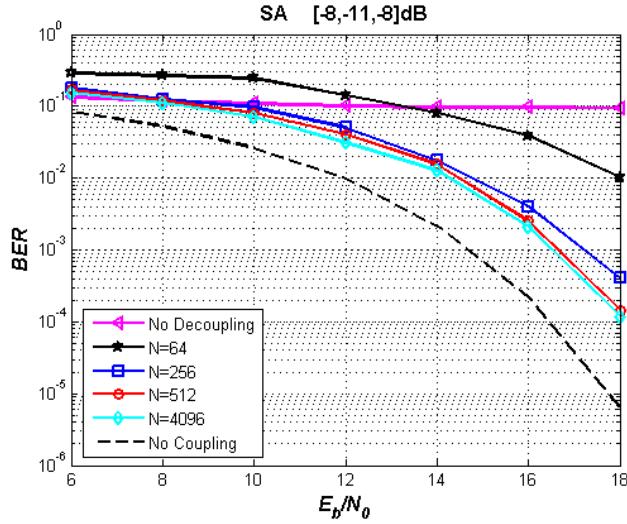


Figure 13: Prestaciones de SA con ruido en la fase de calibración

En la Figura 13 se muestran las prestaciones de SA en este escenario, con el acoplamiento de $[-8, -11, -8]$ dB. Se puede observar que se obtiene una pérdida de algo menos de 2dB incluso con 4096 símbolos de calibración. Sin embargo, al pasar de 512 a 4096 se gana muy poco, por lo que se puede concluir que esta pérdida de prestaciones no se puede disminuir incrementando el tamaño de la trama de calibración.

En este escenario se ha observado que las mejores prestaciones no son las de SA para todos los valores de relación señal a ruido. El algoritmo de IFSA se ha mostrado como el único capaz de producir las mejoras de prestaciones en el caso de acoplamiento fuerte, cuando E_b/N_0 está por debajo de 10dB. Las prestaciones con el algoritmo GML están muy cerca de las de algoritmo de SA, pero sus requerimientos computacionales son mucho más reducidos, por lo que se ha llegado a la conclusión que en este caso sería necesaria una combinación de algoritmos de IFSA y GML para obtener las mejores prestaciones.

Las posibles líneas de investigación futuras

Las contribuciones principales de esta tesis están en el análisis y modelado del acoplamiento en RF y en el desarrollo de un método sencillo, denominado GML para la búsqueda de la superficie de acoplamiento inversa. Por lo tanto, se pueden distinguir claramente dos posibles líneas de investigación:

1. El desarrollo más detallado de GML y el análisis de su eficiencia en otros problemas de acoplamiento, como, por ejemplo en el acoplamiento de sistema audio 3.1 o en las redes de vdd o gnd. Sería interesante ver su comportamiento en el caso de acoplamiento complejo con muchos grados de libertad, pero con la potencia de procesado limitada.
2. El análisis más profundo del acoplamiento de RF. El modelo desarrollado en esta tesis asume que el acoplamiento es invariante en el tiempo e independiente de frecuencia. Estas asunciones se han tomado para simplificar el modelado del sistema, pero una posible investigación futura podría analizar la influencia de su variabilidad y dependencia en el sistema. Sin embargo, un modelado más detallado sería necesario por lo que habrían que realizar medidas exhaustivas para caracterizar el acoplamiento.

Sin embargo, el futuro más práctico y deseable del trabajo realizado en esta tesis sería llevar estos algoritmos a la vida real y comprobar los modelos y las prestaciones de este sistema fuera del entorno simulado.

Las contribuciones científicas

Algunos de los análisis llevados a cabo sobre el frontal de RF se han publicado en las siguientes contribuciones: [Ara07b, Ara07a, Ara07c, Rao06b, Ara02a, Ara02b]. Mientras las primeras tres contribuciones están directamente relacionadas con los problemas del acoplamiento no lineal, las últimas tres son más generales y en ellas se analizan diferentes problemas de los sistemas de comunicaciones. Estas publicaciones han servido para obtener la experiencia y el punto de vista más general de los problemas con los que se encuentran en el diseño de un sistema de comunicaciones. Además de estas publicaciones, el autor ha participado en un número elevado de proyectos de investigación que le han ayudado en gran medida en el trabajo realizado en esta tesis. Algunos de estos proyectos se indican a continuación:

- MATRICE (IST-2001-32620) Multicarrier CDMA Transmission Techniques for Integrated Broadband Cellular Systems; financiado por el quinto programa Marco de la Comisión Europea; duración: desde Enero de 2002 hasta Enero de 2005
- 4MORE (IST-2002-07039) MC-CDMA Multiple Antenna System-on-Chip for Radio Enhancements; financiado por el sexto programa Marco de la Comisión Europea; duración: desde Enero de 2004 hasta Junio de 2006

- Wireless Hybrid Enhanced Mobile Radio Estimators (WHERE); financiado por el séptimo programa Marco de la Comisión Europea; duración: desde Enero de 2008 hasta Junio de 2010
- Arquitecturas para radiocomunicaciones digitales de banda ancha; financiado por el Ministerio de Ciencia y Tecnología; duración: desde 2001 hasta 2003
- Transceptores integrados multielemento y monocircuito para radiocomunicación digital de banda ancha; financiado por el Ministerio de Ciencia y Tecnología; duración: desde Enero de 2004 hasta Diciembre de 2006
- Transceptores integrados multielemento y monocircuito para radiocomunicación multimedia de banda ancha y muy ancha; financiado por el Ministerio de Educación y Cultura; duración desde Enero de 2007 hasta Diciembre de 2009

2 Motivation and background

At the beginning of 21st century, mobile phones are everywhere. In fact, they became so common, that several European countries have more mobile phones than residents. Nowadays, around 80% of the world's population has mobile phone coverage and this figure is expected to increase to 90% by the year 2010. The exact number of world wide mobile phones is unknown. However, according to Paul Budde, a telecoms analyst considered to provide the official subscriber counts used by various industry operators and manufacturers, the total number of mobile phone subscribers reached 3 billion in September 2007 [Bud07]. This number will continue to rise as telecommunication industry growth shows no signs of saturation surpassing 15% increase year after year. This is achieved by constant improvement of mobile phone systems and terminals, and by constant inclusion of new applications and services. In this chapter, the evolution of mobile terminal and technologies that led to present mobile systems is reviewed. It explains the research motivation that resulted in this thesis, and shows the necessity of this research work.

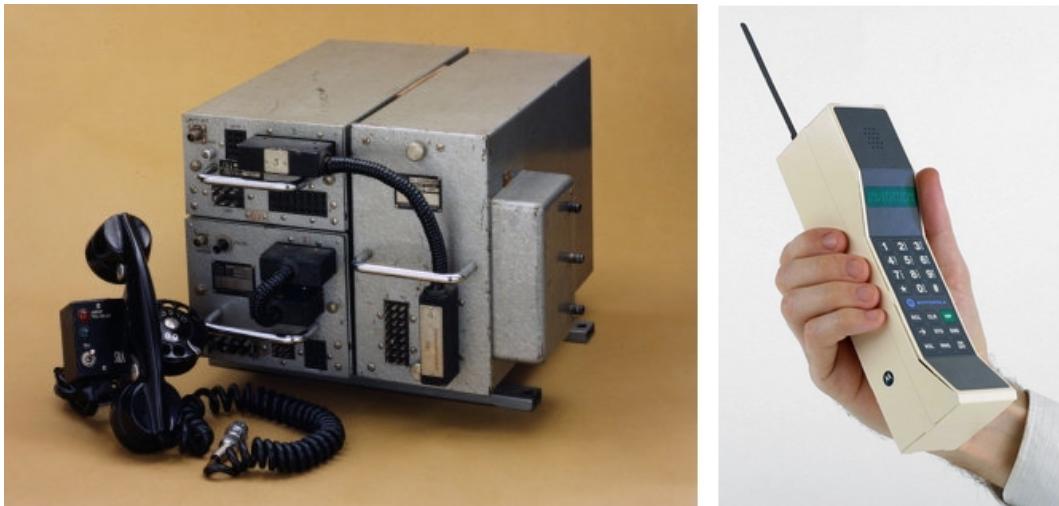


Figure 14: Ericsson's *Mobile Telephone system A* (MTBA) from 1965 and DynaTac 800, the first cellular hand-held mobile phone from 1983

In order to fully appreciate the amazing evolution of mobile communications technology, we should go back to the beginning of 20th century. The history of mobile phones in Europe starts in 1926 on the first-class passenger trains between Berlin and Hamburg that include radio telephony used by trained specialists. Some 25 years later, ships on the Rhine were

among the first to use radio telephony, but this time with an untrained end-customer as a user. The first fully automatic mobile phone system, called MTA (Mobile Telephone system A), was developed by Ericsson and was commercially released in Sweden in 1956. MTA weighted 40 kg (Figure 14), while MTB, its upgraded version based on transistors and released in 1965 , weighted "only" 9 kg. Afterwards came so-called transportable phones or "bag phones" which were equipped with a cigarette lighter plug so that they could be carried around, and were used as true portable phones [Aga05].

The year 1983 marked the end of one era as MTB system with 600 customers was shut down, and the beginning of another era as first authentic handheld mobile phone *DynaTAC 8000X* was introduced (Figure14). The company behind *DynaTAC 8000X* was Motorola which had invested fifteen years of research and 100 million dollars in the advancement of cellular technology. The store price was set to \$3,995, it weighed almost a kilogram and was 33 x 4.4 x 9 centimeters in dimension. *DynaTAC* was known as "the Brick", it supported one hour of talk time and eight hours of standby time, it featured a LED display and marked the beginning of so called first generation (1G) mobile phone systems. Surprisingly, these systems, based on analog technology, are still widely available today (Mexico, Venezuela, Dominican Republic...), although its use has dropped considerably with the introduction of various digital standards.

The second generation (2G) of mobile phone systems such as GSM, IS-136 ("TDMA"), iDEN and IS-95 ("CDMA") was introduced in 90s. This generation is characterized by digital circuit switched transmission and the introduction of advanced and fast phone-to-network signaling. Coinciding with the introduction of 2G systems was a trend toward tiny 100-200g hand-held devices, which soon became the norm. This change was possible through technological improvements such as more advanced batteries and increased energy-efficient electronics [Nol86, Lev04].

Not long after the introduction of 2G networks, projects began to develop the so called 2.5G systems setting the sight on the third generation (3G) systems. Inevitably there were many different standards with different contenders pushing their own technologies. Quite differently from 2G systems, however, the meaning of 3G has been standardized in the IMT-2000 standardization process created by the International Telecommunication Union (ITU). However, this process did not standardize on a technology, but rather on a set of requirements (2 Mbit/s maximum data rate indoors, 384 kbit/s outdoors, for example). At that point, once again, the vision of a single unified worldwide standard broke down and five different standards have

been introduced (W-CDMA, CDMA2000, TD-CDMA/TD-SCDMA, DECT and UWC-136) grouped together under the IMT-2000 label [ITU03, ITU97].

The fourth generation of wireless communications (4G) is still under world wide research. Different manufacturers, academics and organizations are already working on 4G research and development issues. The European Commission carries out different projects for the definition of 4G systems like MATRICE, 4MORE or RACE. Its main goal is to provide multimedia services over the air interface with data rates over 30 Mbps. For these higher data rates, the frequency band of 5 GHz, and perhaps later of 60 GHz, are foreseen. The European Radio communications Office (ERO) has already studied the spectrum assignment needed for the future 4G systems and the work on a standard for operation above 60 GHz had already begun.



Figure 15: Integration of different applications in a mobile terminal

With the introduction of 2G systems, wireless communications have entered the digital world

enabling its evolution and possibility to offer numerous new services and applications. Apart from extremely popular short message systems (SMS) the use of digital technology led to the inclusion of computer like functionalities inside the mobile phones. These, so called *smart-phones* include Internet and e-mail access, alarm clocks, calculators, voice memo recording, personal organizers, FAX, infrared and Bluetooth connectivity, call registers, ability to watch streaming video and above all the ability to read business documents in a variety of formats such as PDF and Microsoft Office. The first in line of smart phones was Nokia 9000, released in 1996. Further advances in digital technology and microelectronics gradually led to the integration of different cellphone-unrelated applications (Figure 15) like FM radio receiver, built-in camera and camcorder, mp3 music player, PlayStation like games, GPS navigation software, live television streaming etc, at the same time keeping the terminal as small as possible. Thus, nowadays a mobile phone can connect wirelessly to a computer, download, store and reproduce music or video of high quality and all that with no keyboard, using just touch screen control. It can take photos with 12Mpixels camera and transfer it wirelessly directly to the printer, or pass it to the computer for image edition. It can guide you to the nearest gas station using GPS system, and you can even use it as a credit card to pay for the gas. Eventually, what once was just a mobile-phone with a suitcase as a battery nowadays is being converted into mobile-camera-television-computer-PlayStation, and finally phone terminal small enough to be carried in the pocket or worn around the wrist.

Inside this short history, two tendencies of mobile phones terminal can be acknowledged: terminal minimization and the increase of value added functionalities and applications through hardware and software. In fact, with the implementation of the non-voice applications, mobile terminal changed into a multi-functional product focusing more on data than on voice transfer. However, as data applications evolve rapidly, higher data rates and consequently more computational power was required with each new terminal generation. This phenomenon can be clearly observed in Figure 16. where the increase in data rates of cellular standards (1G to 2G to 3G to 4G) is followed by exponential increase of algorithmic complexity. The computing power requirements are measured in MOPS (Millions of Operations Per Second). During the transition from 1G to 2G the number of MOPS remained almost flat while an exponential increment started with 3G and continues on with 4G. Therefore, the increase of processing power can be seen as the third tendency of mobile phone terminals. In accordance with the required number of MOPS the first cellular phones were based on application specific integrated circuits (ASICs) while the second generation include terminals with DSP (Digital Signal Processing) chip-sets. However, DSP core chip-set performance could not follow the mobile

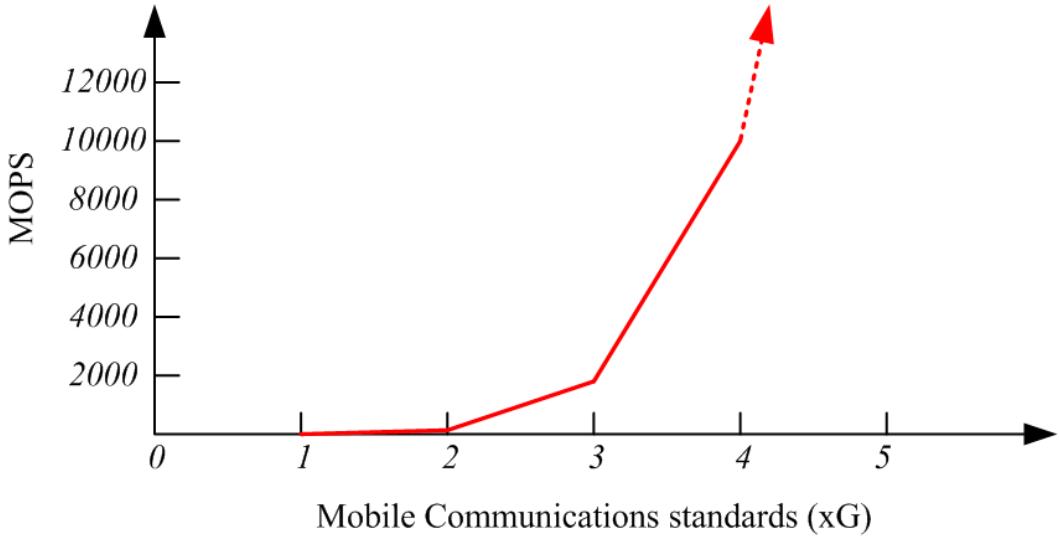


Figure 16: Computational power evolution of different mobile communication generations (xG)

terminal advances, and was surpassed rather quickly. Consequently field programmable gate arrays (FPGAs) were introduced, and lately even general purpose processors (GPP) like Intel Pentium ones are being mounted. Nowadays, a typical high-end smart-phone can match the performance of mid-range laptop computer only five years ago. By taking a peak into the inside of the latest state-of-the-art *iPhone* form Apple more than a dozen different chip-sets are found on the logical board. These chip-sets, apart from the core Samsung 620MHz ARM general purpose processor (Figure 17) also include:

- 2 SDRAM chips
- Wolfson audio chip
- Linear Technology USB Power Li-Ion Battery Charger chip
- Marvell's 802.11b/g chip
- Skyworks GSM/Edge Power amplifier chip
- CSR Bluetooth processor
- Infineon PMB8876 S-Gold 2 multimedia engine chip
- touch screen control chip

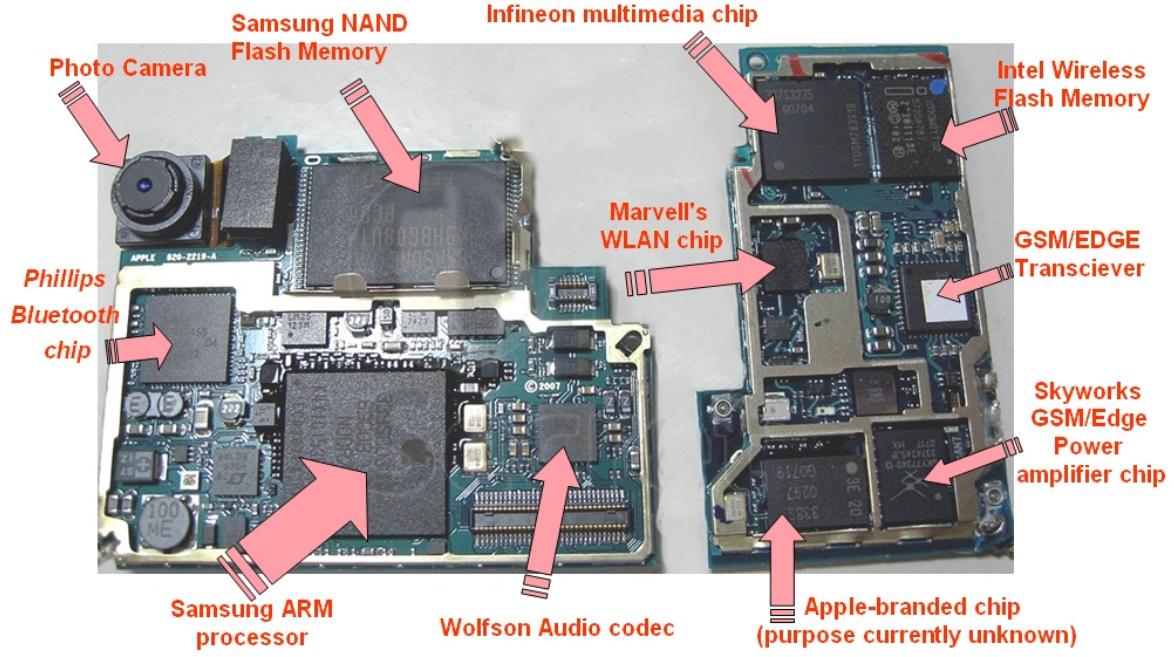


Figure 17: Upper side of iPhone logical board with Samsung core processor and many on-chip components

• ...

Indeed, it looks like the whole mobile terminal has gone digital. However, there is one exception as a unique RF front-end chip is not present. The explanation is rather simple as due to the technological constraints one unique RF front-end chip supporting different wireless communications offered by the single device is still not commercially available. RF front-end is still designed separately for each radio standard, and than each transceiver chain is packed apart in the intention to isolate it from chipsets energy dissipation influence and to keep the analog parts apart. Nevertheless, the current trend, described also as the fourth tendency of mobile phones, is to move the interface between analog and digital parts as close as possible to the antenna. The principal argument for replacing analog with digital signal processing lies in the possibility to *softly* reconfigure the system, increasing its integrability and flexibility. In general, the system of this kind is usually referred to as the software-defined radio or simply software radio (SWR) [Mit95, Bur00, Hen99, Rad04]. The true goal of this trend is to have an all digital and more flexible system. Consequently, the current research emphasis is directed on wireless multi-service, multi-standard, multi-band technologies and configurations suitable

for integration on a multichip package. This multichip package approach is used in *iPhone* terminal as each radio standard is based on its corresponding front-end chip. However, from the technological point of view this is only a middle step as it is very probable that the future wireless terminals will be built around the single silicon RF front-end chip able to adapt to any radio interface simply by changing the software.

Nevertheless, the process of replacing analog with digital signal processing on a single chip, is not as trivial as replacing the external components with on-chip ones. This is mainly due to the technological constraints which cannot be solved easily and rapidly at the moment. At present it is not reasonable to use a single RF stage for a multiband system since it is still impossible to build antenna and low-noise amplifier (LNA) for a whole commercial frequency band ranging from a hundreds of megahertz up to tens of gigahertz. Commercial three or even four-band mobile terminals in fact have each band transmit-receive chain built separately and at the power up initialization, one chain is selected. Other technological challenges are linked to analog-to-digital and digital-to-analog (D/A) converters since a trade-off choice between sampling rate and resolution is a necessity. Today's technology reaching 1 Gsample per second with 10-12 resolution bits is insufficient taking into account high dynamic range of a signal. On the other hand, 16-bit resolution is available only at lower sampling rate of around 200 Msamples/s. Having all this in mind SWR system are not expected in commercially available equipment until 2012-2015 leaving the current mobile radio systems still based on the traditional super hetero-dyne scheme.

The fifth tendency, introduction of multiple antennas at receiver and transmitter, emerged with the new capabilities of mobile terminals (like receiving and processing large images, real time video, 3D graphics, interactive multimedia) which created constant demands for higher bit rates. Eventually, the rates of several thousands of bits per second pushed the traditional RF architecture with one antenna beyond its limits. As single antenna could no longer cope with these bit rates, multiple antenna systems are proposed. This, so called Multiple Input Multiple Output (MIMO) technology is being proposed for the 4G systems where multiple antennas are used in transmitter and at the receiver in order to gain diversity and/or achieve high bit rates [Bar04].

The sixth tendency, represents the everlasting industrial search for more profit keeping aside important quality parameters. Even though the manufacturers will always deny it, anyone that ones had a mobile phone can confirm that after a year or so the mobile battery starts to die out, the keyboard doesn't respond as good as when you bought the phone, and the

1.	terminal minimization
2.	raise of value added functionalities (hardware and software)
3.	increase of processing power (GPP are introduced)
4.	moving the digital interface towards antenna (RF still contains analog elements)
5.	introduction of multiple antennas
6.	cutting the product manufacturing costs (<i>fancy on the outside, cheap on the inside</i>)

Table 1: Six principal mobile tendencies

conversation is not so robust to the interferences as it ones was. The majority of users tolerate this non-functionalities for some time, and then a new phone is bought. According to [Bud07], mobile phones are replaced every 18 months (and this time is still shrinking), so I guess, there is no sense to build them to last longer. Mobile phone prices also confirm this life expectancy, as the newest models cost around 400 euros while the last year models can be found for bargain. Since the success of mobile phone model depends on its outer design, the inner components are seen as less important. Naturally, more profit is reached at the expense of *less* important elements. As a direct consequence the interior of mobile terminal is filled with low cost electronics. I also describe this sixth tendency as *fancy on the outside, cheap on the inside*.

Conclusion

When all six tendencies in Table 1. are put together the result is miniaturized mobile terminal with analog multi-antenna RF front-end, filled with numerous hardware components. Taking into account cost restrictions, low-cost electronic components are slowly overpopulating the physical space of the receiver terminal. Further on, digital components are interlaced with plenty of data metallic wires and different analog elements, which create magnetic fields, increase operating temperature and finally favour the appearance of signal degradation. Since low-cost components tend to dissipate energy, at the same time capturing the energy released by another source, a rather hostile environment is easily created. This kind of environment stimulates the appearance of coupling as these distortions occur mainly inside the places where no energy dissipation protection is present (due to the lack of space), or where low-cost electronic elements are used in order to reduce the product costs. Eventually, the interior of mobile terminal fulfills both assumptions. It can be stated that coupling distortion appears

as a direct consequence of four cellphone tendencies (miniaturization, new hardware functionalities, multiantenna RF front-end with analog elements, low cost electronics). However, two other tendencies (increasing the processing power and moving the digital interface towards the antenna) offer a foundation for solution of this problem by software coupling cancellation at the proper mobile terminal.

Although coupling appears both in transmission and reception paths in mobile terminal, this thesis focuses on the cancellation at the reception and finding the cheap, yet effective way to overcome, or at least reduce, the coupling effect using the resources available at the mobile phone. When coupling is present at transmission in mobile terminal, more sophisticate and computationally complex algorithms can be used at base station for its cancellation. Thus, this situation is not a subject of research here. Furthermore, even though coupling appears in all MIMO transmission at larger or smaller scale, cellular mobile receiver has been chosen for case study as it is the most demanding product from the point of view of processing power and battery restrictions. Therefore, the developed nonlinear coupling cancellation algorithm can easily be implemented in any other MIMO system environment possessing at least the same amount of processing power.

3 RF Front-end description

During the years, as the size of the terminals decreased, large analog components have been replaced one-by-one with on-chip ones moving this way the interface between the digital and analog technology into the RF front-end domain. The actual definition of RF front-end domain is quite generic as it includes all components found between the antenna and the input of the A/D converter at a receiver. Figure 18. summarizes the steps in the migration from traditional architecture with digitalization in base band (BB) stage towards software radio devices based on General purpose processors (GPP).

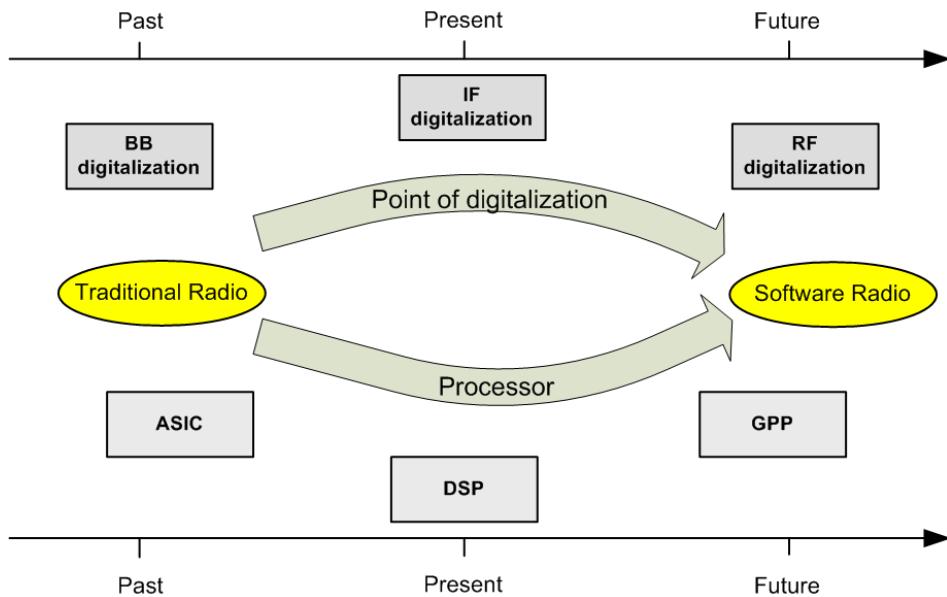


Figure 18: Technical evolution of SWR radio

Actually, the RF area of mobile phones was traditionally dominated by discrete circuits technology but the general course toward miniaturization, low power and low cost terminals directly influenced RF section as well. The RF front-end of these new terminals has experienced significant size reductions (Figure 19) due to evolution of radio architectures enabling semiconductor integrated circuit (IC) technologies, and innovative circuit design techniques [Lok02b]. At present, the employed RF front-end is still mainly based on low-IF or zero-IF architecture with analog signal reception, digitalization at IF stage and signal processing based on DSP technology but the level of integration has radically changed. As DSP processors are becoming faster, many of the radio system functions are implemented in software. Eventually, A/D

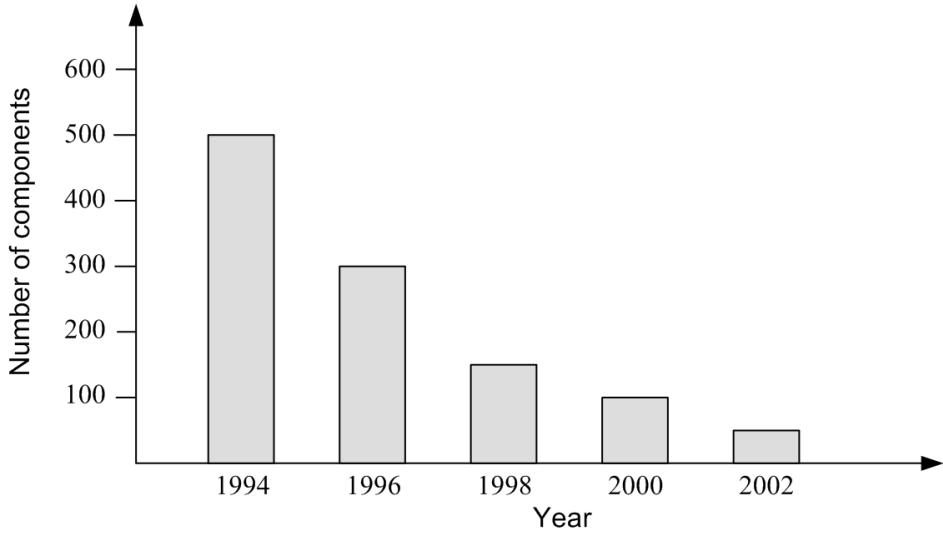


Figure 19: RF front-end component count reduction in GSM mobile handsets over the years

conversion is approaching the antenna leading toward the software defined radio system. It must be stated that moving the digital interface towards the antenna is getting more complex as we closer to the antenna, and it is not as trivial as replacing the external components with on-chip ones [Che96, Chi01, Agn06]. Furthermore, putting two or more antennas influences the complete RF front-end design as additional distortions have to be taken into consideration.

One of the principal problems of MIMO systems is the lack of the physical space inside the terminal since the interfering signals from proper terminal are here more likely to occur and deteriorate the system performance. When using practical electronics, because of the closeness of metallic data wires and different electronic circuits on one hand, and the lack of energy dissipation protection on the other, a phenomenon known as coupling is bound to emerge. The coupling is defined as the desirable or undesirable transfer of energy from one to another physical medium (like metallic wire or an optical fiber), including the transfer of power from one circuit segment to another. Considering the presence of nonlinear elements like nonlinear amplifiers, nonlinear capacitances, resistances etc, the coupling is a nonlinear phenomena. Due to the high system precision request, simplification to the widely used linear modeling is cast aside as not accurate enough. When designing the coupling model, current RF front-end scheme together with future RF architecture expectations are taken into account. It will be shown that analog elements can not be ejected out of RF front-end resulting in the presence of coupling in years to come.

3.1 Traditional RF Front-End Architecture

The vast majority of current transmitters and receivers employed in mobile radio systems is still based on traditional super-heterodyne scheme. Inside this scheme, RF and IF (Intermediate Frequency) stages are completely analog, while the digital component is only present in the base band (BB) stage, usually based on the application specific integrated circuits (ASIC) technology. Traditional RF front-end architecture consists of only two schemes: super-heterodyne (Figure 20), and zero-IF architecture also known as homodyne or direct conversion (Figure 21). Both schemes use analog technology in RF and IF stages while the digital technology is present in the BB stage.

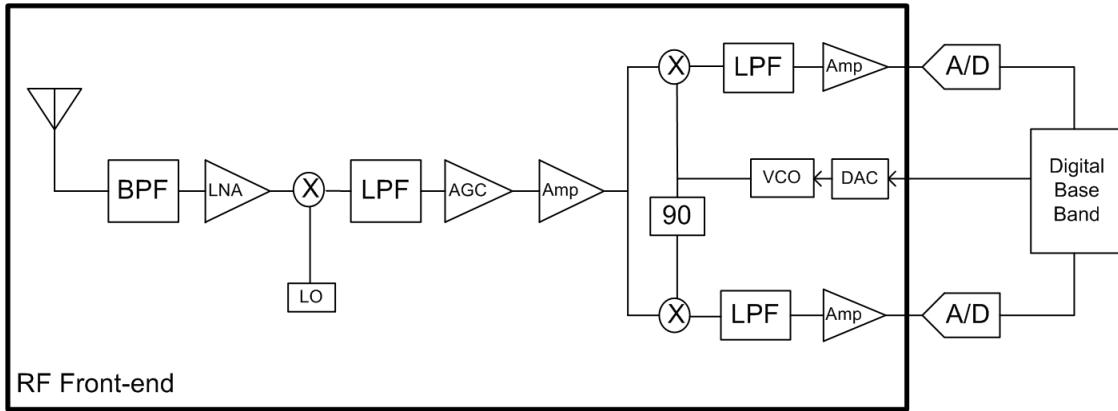


Figure 20: Traditional super-heterodyne receiver receiver architecture

The super-heterodyne receiver shown in Figure 20 consists of many analog elements including Band Pass Filter (BPF), Low Noise Amplifier (LNA), Local Oscillator (LO), three Low Pass Filters (LPF), Amplitude Gain Control (AGC), three analog amplifiers, Voltage Controlled Oscillators (VCO), three multipliers, one phase shifter, two analog-to-digital converters (A/D) and even one digital-to-analog converter (DAC). This architecture faces the problem of image frequency which requires bulky external filter for image cancellation making the system-on-chip integration impossible.

The basic components of homodyne architecture presented in Figure 21 are only a few including RF mixer, two LPF, two AGC components and one LO. The mayor advantage of direct conversion is its high level of integration as the production costs are smaller and energy

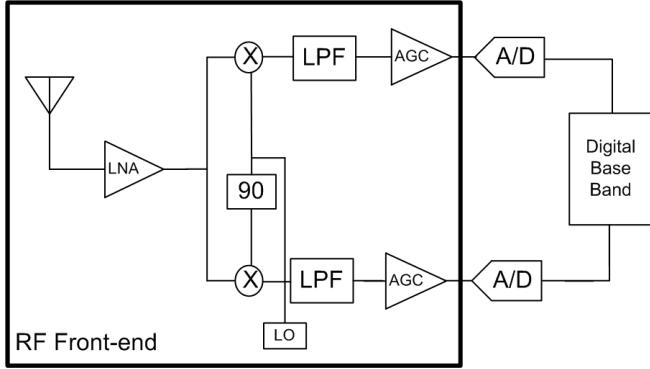


Figure 21: Traditional homodyne receiver architecture

consumption lower. Because the IF frequency is zero, the image of the desired RF signal is the desired RF signal itself, thus discarding the bulky, off chip front-end image filters. Furthermore, as the desired spectrum is down-converted directly to baseband, channel selection is performed either with a simple analog low-pass filter or by converting the signal to digital domain and digitally performing channel selection with DSP. However, when compared to super-heterodyne receiver it had inferior noise characteristics and less stability until the last decade when technological barrier was finally surpassed and direct-conversion receivers started to flourish. Long abandoned in favour of mature super heterodyne architecture, direct frequency conversion has emerged around year 2000 with the introduction of Global System for Mobile Communication (GSM) handsets. Nevertheless these systems can still not be described as true direct conversion receivers. In practical direct conversion receiver (DCR) implementations the RF signal is down-converted to a low intermediate frequency and than the A/D conversion is performed on this IF signal. Apparently, this receiver is designed as a two stage down-converters but with the direct conversion frequency inside the IF area. This way the image frequency problem is reduced and the desired spectrum is *almost* directly converted to the baseband for further digital processing.

Although baseband signal processing is realized digitally in both architectures, parameters related to the digital front-end such as channel spacing, signal bandwidth or carrier frequency are still changed by means of dedicated hardware. As a conclusion, traditional RF front-end is based on a mixture of analog and digital components which interact inside a difficult signal environment. If explored in detail the majority of these analog elements posses the nonlinear characteristic whose level usually increases with the temperature rise. Thus, again nonlinear RF front-end modelling is required if the precise coupling model is to be developed.

3.2 Future Radio System

Future radio systems will evolve around the GPP processor capable of taking charge of all signal processing and control functions. Since radio interface functionalities are controlled by software, this kind of system has been denominated software radio (SWR) system. Amongst many definitions the one found to be the most correct one is defining SWR as “an emerging technology, thought to build flexible radio systems, multiservice, multistandard, multiband, reconfigurable and reprogrammable by software.” [Bur00]

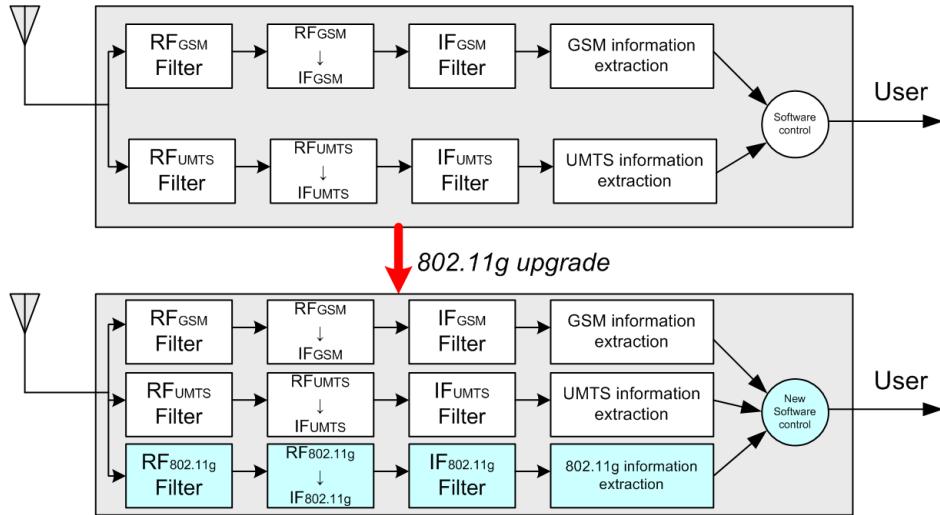


Figure 22: Traditional design of multi band cell phone

The fundamental idea of software radio is the expansion of digital signal processing toward the antenna, and thus to regions where analog signal processing has been dominant so far. By shifting the analog-to-digital interface towards the antenna the possibility of software system reconfiguration is achieved, and consequently system flexibility is increased [Mit97, Cha02, Xu05]. The flexibility of SWR is seen through transparent change of working characteristics, such as operating frequency range, modulation type, signal bandwidth, maximum output power or network level protocols, supported by the alternation of software parameters. Thus, the same SWR device can, for example, be used as cell-phone, FM radio receiver, garage door

opener or even as baby monitor. Indeed, all these functionalities have been technologically possible for many years now. However, their integration into a one device has been based on the assigned hardware receiving chain inside the receiver. Then, once this product has reached the market, adding the additional radio standard is seen as a mission impossible as new components have to be placed inside the terminal itself as shown in Figure 22.

This is not the case with a true SWR terminal as it includes programmable RF bands, programmable channel access modes, channel modulation control etc. The compatibility with any defined radio mobile standard is guaranteed by general-purpose programmable processor (GPP), which is supposed to implement radio interface and upper layer protocols in real time. Thus adding a new radio standard to SWR terminal transforms into a simple software update without the introduction of a new hardware [Lac95, Ber97, Bos99, Chu99]. As is seen in Figure 23 the introduction of a new standard doesn't include any new physical filters or RF-to-IF downconverters as all the processing is done inside the GPP. The development of SWR architecture depends greatly on the production of RF front-end system on a single chip, on the development of wideband AD converters and on the use of powerful GPP systems.

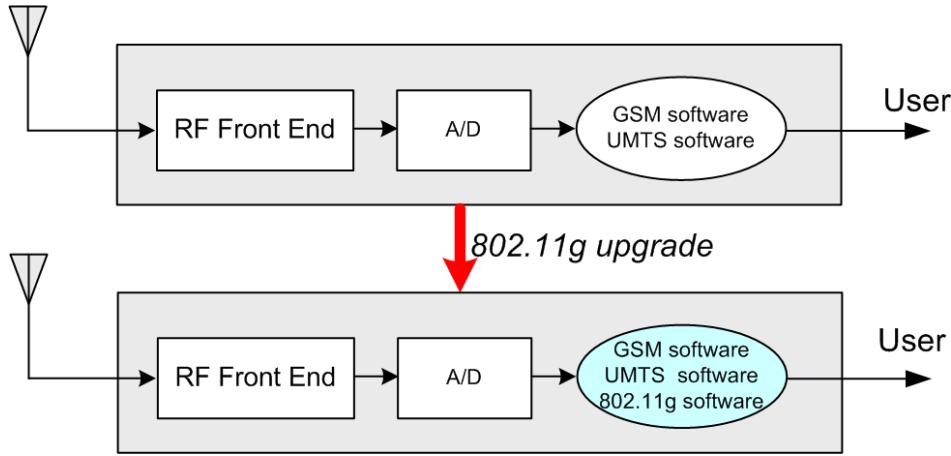


Figure 23: SWR design of dual cell phone

At present, there exist many cellphones offering so called 3-band services, and some of them have even been advertised as the first SWR terminals offering *true* world wide roaming. However, even though these devices are based on digital technology, the programmability of software radio is not present as technological roaming is achieved by tripling the hardware receiving chains. Actually, for up to four radio standards, this kind of traditional implementation is less expensive so we can expect to find even more analog elements inside the mobile terminal

before the next step toward higher level of integration is taken.

3.2.1 Ideal Software Radio

An “ideal” software radio receiver is often considered as an analog/digital (A/D) converter connected directly to an antenna [Rad04]. However, two more analog components: bandpass filter and low noise amplifier (LNA) have to be added in order for this ideal radio to operate. A typical diagram of ideal software radio receiver can be seen on Figure 24 where the analog signal first passes through band pass filter (BPF) and than after the low noise amplifier (LNA) the signal goes directly to analog-to-digital (A/D) converter. Beside the analog part, controlled through feedback connection, DSP controls also the channel selection, demodulation, decoding and other signal processing functions which are implemented in digital domain after the A/D conversion. DSP solutions are chosen ahead of ASIC and FPGA technology because they offer high flexibility and low power consumption [Gra95, Lu00, Lad01, Abe01].

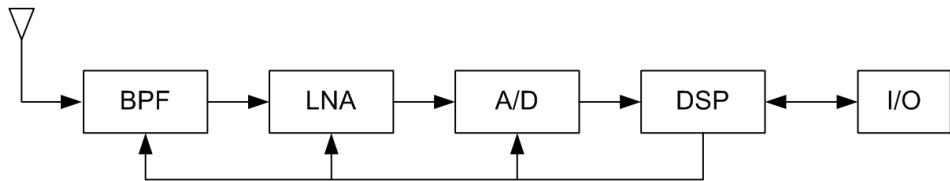


Figure 24: Ideal SWR scheme

In order to see the complexity of the information extraction from an RF signal a case of one one multi-band system is analyzed. For example, in the case of a GSM technology, the mobile terminal receives 124 channels of 200 KHz each situated in 890-915MHz band. The antenna presents the whole RF signal to the receiver which uses RF filtering to select the desired channel and limits the signal to the 200 KHz. The RF filter should be tunable as it must be able to select different channel if the communication channel changes. Since manufacturing accurate tunable filters is expensive, cheap filters are usually placed at the RF stage. In consequence, the output of this filter is not of high quality and the filtering result has more than 200KHz (i.e. has part of the adjacent channels). To eliminate the adjacent bands, the signal is down converted to a lower and fixed intermediate frequency (IF). For GSM signals this would imply down-conversion from 900 MHz to 3 MHz. At IF stage, the signal is filtered again in order

to eliminate the adjacent components. Because all channels are down converted to the same IF, this filter does not need to be tunable and can be highly accurate at lower prices. Now, if the terminal is designed for numerous radio standards, it must be taken into account that various radio standards use different portions of the radio spectrum. As a consequence SWR demands much larger input RF bandwidth than a traditional one-standard radio receiver. Furthermore, since all channelization tasks are performed digitally, the RF front-end and analog-to-digital converter (ADC) have to process the complete signal bandwidth for which the terminal is designed, meaning several parts of the receiver have to process a large number of channels (data) simultaneously. Afterward more complicated functions such as equalization, frequency hopping, error detection and demodulation technique take place which naturally, blows up the processing demands and power requirements of the receiver terminal. Finally, the sole propagation of RF mobile communication signal experience many pitfalls like fading, shadowing, strong blocking, interfering signals (due to the coexistence of several transmit signals) etc. which lead to very high dynamic range reflected in the receiver features. For instance, a wideband signal in the GSM band has a dynamic range in the vicinity of -100 dB. Extending this range to different standards that might appear simultaneously at the receiver, the dynamic range gets even larger. Consequently, a number of narrowband signals at wideband receiver yield a dynamic range far above the one inside the conventional one-standard terminals.

In order to use the ideal SWR scheme illustrated in Figure 24 and achieve reasonable level of RF performance, the components of ideal SWR terminal should meet high technological requirements. However, many of these requirements can not be accomplished in a near future. The limiting factors could be classified in four areas:

- A/D converter parameters
- Power Consumption
- Amplifier linearity
- GPP Processing Power

A/D converter: The available resolution of A/D converter is defined with the thermal noise power present at the converter input as it defines the minimum peak input signal. In order to perform the quantization correctly within the required converter bandwidth, the quantization

noise power must be lower than the sensitivity floor shown in Figure 25. Peak to sensitivity floor ratio, or in other words the dynamic range of an A/D converter, is given by:

$$D_c = 6n + 1.76dB + 10\log_{20} \left(\frac{f_s}{2f_{max}} \right) dB \quad (6)$$

where n is the resolution (number of bits) of the converter, f_s is the sampling frequency and f_{max} is the maximum frequency of the input analog signal. This means, that SWR terminal designed for only GSM and UMTS standards requiring $121.76dB$ dynamic range is achieved with A/D converter of at least 20 bits resolution. The required dynamic range is obtained assuming that at least $100dB$ signal range is needed (e.g., from -20 to $-120dBm$) and that a $12dB$ signal-to-noise ratio is required at minimum sensitivity. [Wep95]

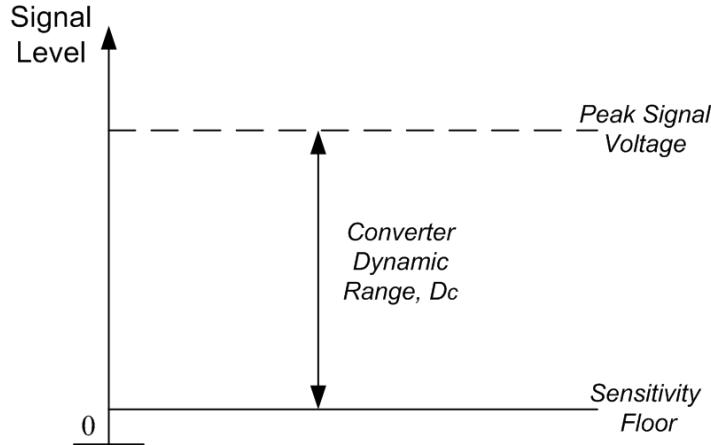


Figure 25: Dynamic range of an A/D converter

As is seen in eq. (6), the dynamic range of memory-less A/D converter raises $6dB$ with every additional bit of quantizer resolution. Likewise, as the sampling frequency is increased beyond the Nyquist rate of $2f_{max}$, the dynamic range also increases. This occurs because the quantization noise power, which is fixed and independent of bandwidth, is spread out over an increasingly wider band as the sampling frequency is increased. Theoretically, additional $3dB$ are obtained by doubling the oversampling factor. However, the true benefit of oversampling is found in the form of an effective increase in the number of bits of A/D resolution. This effective increase denominated N_{eb} , is given by eq. (7) where M is the oversampling factor,

given by eq. (8), B is the RF channel bandwidth and f_s the sampling frequency.

$$N_{eb} = \frac{10\log(M/2)}{6} \quad (7)$$

$$M = \frac{f_s}{B} \quad (8)$$

This means that using a 200 mega samples per second converter on a $200kHz$ bandwidth channel ($M=500$) effectively provides an extra 4.5 bits of A/D resolution. Thus, a 16-bits converter could be used instead of the 20-bits one. Nevertheless this technique is not used in practice as the increase in sampling rate leads to power consumption increase, and this negates the benefits of extra resolution.

Today's technology, allows us to reach up to 1 giga sample per second sampling rate with 10 – 12 resolution bits, decreasing to 400 Msamples with 14 bits, arriving at 150 ksamples with 20 bits. However, this dynamic range is not covering the whole wideband RF signals as it is offered only in a certain frequency bands. Practical systems are typically designed according to the Nyquist criterion with modest oversampling. Nyquist criterion states that the minimum sampling rate f_s of the A/D converter must be at least twice the bandwidth to be digitized (B). Assuming that the SWR terminal supports GSM and CDMA scheme an oversampling with factor 8 should be used as a method for relaxing the analog anti-alias and IF filtering needs according to [Kra97, Ken00]. Thus, the sample rate of approximately 48 mega samples per second is required .

Apart from the dynamic range and the resolution bits, the third parameter describing A/D converter features is the supported converter bandwidth. Here, again the Nyquist criterion limits the theoretical bandwidth that can be digitalized in practice. For example, in order to cover Professional Mobile Radio (PMR), GSM and the UMTS satellite band, the RF input bandwidth in range of $100MHz$ up to $2.2GHz$ is a necessity. Current A/D converters are far from meeting these severe SWR terminal requirements, hence indicating that converter technology has a long way to go before the design of a true SWR terminal may be attempted. Besides, jitter effects make A/D conversion directly at RF very difficult. On the other hand, converter technology is continually improving, and it is not unrealistic to predict that further improvements in resolution and sampling rate will continue to be made. In addition, the trend of migration from bipolar or BiCMOS implementations to CMOS implementations brings further potential benefits and opens new perspectives to hi-tec A/D converters.

Power Consumption: The inclusion of general purpose processors, additional digital processing chipsets and fast A/D converters would increase the energy consumption of any theoretical device. Inside a cable network this problem is usually not even considered, but inside a mobile handset, where autonomous operation could be limited to just a few hours with current batteries, power consumption is of high importance. So far, the most critical part of energy consumption inside a mobile terminal have proven to be an A/D converters as their large energy consumption is also accompanied by intense energy dissipation leading to terminal overheating. This issue is particularly critical inside mobile devices since due to the lack of space no efficient cooling systems can be installed. Consequently, the A/D power consumption for SWR mobile device should be within the range of 50 to 150mW.

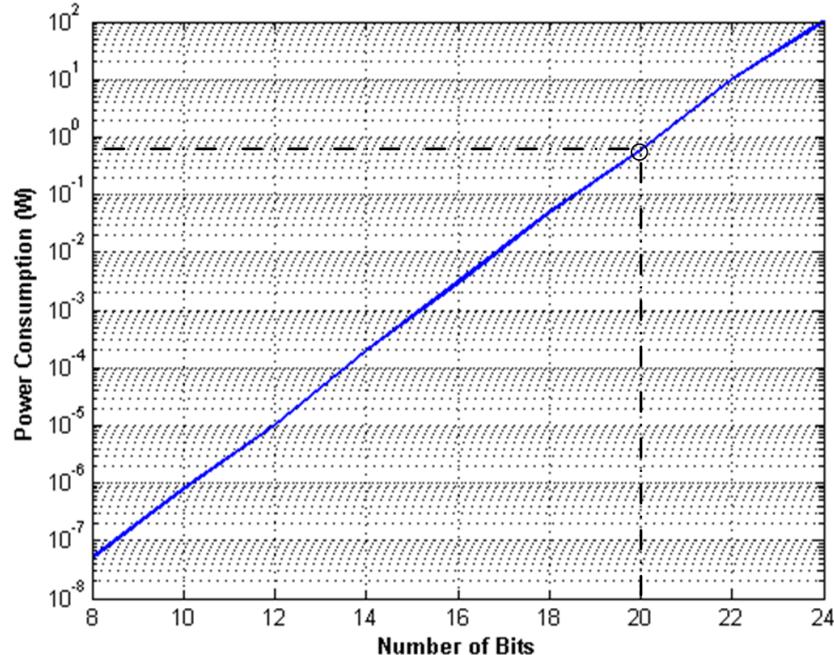


Figure 26: Minimum theoretical power consumption of A/D converter operating at 48MSPS and 20bits resolution

According to [Ken00], theoretical minimum power consumption for A/D converters can be expressed as:

$$P = \frac{kT}{t_s} 10^{(6n+1.76)/10} [W] \quad (9)$$

where k is Boltzmann's constant ($k = 1.38 \times 10^{-23} J/K$), T is the device temperature (in

Kelvin), t_s is the sampling interval and n is the number of A/D converter bits. Thus, the A/D converter operating at 48 mega samples per second with 20 bits resolution (the case of SWR terminal with GSM and CDMA standard supported) has the minimum power consumption of around $600mW$ as shown in the Figure 26. These numbers are far from the maximum allowed A/D power consumption inside the SWR mobile device.

As an example of present technological achievements, the latest A/D converter from analog devices denominated AD9254 uses 14 bits resolution and 50 mega samples per second over the input bandwidth of $650MHz$. In order to reach this performance, it consumes a maximum power of $500mW$. This is clearly a high value for consideration to be included in a handset and would probably be considered excessive in most designs.

Amplifier linearity: Linearity is of high importance in RF front-end of a SWR terminal as it includes RF filters which not only limit the received signal in frequency but also amplify it to compensate the attenuation due to the propagation over the air. Certain critical RF problems, like the need for amplifier linearity and efficiency across the whole access bandwidth, are here more pronounced as signal bandwidth is easily by several orders of magnitude larger than in conventional systems. When coping with large signal bandwidth amplifiers often introduce signal distortions on the edges increasing BER rate and decreasing system performance. Lately, important research efforts are dedicated to obtaining high quality amplification over extensive bandwidths and digital signal processing has been introduced for keeping the RF signals imperfections as low as possible.

GPP Processing Power: As analog IF processing is replaced by digital signal processing, the processing segment which includes IF stage and baseband conversion is done digitally mainly on dedicated chipsets. A typical process of frequency translation, filtering and decimation requires around 100 operations per sample which at 48 mega samples per second is equal to approximately 5000 million instructions per second of processing demand. Although such microprocessors are on the horizon, contemporary implementations offload this computationally intensive demand to dedicated chips such as the Harris Decimating DownConverter (DDC) or Gray digital receiver chip. In addition, spreading and de-spreading functions of CDMA signal, whose computational load is proportional to the bandwidth of the spreading waveform (typically the chip rate) and baseband signal bandwidth, are also handled on dedicated chips as well. The baseband functions including Trellis coding, soft decision parameter estimation, demodulation, frame alignment, bit-stuffing, radio link encryption etc. all have high processing demands together with severe timing requirements, and are at present mainly

done in dedicated chipsets. The complete real time processing demands of a digital part of a transmission chain are still slightly ahead of the computational power offered by hi-tec GPP processor of the ultimate generation. When the price of these GPP processors and their power consumption are taken into account we can expect terminals to use dedicated chipsets in many years to come.

3.2.2 Feasible SWR Receiver

The SW radio receiver illustrated in Figure 24 has been defined as ideal because there are several matters that makes it, at the moment, far from realizable. As previously mentioned, due to technological limitations (power consumption, bandwidth, resolution), it is not reasonable to use a single RF stage for a full band digitalization of a multiband system. Therefore, the bandwidth of the A/D converter has to be reduced, relaxing this way the digital front-end requirements as less digital data have to be handled. This idea leads to the concept of partial band digitalization where in order to cover all radio standards supported by the SWR terminal, a limited band is selected out of the full band by means of analog conversion and intermediate frequency (IF) filtering. This concept results in an architecture employing IF sampling and performing A/D conversion on this IF signal. This kind of a receiver shown in Figure 27 is known as the digital radio transceiver.

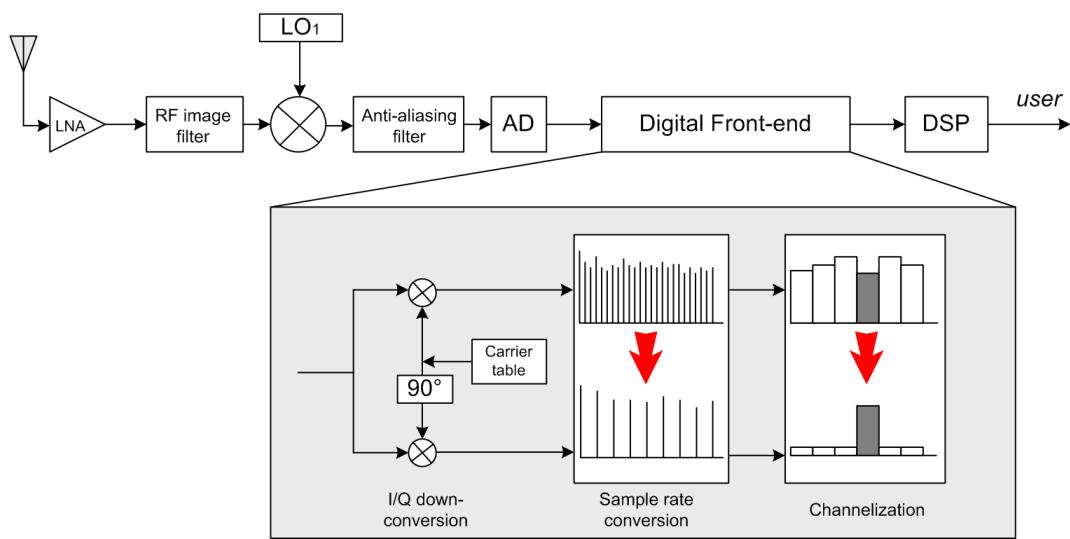


Figure 27: Feasible SWR radio architecture

The structure of feasible SWR receiver is very similar to the wideband transceiver, with the RF stage completely analog and digital domain extending toward the IF area. The A/D converter samples the IF spectrum allocated to the system, while the programmable downconverter (PDC) provides the digital conversion from IF to BB, using a look-up table containing the samples of a sinusoidal carrier. The look up table replaces the local oscillator used in an analog downconverter. After the IQ signals have been separated digital sample rate adaptation takes place, followed by channel selection. Since the output of a digital front-end is a base band signal further digital signal processing can be implemented. Actually, in the future, it seems that the only way to guarantee the multiband feature will be to have more RF stages, depending on the radio bands for which the SWR terminal is designed. For example, if the terminal is to support 2G and 3G mobile communications the first RF stage would choose between $900MHz$ for 2G-GSM mobile systems or $2GHz$ for 3G-UMTS system. The band selection will influence the rest of the receiving chain only in the selection of the adequate software parameters.

Conclusion

Due to the actual technological limitations SWR terminals based on direct A/D conversion over RF signal can not be achieved. Thus, feasible SWR terminal based on scheme similar to the super-heterodyne technology is designed. When compared to the traditional design the difference appears to be in the position of analog-to-digital interface which is moved to IF stage. By keeping the RF stage analog, receivers based on traditional RF front-end architecture, as well as the ones based on future SWR technology, when simplified can be seen as a simple mixture of analog and digital components. This is important from the point of view of coupling model as analog elements like amplifier, RF filters or local oscillator will keep generating the same distortions in SWR terminal as inside the terminal based on traditional technology. Thus, since the same distortions have to be taken into account for both RF front-end architectures when designing the coupling model only one unique model is required. On the other hand, since processing capabilities of the next generation mobile devices will rise, decoupling functions can get more complex and should offer solutions in less time and with higher accuracy.

3.3 Typical single antenna RF Front-end Imperfections

When compared to heterodyne structure, direct conversion topology has become an attractive solution for meeting the commercial market demands for low-cost, low-power, wide bandwidth, and highly integrated RF circuitry. Actually, as of year 2000 the direct conversion receiver with I/Q mixer have been used as a standard for GSM handset design. Quadrature mixing simplifies the RF front-end as it eliminates the need for external bulky analog image rejection filters, voltage-controlled oscillator (VCO) and phaselocked loop (PLL), saving the space and reducing the cost. Precisely for this reasons the overview of RF imperfections is given with a focus on a direct-conversion structure.

Inside a DCR, the received RF signal is amplified by a low-noise amplifier (LNA) and then directly demodulated to baseband in-phase (I) and quadrature (Q) signals. Channel selection and gain control are achieved by on-chip low-pass filters and variable gain amplifier (VGA) at baseband. By filtering at baseband frequencies, device imperfections are less severe, the implementation of power-efficient on-chip filtering is feasible. With the introduction of quadrature mixing in traditional RF front-end architecture apart from standard phase-noise [Vou93, Pol95, Rob95, Nik04, Sch04a, Li07] and thermal noise distortions [Rut71, Sch93, Ou99, CH00, Sch04b], a completely new set of imperfections like second-order nonlinearity, DC offset, I/Q imbalance, and flicker noise gained the attention of scientific research community as they are limiting factors of the entire system.

3.3.1 DC offset

The direct current (DC) offset noise is perhaps the most serious problem inside a direct-conversion receiver as it may be larger than the received signal. Secondly, it does not only corrupt the received baseband signal but also may cause saturation in the baseband amplifiers and in the following stages of the receiver. The main causes of DC offset are component mismatch and self-mixing of the LO signal due to the insufficient on-chip isolation. The offset appears as interference at the center of the desired down-converted base-band signal. It can be divided into two components:

- static or time-invariant offset
- dynamic or time-variant offset

Component imperfections like inherent DC value of the analog circuits or interaction of the input signal with circuit generated second-order nonlinearities generate static DC-offset. For example, if the nonlinearity of practical amplifiers is approximated by the standard Taylor power series:

$$y(t) = k_1 x(t) + k_2 x^2(t) + k_3 x^3(t) + \dots \quad (10)$$

and a single-tone input signal, $x(t) = A\cos(\omega t)$, is applied to the circuit, the resulting output signal is given as:

$$y(t) = k_1 A\cos(\omega t) + k_2 A^2 \cos^2(\omega t) + \dots \quad (11)$$

Assuming that higher order nonlinearities are much smaller than the second order nonlinearity:

$$y(t) \simeq k_1 A\cos(\omega t) + \frac{1}{2} k_2 A^2 [1 + \cos(2\omega t)] \quad (12)$$

$$y(t) \simeq \frac{1}{2} k_2 A^2 + k_1 A\cos(\omega t) + \frac{1}{2} k_2 A^2 \cos(2\omega t) \quad (13)$$

In eq. (13) the DC component ($\frac{1}{2} k_2 A^2$) is related to the second order nonlinearity (k_2) and is magnified by the square of the input signal amplitude. Therefore, any large signal within the passband of a DCR can generate a large DC-offset if the even-order performance of the receiver is poor (i.e., k_2 is relatively large).

The leakage of LO signal generates static as well as dynamic DC offset since LO signal can couple through the substrate and bonding wire to antenna, low noise amplifier and the RF port of the mixer as shown in Figure 28. In order to provide sufficient drive for the mixer switching core, the LO signal is typically strong. Thus the LO with sufficiently high amplitude can leak back to the output of front-end LNA or in the worse case to the entrance of LNA. If LO is represented by $\cos(\omega t)$ and the leakage by $\alpha \cos(\omega t + \theta)$, where α is the relative gain in the leakage path and θ is the relative phase shift of the coupled LO, the self-mixing of the LO signal can then be represented by:

$$I_{SM}(t) = \cos(\omega t) * \alpha \cos(\omega t + \theta) = \dots \frac{\alpha}{2} * \cos(\theta) + \dots \quad (14)$$

$$Q_{SM}(t) = \sin(\omega t) * \alpha \cos(\omega t + \theta) = \dots \frac{\alpha}{2} * \sin(\theta) + \dots \quad (15)$$

These two equations mathematically describe the direct consequence of finite isolation between the LO and RF ports of a mixer allowing the LO to be self-mixed with a phase-shifted version of itself generating static DC offsets as seen in Figure 28a.

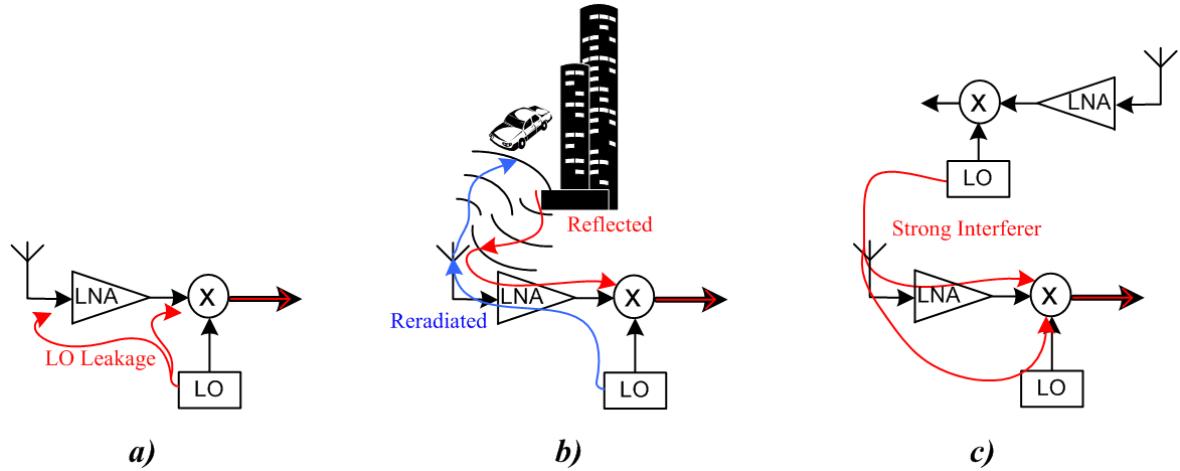


Figure 28: DC offset mechanisms

When LO couples to the antenna, radiates and is then reflected back to the antenna, like in Figure 28b, a dynamic offset is created. It is denominated dynamic as with fading and multi-path reception, the received power level can vary rapidly, which results in a quick time-varying process. This LO re-radiation can also cause problems for other users in the same receive band as this LO signal appears inside their passbands. Thus, as shown in Figure 28c, a strong nearby interferer, such as another user's LO, can also generate a DC-offsets by finding a path to the mixer LO port.

There are several ways to deal with DC-offset issue. Due to its simplicity, the cost-effective straight forward solution is to use high-pass filter based on AC coupling which allows low-frequency disturbances to be removed with minimal distortion to the signal spectrum. This approach has been successfully applied in pager systems to receive frequency shift keying (FSK) modulated signals [Abi95, Nam00]. However this method is not always feasible as several other modulation schemes, like QAM, have significant signal content at DC, thus AC-coupling can corrupt the signal. This effect of corrupting the information bearing low-frequency part of the signal, is typically referred to as baseline wander. It makes the detection of the signal difficult and causes loss of noise margin [Gab97, Hay00]. To compensate for the baseline

wander degradation caused by the AC-coupling, the quantized feedback technique is used in [Sha03a]. Theoretically, AC coupling approach can also be used for wide-band signals, like CDMA system, as almost no signal energy is found near DC component. However, since the corner frequency of the high-pass filter has to be extremely low, of the order of tens of hertz, the required on-chip area for capacitors of these characteristics becomes too large. Large size capacitors also tend to react slowly to dynamically changing DC levels. Therefore, other methods, or combinations of methods, must be employed. Dynamic calibration of DC offset and different digital signal processing techniques used for the calibration of DC offset are some other popular techniques employed to minimize signal degradation [Sam92, Raz97, Yos98, Mat00, Wan03, Sha03b, Pha05, Gil05].

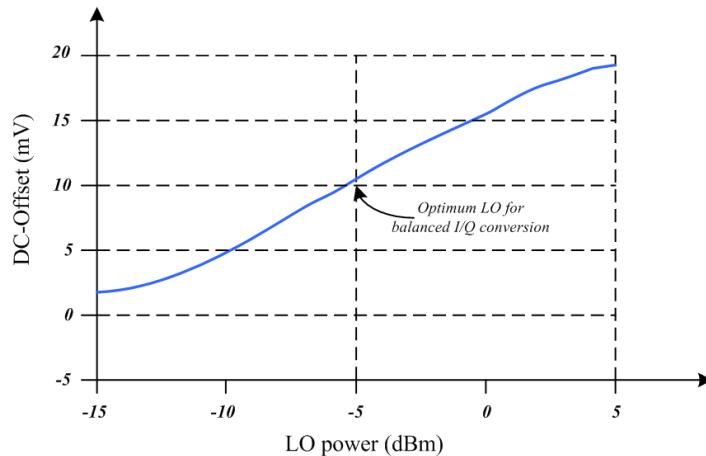


Figure 29: Measured DC-offset for the I-branch of the DCR front-end [Svi05]

It should be stated that with accurate circuit design, adequate linearity and on-chip isolation most LO leakage can be reduced to minimum. However, these guidelines are not followed in practice. Figure 29. shows the measured DC-offset versus LO power at the IF after removing the static offset [Svi05]. The consequence of LO self-mixing is clearly seen as the DC-offset increases with increased LO power indicating the presence of LO self-mixing. As the corners of this dependency are bended the most accurate approach would be to model it with nonlinear function. However, this is not the case, since DC-offset generation is generally modelled as a linear process. Nevertheless this nonlinear dependency only confirms the existence of nonlinear effect inside the DCR receiver and backs the nonlinear RF front-end modelling.

3.3.2 Flicker noise

The flicker noise is found in many natural phenomena such as nuclear radiation, electron flow through a conductor, or even in the everyday life environment [Sch07, Mil03]. It is seen as low-frequency noise which also occurs in almost all electronic devices. It results from a variety of effects, such as impurities in a conductive channel, the flow of electron through a conductor, noise generation in a transistor due to base current, and so on. Flicker noise is essentially random, but because its frequency spectrum is not flat, it is not a white noise. It is often referred to as pink noise because most of the power is concentrated at the lower end of the frequency spectrum corresponding to visible pinkish light.

This kind of noise has only recently gained its attention as its influence in bipolar transistors was negligible. However, with the introduction of CMOS integrated receivers its significance grows as its presence was associated with crystal surface defects in semiconductors [Raz97, Mir00a, Nam01]. Measurements show that flicker noise is most pronounced in surface-transport devices such as the CMOS due to random trapping/detrapping of charge carriers at the oxide-silicon interface [Raz99]. It is typically in the range of $4 - 8\text{kHz}$ for the BiCMOS process, while it is in the vicinity of 1MHz for MOSFET devices [Raz97]. Although there is no unifying mechanism for flicker noise since it is semiconductor processes dependent, there exist an empirical formula, which is adopted in [Tsi99, Liu98b] for the CMOS BSIM3V3.2 model given as:

$$V_n^2 = \frac{1}{f^{ef}} \frac{K_f I_{ds}^{af}}{C_{ox} L^2} \quad (16)$$

where C_{ox} is the unit capacitance of oxide, K_f is a device-specific constant, L is the length of the device, and I_{ds} is the channel bias current while superscript af and ef are current and frequency index. From eq. (16) flicker noise is proportional to the channel current and inversely proportional to the transistor length. Thus, as the size of the terminal device (and transistors) is reduced the flicker noise is increased [Cha94]. Since noise spectral density is inversely proportional to the frequency, the flicker noise is also denominated as $1/f$ noise. The inverse proportionality with frequency is almost exactly $1/f$ for low frequencies, whereas for frequencies above a few kHz, the noise power is weak but essentially flat. Flicker noise is usually defined by the corner frequency f_α [Lok02a], a point where flicker noise is equal to white noise or to cascaded thermal noise floor as shown in Figure 30.

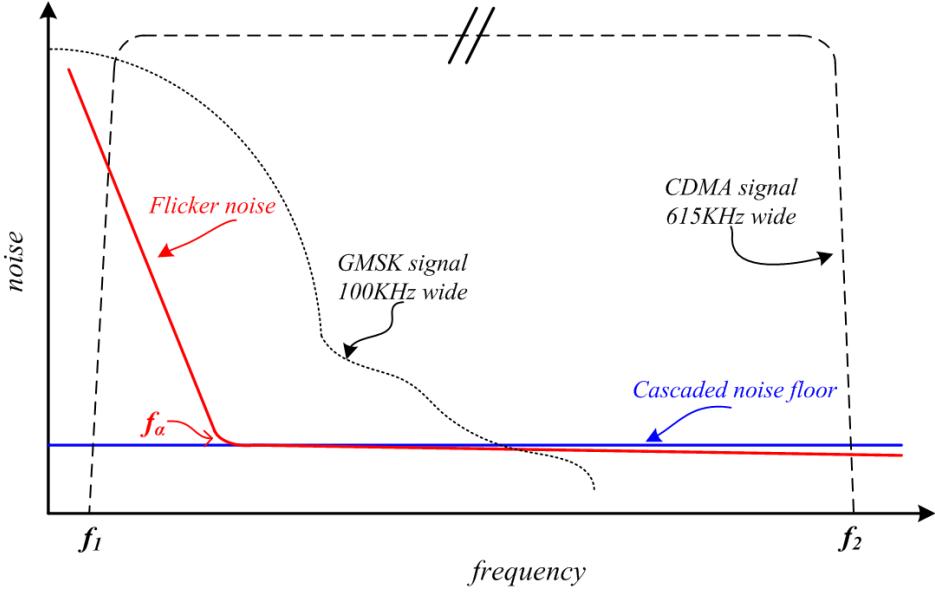


Figure 30: Effect of flicker ($1/f$) noise [Lok02a]

Several publications in the past noted the low flicker-noise property of the passive mixer [Lee98, Raz99], but none of those has carefully investigated its influence on system performance. In narrowband RF systems, bandwidth utilization is a critical issue and it can be maximized by using a reduced flicker-noise receiver system since more channels around the DC area can be used. In a multicarrier modulation system like orthogonal frequency division multiplexing (OFDM) flicker-noise presents smaller problem. Nevertheless, spectral efficiency can also be improved by using a reduced flicker-noise receiver system. It is evident that for a given, narrow-band signals like GSM (100kHz) are more susceptible to flicker noise compared to wider band signals like CDMA 2000 (615kHz) and WCDMA (1.92MHz) as seen in Figure 30. Its effects can be minimized by selection of the right semiconductors and providing the adequate mixer design [Pre02, Sac03].

>From the point of view of this thesis, flicker noise presents yet another nonlinear noise process present in DCR receivers adding weight to nonlinear RF front-end modelling.

3.3.3 Second Order Intermodulation

Besides l/f noise and DC offset, one of the major problem of direct conversion receivers is the second order intermodulation (IM2) which appears as interference around the DC signal component of the original signal. It is caused by the shortcomings of direct-conversion technology with the most dominant source found in the downconversion mixer [Abi95, Kiv01]. In a perfectly balanced mixer, the even-order distortions do not exist. However, in real life, as very linear circuits can not be constructed and both, the RF input transconductors and the switching devices contribute to the overall mixer nonlinearity. Thus, the theoretical perfect balance is undermined and second-order intermodulation appears in base band.

Two distinct mechanisms are responsible for the appearance of IM2, namely, LO coupling and device non-linearities in combination with branch mismatches [Mey86, Hul93, Lau99, Liw00]. The second cause of IM2 depends primarily on a good isolation between the mixing ports, and appears due to physical limitation of fabrication technology [Dar00, Man02], while the first one is generated by the coupling from the RF input port into the local oscillator port. This is very common in frequency division duplexing access techniques or in the presence of strong continuous wave (CW) blocker as shown in Figure 31 where amplifier with poor IIP2 characteristics in the presence of CW blocker generates a DC component.

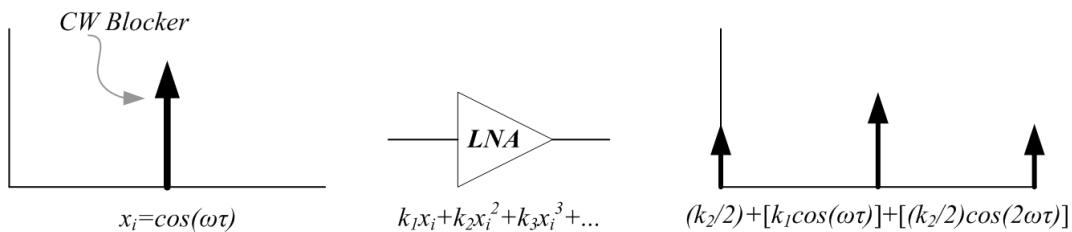


Figure 31: Second order intermodulation caused by continuous wave blocker

If the practical amplifier is approximated by the standard power series:

$$y(t) = k_1x(t) + k_2x^2(t) + k_3x^3(t) + \dots \quad (17)$$

and the $x_i = \cos(\omega\tau)$ is the input CW blocker, the circuit second-order nonlinearity generated by the second element in the eq. (17) is the interference signal at DC:

$$k_2 x_i^2 \Rightarrow a_2 [\cos(\omega\tau)]^2 \Rightarrow \left(\frac{a_2}{2}\right) + \left[\left(\frac{a_2}{2}\right) * \cos(2\omega\tau)\right] \quad (18)$$

Similarly, it can be shown that multiple CW jammer signals can generate interference not only at DC component but also within the desired channel of interest. For instance, if two strong interferer's at frequencies ω_1 and ω_2 experience even-order distortion, they generate a low-frequency IM2 distortion component at frequency $(\omega_1 - \omega_2)$ which can be inside the signal bandwidth and result in signal distortion [Lok02a].

Theoretically, the IM2 interference can easily be eliminated by AC-coupling the RF input stage like DC offset [Raz97]. However, once again, inside a DCR this is not an option, since high integration level is one of the advantages of DCR and the AC coupling components are too large. Since the device matching and offsets cannot be improved beyond certain limits [Man03, Siv05] many different techniques based on DSP processing have been proposed for linear performance enhancement [Yam98, Bau00, Ter00, Dar00, Cof01, Kiv02, Hwa04, Hot04].

3.3.4 I/Q Imbalance

The in-phase/quadrature (I/Q) signal processing is known for its solution to the inherent image signal problem without the need for bulky RF image reject filtering. This results in a clearly simplified analog front-end. However, all downconverters based on this topology share the same problem of matching the amplitudes and phases of the I and Q branches. Ideally, the branches should have the same signal amplitude level and a phase difference of exactly 90 degrees, providing infinite attenuation of the image-band. However, due to finite tolerances of capacitor and analog resistor components, the imbalances of analog front-end are unavoidable in practice. This causes finite and usually insufficient rejection of the image frequency band and as a consequence image signal appears as interference on top of the desired signal [Abi96, Kol96, Raz97, Cro98, Raz98b, Mir00b, Sev00]. When using non-zero IF frequency, the image signal can be up to 100dB stronger than the desired one and the attenuation provided by the I/Q downconversion is just not enough. In this case the complete image rejection is achieved through posterior base band digital signal processing [Cav93, Fre94, Abi95, Gla98, Val00b, Val00a, Val02].

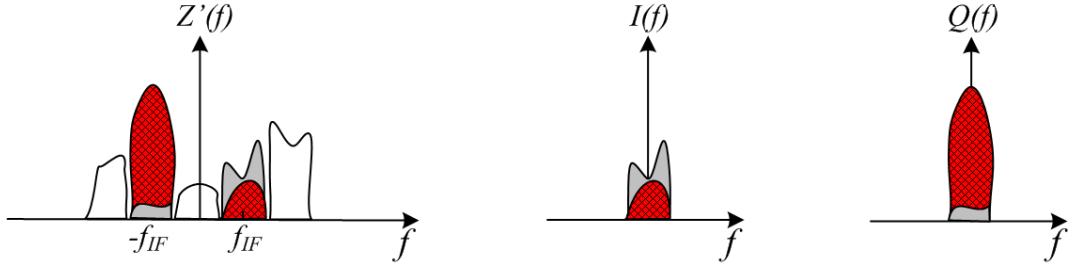


Figure 32: Intermixing of the image signal and the desired signal in I/Q downconverter

In general, if the received signal is denoted as $z'(t)$, baseband equivalent of the sent signal as $z(t)$, and central frequency as f_c , the received signal can be expressed as:

$$z'(t) = 2Re[z(t)e^{j2\pi f_c t}] = z(t)e^{j2\pi f_c t} + z^*(t)e^{-j2\pi f_c t} \quad (19)$$

In order to recover the signal of interest, the baseband equivalent of the sent signal $z(t)$ should be passed, and the image signal $z^*(t)$ rejected. Using analog I/Q downconversion, the mismatches result in an imbalanced baseband equivalent which can be expressed in frequency domain as:

$$Z'(f) = G_1 Z(f) + G_2 Z^*(-f) \quad (20)$$

where $Z^*(-f)$ represents the effect of finite image rejection while G_1 and G_2 correspond to band attenuations [Val01]. This phenomena is also depicted graphically in Figure 32 where frequency spectrums of $Z'(f)$ and the resulting baseband observations $I(f)$ and $Q(f)$ are presented. The intermixing of the image signal and the desired signal is clearly visible as one image signal is superposed over the other. With careful analog design, phase imbalance of $1 - 2^\circ$ and amplitude imbalance of $1 - 2\%$ are realistic, resulting in $30 - 40dB$ attenuation of the image signal [Cro98, Val00a].

This is the first RF imperfection which is actually modelled as nonlinear phenomena as I/Q mismatches are mainly represented as gain and phase mismatches of the imbalanced quadrature mixer with the LO signal $x_{LO}(t)$ presented as:

$$x_{LO}(t) = \cos(2\pi f_{LO} t) - j\beta \sin(2\pi f_{LO} t + \phi) \quad (21)$$

where β is the amplitude imbalance factor and ϕ models the phase imbalance [Sch01b]. Based

on this model, the analysis of QPSK modulated signal in AWGN, is shown in Figure 33. The image reveals that amplitude imbalance of $1dB$ and phase imbalance of 5 degrees cause the degradation of BER performance of approximately $0.5dB$ for $BER = 10^{-6}$ [Liu98a].

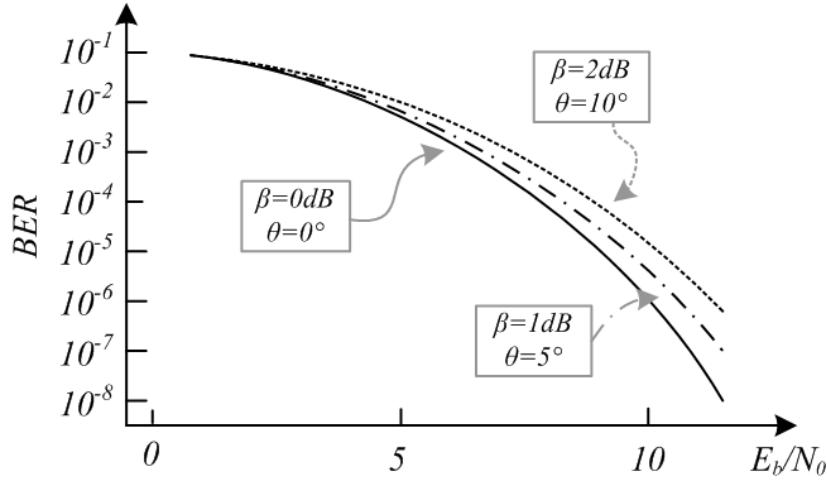


Figure 33: I/Q imbalance impact on BER for different imbalance factors (β, ϕ)

Extensive analysis of the performance of SISO-OFDM systems in the presence of I/Q mismatch can be found in [Sch01a, Val01, Pun01, Rao04, Win04, Tar05a]. Lately, there have been a few papers that discuss the impact of these impairments on MIMO-OFDM systems [Liu98a, Liu04, Rao06a, Tar05c, Tar05b]. Different estimation and calibration methods have been proposed for I/Q gain and phase mismatch compensation [Raz98a, Cro98, Gla98, Li99, Val00b] with the emphasis moving towards baseband digital signal processing in the last few years [Val00c, Val00a, Val02, Hua04].

Conclusion

In this section different signal distortions of DC receiver have been presented. All of them appear to present the nonlinear behaviour, which the majority of authors tries to bypass in order to present the system as simple as possible. It is interesting to notice that all presented distortions can be combated with both, analog and digital approach. Still, the cancellation methods are converging towards digital signal processing as processors are getting more powerfull.

When building a system model, the emphasis is usually made on a certain distortion while all the other imperfections are neglected as if they stop to exist. Inside RF front-end with multiple antennas all these imperfections are intensified as the environment gets more hostile. In addition, when multiple antennas are put close together, due to coupling, the imperfections caused at one antenna can be observed on the other and vice versa. The developed RF front-end model should be able to adapt easily to different distortion scenarios, being able to support the appearance of all distortions at the same time together with the coupling phenomena. This rules out linear coupling model if we are to offer accurate coupling cancellation module.

3.4 Multi-antenna RF Front-end

In January 2006, a new pre-standard 802.11n for WLAN networks was approved. The main idea was to increase the transmission rates in order to reduce the transmission difference when compared to traditional cable networks. However, apart from offering new transmission rates (up to $250Mb/s$) the real revolution was inside the RF front-end which is based on multiple antennas technology. Actually, this was not the first time multiple-antenna systems emerged as this approach has been used since long ago for improving communication system performances. in this standard, multiple antennas in transmitter and receiver are employed in order to gain diversity and to reach high bit rates through multiple-input-multiple-output (MIMO) systems.

It was the paper written by Foschini [Fos98] that demonstrated that multiple antennas can lead to significant capacity improvement in rich scattering environments where diversity gain is exploited, and different signal propagation paths are used to improve the system data rate. The second enhancement of MIMO system is found when spatial diversity is used to improve the reliability of the communication link. The spatial dimension can thus be used in different ways: beamforming (e.g. Space Division Multiple Access, SDMA), spatial diversity (e.g. Space Time Coding, STC) and spatial multiplexing (e.g. BLAST). The two former approaches are adequate in systems with separated antennas and smaller degree of correlation between them. The difference between diversity and multiplexing is that the first one improves system performance while the second one is used to increment data rates. At the moment, beamforming is by far the most used multi-antenna configuration. It is convenient when the distance between antennas is small, and it basically weights adequately the signals in transmitter and/or at the receiver in order to form constructive sum of all signals. Multipath

propagation is an obstacle for beamforming, and usually only the strongest path is taken advantage of. As the mobile system pursue high data rates, in following, the capacity analysis of multiplexing systems will be shortly reviewed in order to demonstrate the advantage of such MIMO system.

One of the most important measures of the channel in communications is its capacity, term introduced by Shannon, that, for a given channel, can be interpreted as the maximum information transmission rate at which the transmitter and receiver can communicate with an arbitrarily low probability of error and no delay constraints [Sha48]. His famous formula for the capacity is:

$$C = W \Delta \log \frac{(P_T + P_N)}{P_N} \quad (22)$$

where W , P_T , and P_N represent the bandwidth, the average transmitted power, and the noise power, respectively.

According to this formula, the channel capacity can be incremented by transmitting in larger bandwidth or by radiating more power. However, none of these two approaches is very recommendable. As inside wide electromagnetic spectrum many different signal coexists, only the assigned bandwidth portion can be used. On the other side, radiation of more power may not be permitted by regulation power masks. Furthermore, due to the logarithmic dependence of the spectral efficiency on the transmitted power, the gain in capacity increment is not very substantial. At high SNR, doubling the transmitted power leads the efficiency increment of just one bit per second per Hertz.

What we are interested in this thesis is the transmission over realistic, fading channels. As the analysis of the channel capacity conducted by Shannon assumed time-invariant channels, it cannot be applied to time-variant channels like wireless channel. Therefore, two different types of capacity are defined for time-variant channels: ergodic and outage capacity. The first one is used when channel varies continuously which is valid, in discrete time models, for block-constant fading (limit value is block of single symbol duration), with independent fading over different blocks. This holds true because of the channel coding that can be performed over independent blocks of fading during the periods that are long enough to reach channel ergodic capacity. The outage capacity should be calculated when the channel is slowly varying and can be approximated as fixed during the whole transmission. Special attention is given to ergodic capacity and realistic channels with fading that vary during the transmission. Foschini

and Shiu [Shi00] calculated the capacity for the case of fixed Rayleigh MIMO channels as:

$$C = \log_2 \det(\mathbf{I}_{N_R} + (\rho/N_T)\mathbf{H}\mathbf{H}^H) \text{ bits/sec/Hz} \quad (23)$$

where ρ stands for average SNR per branch, N_T and N_R are the number of transmitter and receiver antennas, respectively, and \mathbf{H} is the channel matrix.

$$\mathbf{H} = \begin{bmatrix} h_{11} & \cdots & h_{N_T 1} \\ \vdots & \ddots & \vdots \\ h_{1 N_R} & \cdots & h_{N_T N_R} \end{bmatrix} \quad (24)$$

where h_{ij} is the channel response from i-th transmitter to j-th receiver antenna. In case of random MIMO channel, this formula becomes

$$C = E_H \log_2 \det(\mathbf{I}_{N_R} + (\rho/N_T)\mathbf{H}\mathbf{H}^H) \text{ bits/sec/Hz} \quad (25)$$

where E_H stands for mathematical expectation over different channel realizations (\mathbf{H}).

In Figure 34, the capacity for $SNR = 20dB$ is calculated for different number of transmit and receive antennas. Independent Rayleigh fading in channels is assumed. From the results it can be observed that with probability of 0.8 the channel capacity will be greater than 4.5bps/Hz for SISO system, 9.6 for 2×2 MIMO, 20.5 for 4×4 MIMO and 42.4 for 8×8 MIMO system. This shows the important capacity gain with MIMO systems. Actually, it can be seen that the capacity grows linearly with the number of antennas. This calculation is valid in scenarios where channel information is available in transmitter and in the receiver and the uniform power allocation in all subchannels [Shi00].

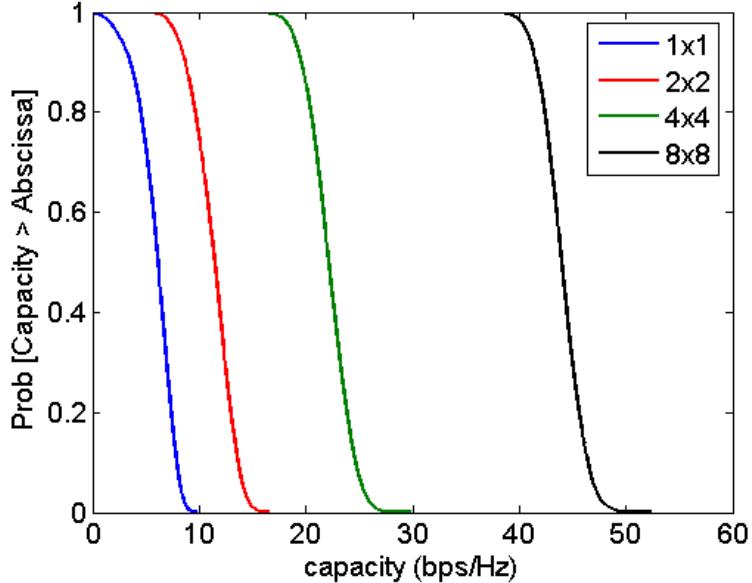


Figure 34: Capacity of MIMO systems in independent Rayleigh fading with $SNR = 20dB$.

>From this short review of some aspects of MIMO systems, it can be seen that a substantial capacity gain can be obtained by incrementing the number of antennas in the communication system. That is precisely the reason why it is one of the topics on which intensive research is conducted. In this chapter some important aspects of MIMO, as possible ways of taking advantage of multiple antennas in cellular, multiuser system are not mentioned. The presented calculus does not take into account the correlation between antenna signals or any type of MIMO RF front-end imperfections. When compared to traditional single antenna RF front-end architecture the imperfections inside the multi-antenna system are stronger and more pronounced, and some new distortions like coupling arise. This is especially important with antennas being so close one to another as is the case in the mobile where, as will be shown in this thesis, the coupling can degrade significantly the system performance. However, the analysis conducted here shows that multiple antennas are one of the key aspects for reaching high data rates offered by future mobile systems.

4 Direct Coupling Model

4.1 Nonlinear Coupling Description

Throughout the years, it has been shown in many scientific areas that nonlinear foundations adapt better to real life situations and offer higher accuracy solutions. However, nonlinear approach is basically avoided at sight, favouring the simpler linear system model with less satisfactory results. This is largely due to the fact that while the first-order, linear behavior is relatively easy to model, the non linear behavior, and especially inverse nonlinear modelling requires careful study and more advanced mathematical methods. This is the main reason why nonlinear engineering is regarded as “too difficult”, “too demanding” and basically confusing.

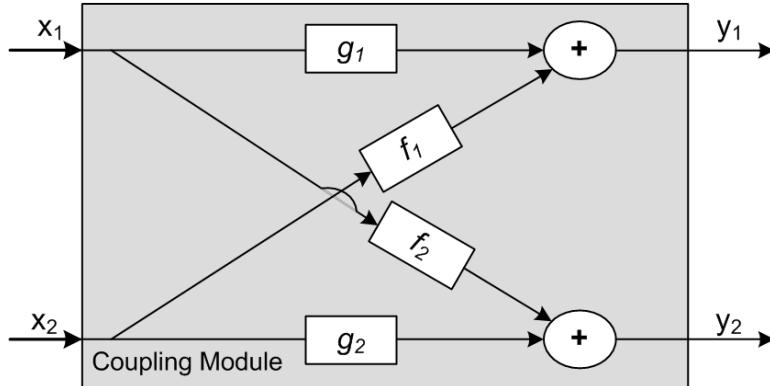


Figure 35: The simplified nonlinear coupling module scheme

The idea that lies in this thesis, beneath multi-antenna RF front-end coupling cancellation is of a multi-functional coupling cancellation module able to separate any two energy signals regardless of their type, their coupling form or their coupling source. Therefore the coupling is intended, as much as possible, to be seen like a simple *black box* with its input and output values, as the one shown in Figure 35. This approach makes the model independent of the actual coupling source. and since it is not limited to multi-antenna systems, the model is of general importance.

The problem of energy coupling can be described with equations:

$$\begin{aligned} y_1 &= g_1(x_1, Dist_1) + f_1(x_2, Coup_{21}) \\ y_2 &= g_2(x_2, Dist_2) + f_2(x_1, Coup_{12}) \end{aligned} \quad (26)$$

where x_1 and x_2 represent the input signals, while y_1 and y_2 the output of coupling module. $Coup_{ij}$ describes the coupling experienced on the signal path of the j -th input generated by the signal corresponding to the i -th input, while the $Dist_i$ specifies the distortion parameters on the corresponding direct signal path. The functions g_1 and g_2 form the direct signal path functions consisting of specific signal distortions characteristic to the regarded systems with only one input and without coupling. The form and complexity of these functions depends entirely on the system environment and the required precision of the system approximation. The functions f_1 and f_2 are denominated direct coupling functions as they describe the distortions experienced when passing through the coupling environment. The amplitude and shape of direct coupling functions change depending on strength of different signal distortions phenomena characteristic to corresponding system. In order to be able to reveal the dominant nonlinear distortion source different nonlinear contributions should be separated and observed individually. However this is too complex to be modelled, and it presents a key factor only when modelling the specific direct system. Since the focus here is on the general inverse nonlinear function, the overall nonlinearity is to be modelled while the source of nonlinearity is irrelevant to the model.

In most practical cases the reasonable assumption for the direct functions is to model them as polynomials and to limit them to third order nonlinear behavior. Hence, inside these functions, different nonlinear interference contributions are modelled with one polynomial of a certain order:

$$\begin{aligned} f_1(x_2, Coup_{21}) &= c_{11}x_2 + c_{12}x_2^2 + c_{13}x_2^3 \\ f_2(x_1, Coup_{12}) &= c_{21}x_1 + c_{22}x_1^2 + c_{23}x_1^3 \end{aligned} \quad (27)$$

with c_{ij} corresponding to the coupling of the j^{th} order experienced on the i^{th} coupling module input. As coupling describes energy absorption/dissipation, the parameters of direct function should physically correspond to the power of the corresponding coupling interference order. The logical approach is to introduce another parameter denominated as C_{ij} given in [dB],

which defines the distortion of the j^{th} order experienced on the i^{th} coupling module input, as a log ratio of the signal and the interference power:

$$C_{ij} = 10 \log \frac{P_j}{P_i} [\text{dB}] \quad (28)$$

with P_i the power of the received signal at input i , P_j the power of the interference of the j^{th} order and C_{ij} the coupling parameter expressed in [dB]. From this equation, the corresponding coupling parameter of the direct function is calculated as:

$$c_{ij}^2 = \frac{P_i}{P_{j0}} 10^{C_{ij}/10} \quad (29)$$

where c_{ij} is the coupling amplitude distortion of the j^{th} order and P_{j0} the *normalized* power of interference of the j^{th} order expressed as:

$$P_{j0} = \frac{1}{N} \sum_{i=1}^N [c_j x^j]^2 \quad (30)$$

where x is the output of I/Q up-converter whose input power is normalized to one and N is the number of signals. The *normalized* power of the interference depends on the form and strength of the received signal and should be calculated separately for each signal type and for each interference order. The exact form of the normalized interference power for DC receiver is analyzed in the second part of this chapter.

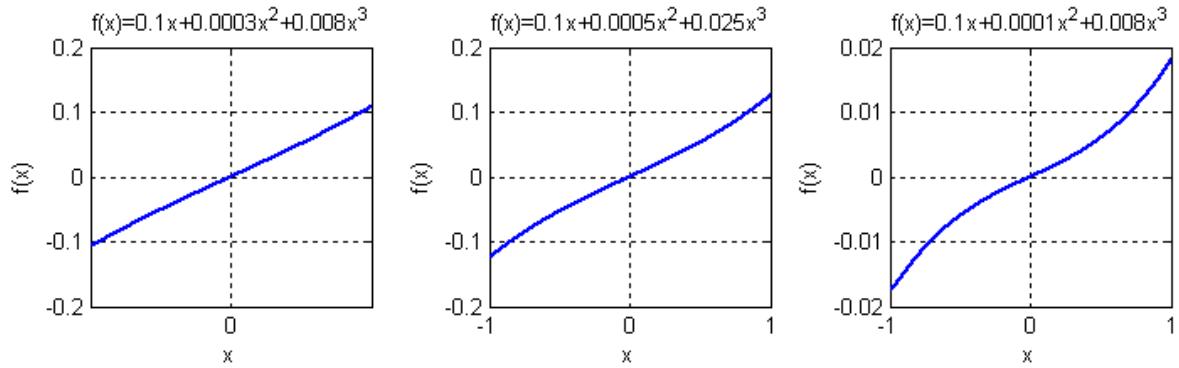


Figure 36: Direct functions with different distortion parameters

According to eq. (27), the direct function is defined with three parameters (c_{i1}, c_{i2}, c_{i3}) which will be given in [dB] as an array of size three [C_{i1}, C_{i2}, C_{i3}], with i defining the direct function. For example, in a case with $P_i = 1$, $P_{10} = 1$, $P_{20} = 2$ and $P_{30} = 4$, the coupling of order $[-20, -30, -10]$ dB is mathematically described with direct function f of the following form:

$$f(x) = 0.1x + 0.0005x^2 + 0.025x^3 \quad (31)$$

The form of direct functions with different nonlinear distortion values (c_{ij}) can be seen on Figure 36. It can be seen that a small level of nonlinearity influences significantly the function characteristics.

Conclusion

In the presented coupling description, the nonlinear coupling distortion is modeled with polynomials limited to the third order. Two direct nonlinear coupling functions are completely defined with six coupling parameters expressed in [dB], the power of the received signals, the power of interference of each order which is established with the signal type. The direct signal path distortion functions, depend on the system environment, and are not defined by any closed mathematical form. This way, the module is generic and can easily be adopted to any coupling environment.

4.2 Multi-antenna RF front-end coupling model

At physical level RF front-end coupling depends on many parameters, starting with the selection of electronic elements, throughout the RF front-end design, up to the working conditions of the system which include system bandwidth, intermediate frequency (IF), channel parameters etc. When designing RF front-end, the commonly used components are basically bipolar and MOS (Metal Oxide Semiconductor) transistors, resistors, capacitors and diodes. All of these components have nonlinear input-output characteristics which regardless of their individual amount all contribute to the global nonlinear behaviour of the RF front-end.

This nonlinear RF front-end coupling model is developed for the direct conversion receiver with implemented I/Q downconverter. The I/Q technology concept is based on complex-valued signals that contain only positive frequencies, referred to as analytic signals. One of these signals is complex oscillator which enables bandpass transmission as positive and negative frequencies are processed independently:

$$e^{j2\pi f_c t} = \cos(2\pi f_c t) + j\sin(2\pi f_c t) \quad (32)$$

In general, a complex-valued baseband signal $x(t) = x_I(t) + jx_Q(t)$, composed of two real-valued messages $x_I(t)$ and $x_Q(t)$ can be transmitted in a real-valued channel by means of this complex oscillator. In this case both signals are transmitted over the same bandwidth resulting in increased spectral efficiency. The sent signal is written as:

$$s(t) = 2\operatorname{Re}[x(t)e^{j2\pi f_c t}] = 2x_I(t)\cos(2\pi f_c t) - 2x_Q(t)\sin(2\pi f_c t) \quad (33)$$

where f_c denotes the central frequency, and $\operatorname{Re}[x(t)]$ refers to the real part of a complex-valued quantity $x(t)$.

In frequency domain, a general complex-valued signal $x(t)$ with two different real valued messages implies no particular spectrum symmetry around the center-frequency f_c of $s(t)$. However, as illustrated in Figure 37a, the spectral components of $s(t)$ around $+f_c$ and $-f_c$ are mirror images of each other, each one of them containing all the necessary information about $s(t)$.

Thus, only one of the above spectral components is required for the complete information recovery. The corresponding mirror component is generally rejected using some kind of bandpass filter to avoid signal overlapping. This is seen more clear if eq. (33) is rewritten using the Euler formula ($\cos(x) = \frac{e^{jx} + e^{-jx}}{2}$; $\sin(x) = \frac{e^{jx} - e^{-jx}}{2j}$):

$$s(t) = 2\operatorname{Re}[x(t)e^{j2\pi f_c t}] = x(t)e^{j2\pi f_c t} + x^*(t)e^{-j2\pi f_c t} \quad (34)$$

In order to select the positive frequency component of a bandpass signal, the most commonly used approach is to simply shift the spectrum of the signal by multiplying it with a complex exponential $e^{-j2\pi f_c t}$. This causes a pure frequency translation, resulting in an I/Q signal pair

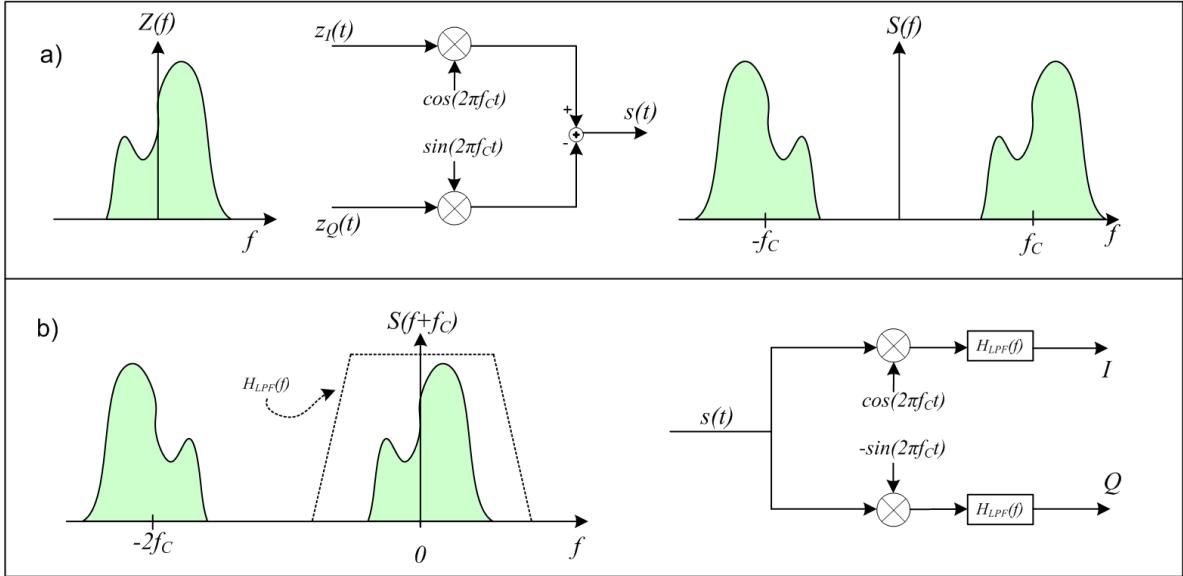


Figure 37: Bandpass signal transmission and single frequency component extraction at the receiver

for which the positive and negative frequency parts of $s(t)$ are shifted around the zero and $-2f_c$, respectively. The induced high frequency component around $-2f_c$ can be attenuated by applying lowpass filtering (LPF) on the I and Q signals. This concept, depicted graphically in Figure 37b, is generally known as the quadrature or I/Q downconversion approach.

4.3 Influence of RF front-end nonlinear coupling on the output of I/Q downconverter

Even though coupling appears inside RF front-end, at the output of I/Q downconverter its effects are altered when LPF is applied. Since coupling components at higher frequencies are rejected the coupling interference at output of I/Q downconverter is partially cancelled. Actually, the coupling in a system based on the I/Q receiver is only influenced with direct interference components which only appear in linear and third order coupling interference. This is shown mathematically in what follows.

If the sent signal of the form as in eq (33) experiences linear coupling of the amplitude c_{21} the

interference signal is of the form:

$$f(x, Coup_{12}) = c_{21} [2x_I(t)\cos\omega_c t - 2x_Q(t)\sin\omega_c t] \quad (35)$$

where x_I and x_Q are the real and imaginary part of the signal constellation, f_c the central frequency, n denotes the symbol position, and $\omega_c = 2\pi f_c$. When this interference signal passes through I/Q downconverter the corresponding I and Q branches are of the following form:

$$\begin{aligned} I_1(t) &= c_{21}\cos\omega_c t [2x_I(t)\cos\omega_c t - 2x_Q(t)\sin\omega_c t] \\ Q_1(t) &= -c_{21}\sin\omega_c t [2x_I(t)\cos\omega_c t - 2x_Q(t)\sin\omega_c t] \end{aligned} \quad (36)$$

which can be rewritten as:

$$\begin{aligned} I_1(t) &= c_{21}x_I(t) + c_{21}x_I(t)\cos 2\omega_c t - c_{21}x_Q(t)\sin 2\omega_c t \\ Q_1(t) &= c_{21}x_Q(t) - c_{21}x_I(t)\sin 2\omega_c t - c_{21}x_Q(t)\cos 2\omega_c t \end{aligned} \quad (37)$$

After the low-pass filtering is implemented, the second and third elements of the both equations are rejected, and the outputs are:

$$\begin{aligned} I_1(t) &= \underset{LPF}{c_{21}x_I(t)} \\ Q_1(t) &= \underset{LPF}{c_{21}x_Q(t)} \end{aligned} \quad (38)$$

When third order coupling is present in the system, the direct coupling function converts into the following form:

$$f(x, Coup_{12}) = c_{23} [2x_I(t)\cos\omega_c t - 2x_Q(t)\sin\omega_c t]^3 \quad (39)$$

while I and Q branches carry the signal of:

$$\begin{aligned} I_3(t) &= c_{23}\cos\omega_c t [2x_I(t)\cos\omega_c t - 2x_Q(t)\sin\omega_c t]^3 \\ Q_3(t) &= -c_{23}\sin\omega_c t [2x_I(t)\cos\omega_c t - 2x_Q(t)\sin\omega_c t]^3 \end{aligned} \quad (40)$$

which are transformed into:

$$\begin{aligned}
I_3(t) = & c_{23}x_I^3(t)[\cos 4\omega_c t + 4\cos 2\omega_c t + 3] - 2c_{23}x_Q^3(t)[1 - \cos 2\omega_c t]\sin 2\omega_c t \\
& - 3c_{23}x_I^2(t)x_Q(t)[1 + \cos 2\omega_c t]\sin 2\omega_c t + 3c_{23}x_I(t)x_Q^2(t)\frac{1}{2}[1 - \cos 4\omega_c t] \\
Q_3(t) = & -2c_{23}x_I^3(t)[1 + \cos 2\omega_c t]\sin 2\omega_c t + c_{23}x_Q^3(t)[\cos 4\omega_c t + 4\cos 2\omega_c t + 3] \\
& + \frac{3}{2}c_{23}x_I^2(t)x_Q(t)[1 - \cos 4\omega_c t] - 3c_{23}x_I(t)x_Q^2(t)[1 - \cos 2\omega_c t]\sin 2\omega_c t
\end{aligned} \tag{41}$$

The frequency composition is seen more clearly in the following equations:

$$\begin{aligned}
I_3(t) = & 3c_{23}x_I^3(t) + \frac{3}{2}c_{23}x_I(t)x_Q^2(t) - [2c_{23}x_Q^3(t) + 3c_{23}x_I^2(t)x_Q(t)]\sin 2\omega_c t \\
& + 4c_{23}x_I^3(t)\cos 2\omega_c t + [c_{23}x_Q^3(t) - \frac{3}{2}c_{23}x_I^2(t)x_Q(t)]\sin 4\omega_c t \\
& + [c_{23}x_I^3(t) - \frac{3}{2}c_{23}x_I(t)x_Q^2(t)]\cos 4\omega_c t \\
Q_3(t) = & 3c_{23}x_Q^3(t) + \frac{3}{2}c_{23}x_I^2(t)x_Q(t) - [2c_{23}x_I^3(t) + 3c_{23}x_I(t)x_Q^2(t)]\sin 2\omega_c t \\
& + 4c_{23}x_Q^3(t)\cos 2\omega_c t - [c_{23}x_I^3(t) - \frac{3}{2}c_{23}x_I(t)x_Q^2(t)]\sin 4\omega_c t \\
& + [c_{23}x_Q^3(t) - \frac{3}{2}c_{23}x_I^2(t)x_Q(t)]\cos 4\omega_c t
\end{aligned} \tag{42}$$

When these two interference signals are passed through LPF the resulting output is:

$$\begin{aligned}
I_3(t) &= \overset{LPF}{3c_{23}x_I^3(t) + \frac{3}{2}c_{23}x_I(t)x_Q^2(t)} \\
Q_3(t) &= \overset{LPF}{3c_{23}x_Q^3(t) + \frac{3}{2}c_{23}x_I^2(t)x_Q(t)}
\end{aligned} \tag{43}$$

On the other hand, the second order coupling interference signal is completely rejected by LPF and in this system has no influence on the output of I/Q downconverter. Similar to previous analysis, if the direct coupling function consists of only second order interference the received

interference signals inside the I branch can be expressed as:

$$\begin{aligned} I_2(t) &= c_{22}\cos\omega_c t [2x_I(t)\cos\omega_c t - 2x_Q(t)\sin\omega_c t]^2 \\ Q_2(t) &= -c_{22}\sin\omega_c t [2x_I(t)\cos\omega_c t - 2x_Q(t)\sin\omega_c t]^2 \end{aligned} \quad (44)$$

which before passing through the LPF can be written as:

$$\begin{aligned} I_2(t) &= \left[3c_{22}x_I^2(t) + c_{22}x_Q^2(t) \right] \cos\omega_c t - c_{22}x_I(t)x_Q(t)\sin\omega_c t \\ &\quad + \left[c_{22}x_I^2(t) - c_{22}x_Q^2(t) \right] \cos 3\omega_c t + c_{22}x_I(t)x_Q(t)\sin 3\omega_c t \\ Q_2(t) &= - \left[c_{22}x_I^2(t) + 3c_{22}x_Q^2(t) \right] \sin\omega_c t + c_{22}x_I(t)x_Q(t)\cos\omega_c t \\ &\quad - \left[c_{22}x_I^2(t) - c_{22}x_Q^2(t) \right] \sin 3\omega_c t - c_{22}x_I(t)x_Q(t)\cos 3\omega_c t \end{aligned} \quad (45)$$

When these two signals are passed through LPF, the output is zero as the whole signal is found at the higher frequencies. In other words, it is mathematically confirmed that the coupling of second order has no influence in DCR receivers as all generated components are rejected in LPF. This phenomena is mathematically demonstrated for the ideal case but the same holds for any amplitude and phase imbalance inside the DCR receiver.

$$\begin{aligned} I_2(n) &\stackrel{LPF}{=} 0 \\ Q_2(n) &\stackrel{LPF}{=} 0 \end{aligned} \quad (46)$$

Eventhough second order distortion has no influence on the multi antenna RF Front-end model of DCR receiver, the distortions physically exists. There is no necessity to eliminate it, but it should be included in direct coupling functions expressed as:

$$\begin{aligned} f_1(x_2, Coup_{21}) &= c_{11}x_2 + c_{12}x_2^2 + c_{13}x_2^3 \\ f_2(x_1, Coup_{12}) &= c_{21}x_1 + c_{22}x_1^2 + c_{23}x_1^3 \end{aligned} \quad (47)$$

As a consequence direct coupling function of RF front-end coupling module is defined with three parameters as an array of size three $[C_{i1}, C_{i2}, C_{i3}]$.

4.4 The power of the coupling interference inside RF front-end

As already mentioned, the nonlinear coupling parameters are limited to the third order non-linearity factor as this order is considered to be a good approximation of the physical coupling phenomena. According to eq. (28) the direct coupling function is defined with three parameters $[C_{i1}, C_{i2}, C_{i3}]$ expressed in $[dB]$, where i defines the antenna. The relation between interference amplitude in the direct coupling function and the interference power at physical level is given by:

$$C_{ij} = 10 \log \frac{c_{ij}^2 P_{j0}}{P_i} [dB] \Leftrightarrow c_{ij}^2 = \frac{P_i}{P_{j0}} 10^{C_{ij}/10} \quad (48)$$

with P_i the power of the received signal at input i , P_{j0} the normalized power of the interference of the j^{th} order and C_{ij} the coupling parameter expressed in $[dB]$ for the i^{th} antenna and j^{th} coupling order. Since signal power is defined as the mean power of all symbols over the corresponding symbol length, the normalized power of linear distortion can be calculated as:

$$P_{10} = \frac{1}{N} \sum_{n=1}^N \left(\frac{1}{T} \right) \int_0^T (c_j x_2^j)^2 dt \quad (49)$$

where T is the symbol duration, $\omega_c T = 2\pi$, and N is the number of received signals. The analyzed signal x_2 is the interference generated on second antenna and experienced on the first one consisting of in-phase and in-quadrature signals $x_2 = x_{2I} + x_{2Q}$. To make the indexing simpler, since the same process appears at both antennas the index i corresponding to the number of antenna is cast aside. Hence, $x_2 = x = x_I + x_Q$, and the same holds for numerical coupling parameters $c_{11} = c_{21} = c_1$, $c_{13} = c_{23} = c_3$ as for the ones expressed in $[dB]$ $C_{11} = C_{21} = C_1$, $C_{13} = C_{23} = C_3$

$$P_{10} = \frac{1}{N} \sum_{n=1}^N \left(\frac{1}{T} \right) \int_0^T (c_1 2x_I(n) \cos \omega_c t - c_1 2x_Q(n) \sin \omega_c t)^2 dt \quad (50)$$

with $x_I(n) = x_I(t - nT)$ and $x_Q(n) = x_Q(t - nT)$ where n is the symbol index. The total number of symbols N is in this case equal to the number of symbols of the corresponding constellation as x_I, x_Q range over the whole constellation.

$$P_{10} = \frac{1}{N} \sum_{n=1}^N c_1^2 \left(\frac{1}{T} \right) \left\{ \left[(2x_I(n))^2 \left(\frac{1}{2}t + \frac{1}{4\omega_c} \sin 2\omega_c t \right) \right]_0^T + \right. \\ \left. + \left[(2x_Q(n))^2 \left(\frac{1}{2}t - \frac{1}{4\omega_c} \sin 2\omega_c t \right) \right]_0^T \right. \\ \left. - \left[(2x_I(n))(2x_Q(n)) \frac{1}{4\omega_c} \cos 2\omega_c t \right]_0^T \right\} \quad (51)$$

As the duration of the symbol is equal to one period of the complex local oscillator the following calculation holds:

$$P_{10} = \frac{1}{2} c_1^2 \frac{1}{N} \sum_{n=1}^N \left[(2x_I(n))^2 + (2x_Q(n))^2 \right] \quad (52)$$

Numerically, this also represents the power of the received signal ($P_i = P_{10}$) and is equal to the power of received signal in ideal conditions prior to the LPF. The power of the third order interference is obtained with similar calculation:

$$P_{30} = \frac{1}{N} \sum_{n=1}^N \left(\frac{1}{T} \right) \int_0^T \left(c_2 (2x_I(n) \cos \omega_c t - c_1 2x_Q(n) \sin \omega_c t)^3 \right)^2 dt \quad (53)$$

which can be rewritten as:

$$\begin{aligned}
P_{30} = & \frac{1}{N} c_3^2 \sum_{n=1}^N \left(\frac{1}{T} \right) \\
& \left\{ (2x_I(n))^6 \left[\frac{1}{6\omega_c} \cos^5 \omega_c t \sin \omega_c t + \frac{5}{6} \left(\frac{3}{8}t + \frac{1}{4\omega_c} \sin 2\omega_c t + \frac{1}{32\omega_c} \sin 4\omega_c t \right) \right]_0^T \right. \\
& - 6 (2x_I(n))^5 (2x_Q(n)) \left[\frac{-1}{6\omega_c} \cos^6 \omega_c t \right]^T \\
& - 20 (2x_I(n))^3 (2x_Q(n))^3 \frac{1}{8} \left[\frac{-1}{2\omega_c} \cos 2\omega_c t + \frac{1}{6\omega_c} \cos^3 2\omega_c t \right]_0^T \\
& + 15 (2x_I(n))^2 (2x_Q(n))^4 \frac{1}{8} \left[\frac{1}{2}t - \frac{1}{8\omega_c} \sin 4\omega_c t - \frac{1}{6\omega_c} \sin^3 2\omega_c t \right]_0^T \\
& - 6 (2x_I(n)) (2x_Q(n))^5 \left[\frac{-1}{6\omega_c} \sin^6 \omega_c t \right]_0^T \\
& \left. + (2x_Q(n))^6 \left[\frac{-1}{6\omega_c} \sin^5 \omega_c t \cos \omega_c t + \frac{5}{6} \left(\frac{3}{8}t - \frac{1}{4\omega_c} \sin 2\omega_c t + \frac{1}{32\omega_c} \sin 4\omega_c t \right) \right]_0^T \right\} \\
\end{aligned} \tag{54}$$

Again, since $T = 2\pi/\omega_c$, the following calculation can be made:

$$P_{30} = \frac{5}{16} c_3^2 \frac{1}{N} \sum_{n=1}^N \left\{ (2x_I(n))^6 + (2x_Q(n))^6 + 3 (2x_I(n))^2 (2x_Q(n))^2 \left[(2x_I(n))^2 + (2x_Q(n))^2 \right] \right\} \tag{55}$$

Finally, the power of second order interference which has no influence the signal at the output of I/Q demodulator is expressed as:

$$P_{20} = \frac{1}{N} \sum_{n=1}^N \left(\frac{1}{T} \right) \int_0^T \left(c_2 (2x_I(n) \cos \omega_c t - c_1 2x_Q(n) \sin \omega_c t)^2 \right)^2 dt \tag{56}$$

$$P_{20} = \frac{3}{8} c_2^2 \frac{1}{N} \sum_{n=1}^N \left[(2x_I(n))^4 + (2x_Q(n))^4 + 2 (2x_I(n))^2 (2x_Q(n))^2 \right] \tag{57}$$

The eq.(57) confirms the existence of square distortions, which has to be taken into account inside the decoupling module.

According to the equations (52),(55) and (57), numerical calculation of different signal constellations can be made.

	P_{10}, P_i	P_{20}	P_{30}
Binary	2.00	6.00	20.00
QPSK	2.00	6.00	20.00
8-PSK	2.00	6.00	20.00
16QAM	2.00	7.92	39.20
64QAM	2.00	7.92	39.20

Table 2: The unit power of three interference orders for different constellations

The obtained values, presented in Table 2, demonstrate that the influence of nonlinear coupling raises as the constellation consisting of signals with different amplitudes is introduced. These numbers can be used for simulation purposes only when the power of signal in the transmitter (P_s) is normalized:

$$P_s = \frac{1}{N} \sum_{n=1}^N [x_I^2(n) + x_Q^2(n)] = 1 \quad (58)$$

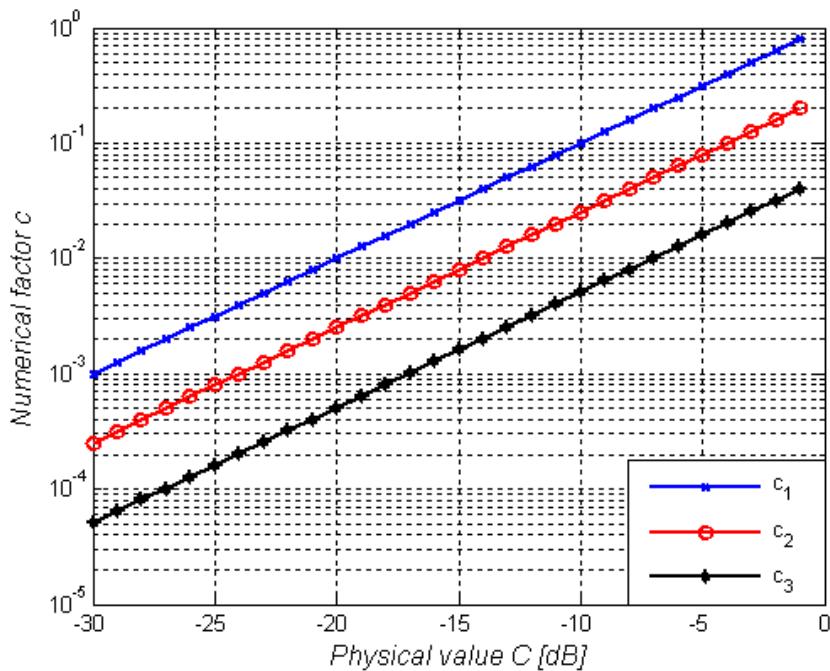


Figure 38: Mathematical factor c_i expressed as a function of physical factor C_i

The relationship of physical (C_i) and mathematical factors (c_i) when $P_s = 1$ is presented

in Figure 38. It can be seen that for a given physical interference power factor expressed in dB the mathematical factor, used in direct function definition, has higher value for linear distortion than for the cubic one, while the second order interference factor value is situated between those two. As in the following the nonlinearity values will be specified with physical factors, it should be stressed that the same strength of coupling interference corresponds to lower numerical values for higher nonlinearity orders.

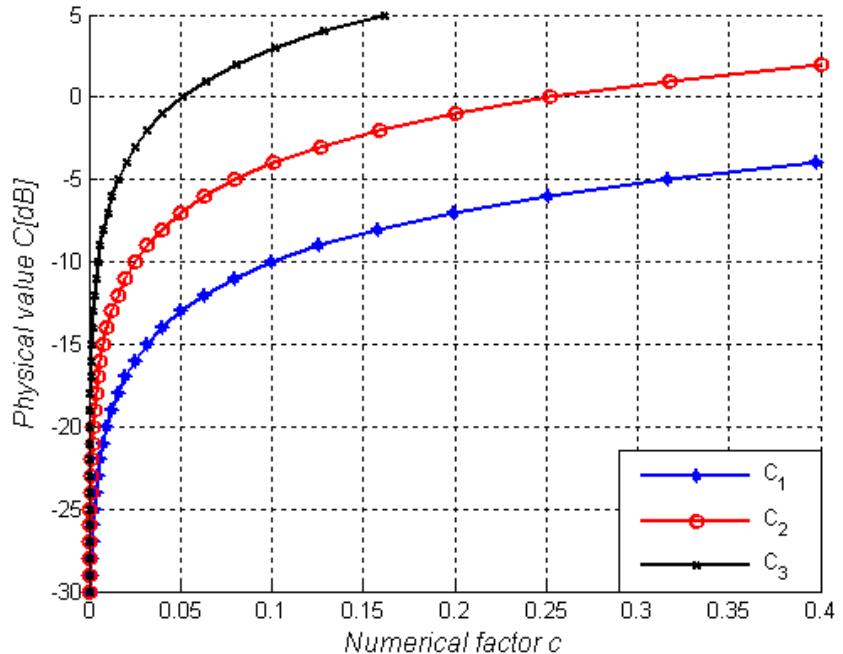


Figure 39: The physical factor C_i expressed as a function of mathematical factor c_i

Actually, if the same numerical factor is selected, for example 10^{-2} , it corresponds to physical values of approximately $-20, -14$, and -7 dB for linear, quadratic and cubic interference. As the lines are parallel, it can be seen that the difference is 6 and 13 dB between the first and second and third order respectively. In linear scale, power difference becomes more striking as seen in Figure 39 where physical factor C_i is expressed as a function of mathematical factor c_i . This power difference confirms the importance of nonlinear coupling analysis.

5 Inverse coupling function

The straight forward solution to the RF coupling cancellation is the use of its inverse function. In this chapter different approaches to the search of inverse coupling function will be introduced and analyzed. It will be demonstrated that the generic form of analytical inverse function can not be derived as the form changes with different sent signals. Furthermore, it will be shown that the inverse solution might not exist at all, or it may offer more than one mathematical solution to same set of data. Since this is unacceptable, the Sturm chain is used for single solution confirmation, and afterwards Newton numerical approximation method is applied for reaching the inverse function solution on symbol to symbol bases. As this approach requires huge amount of processing power, heuristic approximation method in combination with calibration process is introduced as the inverse coupling solution. This method is chosen as it is the only one guaranteeing the optimal solution if enough number of iterations is available. Since this method converges too slowly it is not considered as an option for the search algorithm inside the decoupling module, but it is used for setting correctly the required system parameters like the calibration data size, or the correct cost function form inside this chapter. Furthermore, the decoupling results obtained with this approach are taken as top accuracy reference for other decoupling module approximation methods.

5.1 Analytical solution

Finding the analytical inverse function of a high order nonlinear process is complicated and requires deep mathematical combined with a lot of patience. In order to simplify as much as possible the analytical analysis, the direct distortion functions in eq. (26):

$$\begin{aligned} y_1 &= g_1(x_1, Dist_1) + f_1(x_2, Coup_{21}) \\ y_2 &= g_2(x_2, Dist_2) + f_2(x_1, Coup_{12}) \end{aligned} \tag{59}$$

are set to describe the ideal case, meaning that only the signals, without any imperfection, are passed:

$$\begin{aligned} g_1(x_1, Dist_1) &= x_1 \\ g_2(x_2, Dist_2) &= x_2 \end{aligned} \tag{60}$$

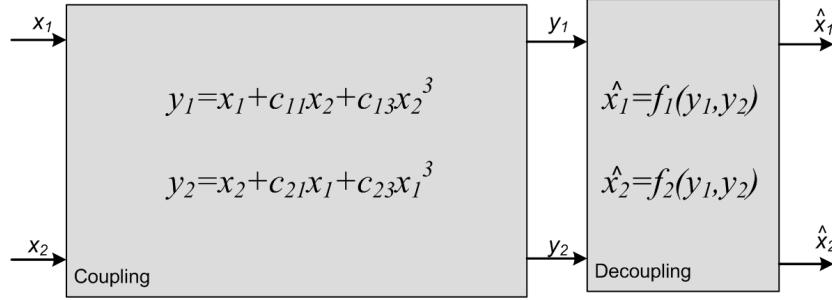


Figure 40: Inverse RF front-end nonlinear coupling function

With the direct coupling functions defining the RF front-end coupling set as in (47):

$$f_1(x_2, Coup_{21}) = c_{11}x_2 + c_{12}x_2^2 + c_{13}x_2^3 \quad (61)$$

$$f_2(x_1, Coup_{12}) = c_{21}x_1 + c_{22}x_1^2 + c_{23}x_1^3$$

the analytical system, shown in Figure 40, can be described by following eq. (62) and (63):

$$y_1 = x_1 + c_{11}x_2 + c_{12}x_2^2 + c_{13}x_2^3 \quad (62)$$

$$y_2 = x_2 + c_{21}x_1 + c_{22}x_1^2 + c_{23}x_1^3 \quad (63)$$

Taking into account that at the output of decoupling module, \hat{x}_1 and \hat{x}_2 , should theoretically be equal to x_1 and x_2 the nonlinear functions take the form of:

$$\hat{x}_1 = y_1 - c_{11}\hat{x}_2 - c_{12}\hat{x}_2^2 - c_{13}\hat{x}_2^3 \quad (64)$$

$$\hat{x}_2 = y_2 - c_{21}\hat{x}_1 - c_{22}\hat{x}_1^2 - c_{23}\hat{x}_1^3 \quad (65)$$

Assuming that all distortion parameters c_{ij} are a priori known, a problem of two nonlinear equations with two unknown variables has arisen. In this case, the simplest possible algorithm for the nonlinear coupling cancellation would be to find the direct analytical expression for \hat{x}_1 and \hat{x}_2 . By extracting \hat{x}_2 from eq. (65) and introducing it in eq. (64) the problem is reduced

to single variable:

$$\begin{aligned}\hat{x}_1 = & y_1 - c_{11} [y_2 - c_{21}\hat{x}_1 - c_{12}\hat{x}_2^2 - c_{23}\hat{x}_1^3] \\ & -c_{12} [y_2 - c_{21}\hat{x}_1 - c_{22}\hat{x}_1^2 - c_{23}\hat{x}_1^3]^2 \\ & -c_{13} [y_2 - c_{21}\hat{x}_1 - c_{22}\hat{x}_1^2 - c_{23}\hat{x}_1^3]^3\end{aligned}\tag{66}$$

Afterwards, this equation can be rewritten as a nonlinear problem of the ninth order as:

$$\sum_{i=0}^{i=9} k_i \hat{x}_1^i = 0\tag{67}$$

with:

$$\begin{aligned}k_0 &= y_1 - c_{11}y_2 - c_{12}y_2^2 - c_{13}y_2^3 \\ k_1 &= c_{11}c_{21} + 2c_{12}c_{21}y_2 + 3c_{13}c_{21}y_2^2 - 1 \\ k_2 &= c_{11}c_{22} - c_{12}c_{21}^2 + 2c_{12}c_{22}y_2 - 3c_{13}c_{21}^2y_2 + 3c_{13}c_{22}y_2^2 \\ k_3 &= c_{11}c_{23} - 2c_{12}c_{21}c_{22} + 2c_{12}c_{23}y_2 + c_{13}c_{21}^3 - 6c_{13}c_{21}c_{22}y_2 + 3c_{13}c_{23}y_2^2 \\ k_4 &= -2c_{12}c_{21}c_{23} - c_{12}c_{22}^2 + 3c_{13}c_{21}^2c_{22} - 3c_{13}c_{22}^2y_2 - 6c_{21}c_{23}c_{13}y_2 \\ k_5 &= -2c_{12}c_{22}c_{23} + 3c_{13}c_{21}c_{22}^2 + 3c_{13}c_{21}^2c_{23} - 6c_{13}c_{22}c_{23}y_2 \\ k_6 &= -c_{12}c_{23}^2 + 6c_{13}c_{21}c_{22}c_{23} + c_{13}c_{22}^3 - 3c_{13}c_{23}^2y_2 \\ k_7 &= 3c_{13}c_{21}c_{23}^2 + 3c_{13}c_{22}^2c_{23} \quad k_8 = 3c_{13}c_{22}c_{23}^2 \quad k_9 = c_{13}c_{23}^3\end{aligned}$$

Mathematically, the eq.(67) is an algebra equation of the odd order with real coefficients. Due to the odd order the equation possesses at least one real solution, and since all coefficients are real the complex solutions can only appear in complex conjugate pairs. Theoretically, with each input pair (y_1, y_2) , this kind of equations can have up to nine different real solutions satisfying the main equation. However, even though all nine solutions might correspond to signal constellation domain, according to signal theory only one out of these nine different solutions

is the correct one, and all other solutions only introduce transmission errors. This makes the unique solution of the inverse functions of the highest importance as no ambiguousness is allowed. Thus a procedure that determines the number of unique real roots of a polynomial, named after its discoverer Jacques Charles François Sturm, and commonly known as Sturm's theorem is introduced [Edw64].

According to Sturm theorem the number of real roots of the equation $P(x) = 0$, situated between the two real numbers $\langle a, b \rangle$, where a and b are not the solution of the equation, can be established in three steps:

1. If polynomial $P(x)$ is of the following form:

$$P(x) = (x - \alpha)^k (x - \beta)^l (x - \gamma)^m \dots \quad (68)$$

it is transformed into:

$$P(x) = Q(x)T(x) \quad (69)$$

where:

$$Q(x) = (x - \alpha)^{k-1} (x - \beta)^{l-1} (x - \gamma)^{m-1} \dots \quad (70)$$

$$T(x) = (x - \alpha) (x - \beta) (x - \gamma) \dots \quad (71)$$

$Q(x)$ is obtained as the greatest common divisor of the $P(x)$ and its derivation $P'(x)$, while $T(x)$ is formed by dividing $P(x)$ with $Q(x)$. Eventually, all roots of the polynomial $T(x)$ are also the solutions to the polynomial $P(x)$. The only difference is that they are acquired out of a square-free polynomial $T(x)$ as all repeated roots are eliminated inside $Q(x)$.

2. A Sturm chain is constructed from a square-free polynomial $T(x)$:

$$T(x), T'(x), T_1(x), T_2(x), \dots, T_m = \text{const.} \quad (72)$$

where $T'(x)$ is the derivation of $T(x)$, $T_1(x)$ is the residual of the division of $T(x)$ with $T'(x)$ taken with the opposite sign. $T_2(x)$ is the residual of the division of $T'(x)$ with $T_1(x)$ taken with the opposite sign and so on. The last residual is constant and is denoted with T_m .

3. The number of sign changes inside a Sturm chain denominated A and B is calculated

for both border values a and b respectively:

$$\begin{aligned} A : & T(a), T'(a), T_1(a), T_2(a), \dots, T_m \\ B : & T(b), T'(b), T_1(b), T_2(b), \dots, T_m \end{aligned} \quad (73)$$

Finally, the difference $(A - B)$ is equal to the number of different real roots of the equation $P(x)$ inside the interval $\langle a, b \rangle$. This way, simply by counting the sign changes in the Sturm chain, the number of distinct real roots of a polynomial is obtained.

Using this method the existence of the unique inverse solution of the coupling functions, defined by equations (62) and (63), can be analyzed. This is shown in Figure 41 where the existence of the unique root inside two different border intervals is presented. The unique root existence of x_1 is depicted in blank, while the signals are irreversible mixed, thus making the inverse operation impossible, in gray area. The analysis is made for different sets of coupling coefficients (c_1, c_3) with $c_2 = 0$, and input signals set to $x_1 = 0.1$ and $x_2 = -0.7$. The difference between figures a) and b) is in the root search space. In the figure on the left the root search space of x_1 is defined inside the interval $[-2, 2]$, which is quite larger than interval $[0, 0.3]$ used for obtaining the figure on the right. Since the root search space is built around the solution ($x_1 = 0.1$), inside a larger space the unique solution can be obtained in smaller number of cases, while inside a reduces space the possibility of finding the unique solution increases. Hence, restricting the root space as much as possible around the solution of interest boosts

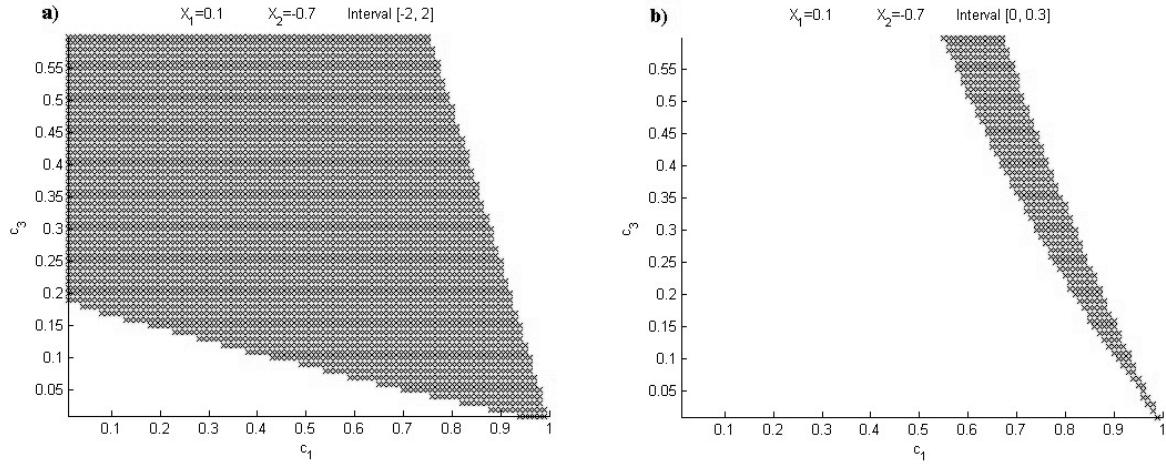


Figure 41: The loss of unique solution for different root spaces with $c_2 = 0$

the existence of the unique inverse function. However, the root search is strictly defined with the signal constellation and can not be restricted further. Thus for the normalized 64QAM signal the root search space is defined inside the interval $[-1.09, 1.09]$ and for 16QAM the roots lie inside $[-0.95, 0.95]$. Similar results are obtained in Figure 42 for different sets of coupling coefficients (c_1, c_3) with $c_2 = 0.3$. The unique solution space again changes as the search space is narrowed, with the difference that in this case isolated solutions exists for several (c_1, c_3) sets of coupling coefficients. For example, $(c_1 = 0.6, c_3 = 0.05)$ offers unique solution when x_1 is searched inside the interval $[-2, 2]$ and $(c_1 = 0.3, c_3 = 0.1)$ offers unique solution when x_1 is searched inside the interval $[0, 0.3]$.

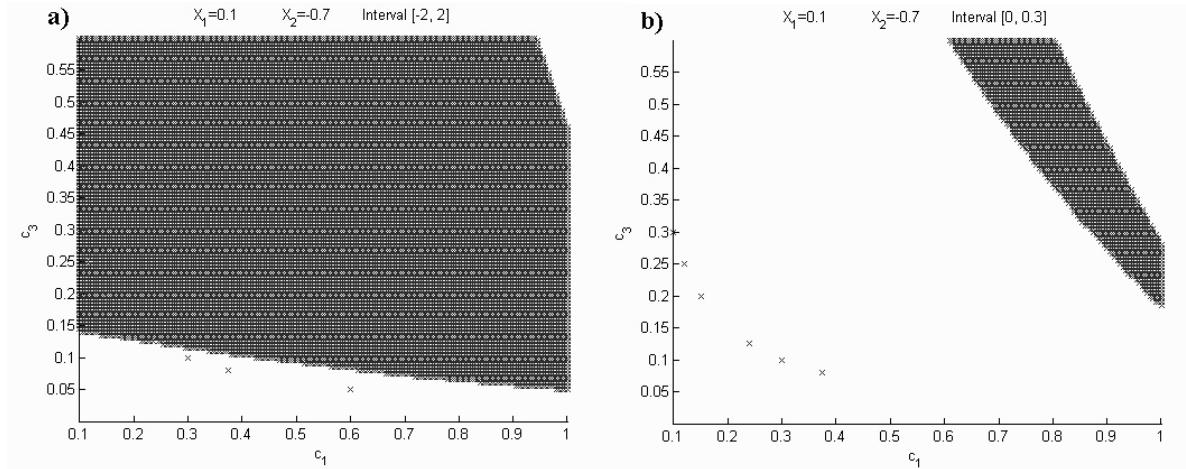


Figure 42: The loss of unique solution for different root spaces with $c_2 = 0.3$

Transforming the mathematical coefficients c_1, c_2 and c_3 into physical values C_1, C_2 and C_3 according to eq. (48) the Sturm's theorem can be applied to nonlinear RF front-end coupling of normalized 64QAM signals. The Figure in 43a is obtained for root interval set to $[-1.09, 1.09]$, C_2 decreased to $C_2 = -100dB$, and input signals $x_1 = 0.1, x_2 = -0.7$. It can be seen that the unique inverse solution exists in majority of situations but still many realistic cases do not have the just one solution, or have no real solution inside the required interval. Furthermore, comparing this figure to Figure 43b, which shows the same system situation only with different input signals set to $x_1 = 1.08$ and $x_2 = -1.08$, the change in operating conditions of the unique solutions can be noticed. Again, inside gray area the unique solution exists, while inside white area inverse function has no unique solution. It is clearly seen that on the left figure the unique solution is almost completely lost for C_3 between $1dB$ and $6dB$ while this is not the case in the configuration depicted on the right. This means that the existence of the unique solution

not only depends on the coupling parameters and constellation root space, but also on the input signal pair (x_1, x_2) .

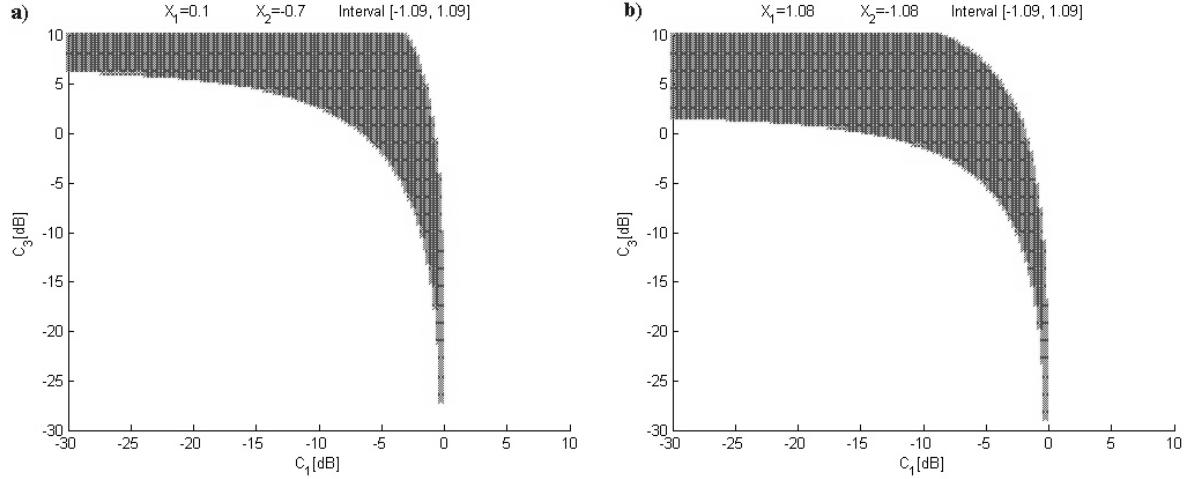


Figure 43: The loss of unique solution for different input signals with $C_2 = -100dB$

The area of the unique solution existence changes when C_3 or C_1 are kept constant as shown in Figure 44a and 44b respectively. The image on the left is obtained for root interval set to $[-1.09, 1.09]$, C_3 is kept constant on $C_3 = -15dB$, and input signals $x_1 = 0.1, x_2 = -0.7$. The loss of the unique solution is presented in two separated areas covering mainly the area of strong coupling with C_2 higher than $5dB$ and C_1 higher than $0dB$. The image on the right proceeds from different sets of (C_2, C_3) coefficients with $C_1 = -5dB$, root interval kept to $[-1.09, 1.09]$,

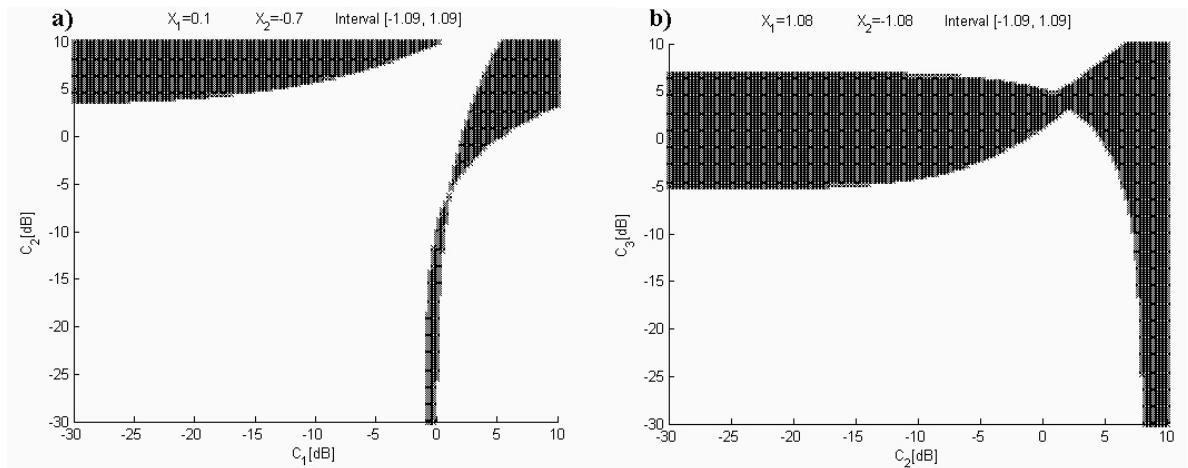


Figure 44: The loss of unique solution with $C_3 = -15dB$ in a) and $C_1 = -5dB$ in b)

and input signals $x_1 = 1.08$, $x_2 = -1.08$. The loss of unique solution is presented in one compact area for all values of C_2 with C_3 kept between approximately $[-5, 5]dB$.

The differences between the last four figures show that the existence of the unique solution changes significantly if only a single system parameter is altered. Since the unique solution area changes with every input pair, it is impossible to construct one analytical solution to the inverse coupling problem, which would offer unique solution to all inputs. However, there exists the possibility of finding the solution for each input pair using numerical methods for approximation of real solutions. Actually, this is the approach commonly used for solving the higher order nonlinear equations like the one in eq. (67). One of the popular and easy to implement methods is known as Newton method [Den83] which iteratively produces a sequence of approximations towards the root of equation. The rate of convergence to the root is quadratic and the starting point of the search is based on an initial guess which is reasonably close to the true root (this can be any point inside the signal constellation domain). If the function has multiple roots, the Newton method will converge to the closest one and will present it as the solution without any further information.

Combining Sturm functions, used for the confirmation of the unique solution, with Newton method applied for the root search the cancellation of the coupling interference can be analyzed for each received signal pair (y_1, y_2) . The result of this approach can be seen in Figure 45 where the domain of y_i is constructed out of the range of $[-1.5, 1.5]$ with 0.025 density. The domain for the estimated solution is chosen to be equal to the normalized 64QAM constellation domain, that is approximately $[-1.09, 1.09]$. Then for each pair of (y_1, y_2) Sturm functions are developed using coupling distortion parameters $[C_1, C_2, C_3]$ for x_1 and x_2 separately. This part of the process checks if the (y_1, y_2) pair produces the unique solution inside the constellation domain under the corresponding distortion parameters. If the unique solution exists for x_1 as well as for x_2 the Newton method is launched for each variable x_1, x_2 from the starting point of $x_0 = 1.0$ with the maximum number of consecutive iterations set to 10. The existence of the unique solution for the pair (x_1, x_2) is traced with a blue dot in the three images on the left inside Figure 45. From the mathematical point of view, under lower coupling conditions ($[-10, -10, -20]dB$), the inverse unique solution exist for any pair of sent data (x_1, x_2) . This is seen in Figure 45a where nonlinear coupling has not irreversibly destroyed the signal information and the whole constellation domain of two independent signal sources can be retrieved. However, as coupling increases the mathematical conditions also change and Sturm/Newthon method cannot guarantee the estimation of the correct sent signal in whole

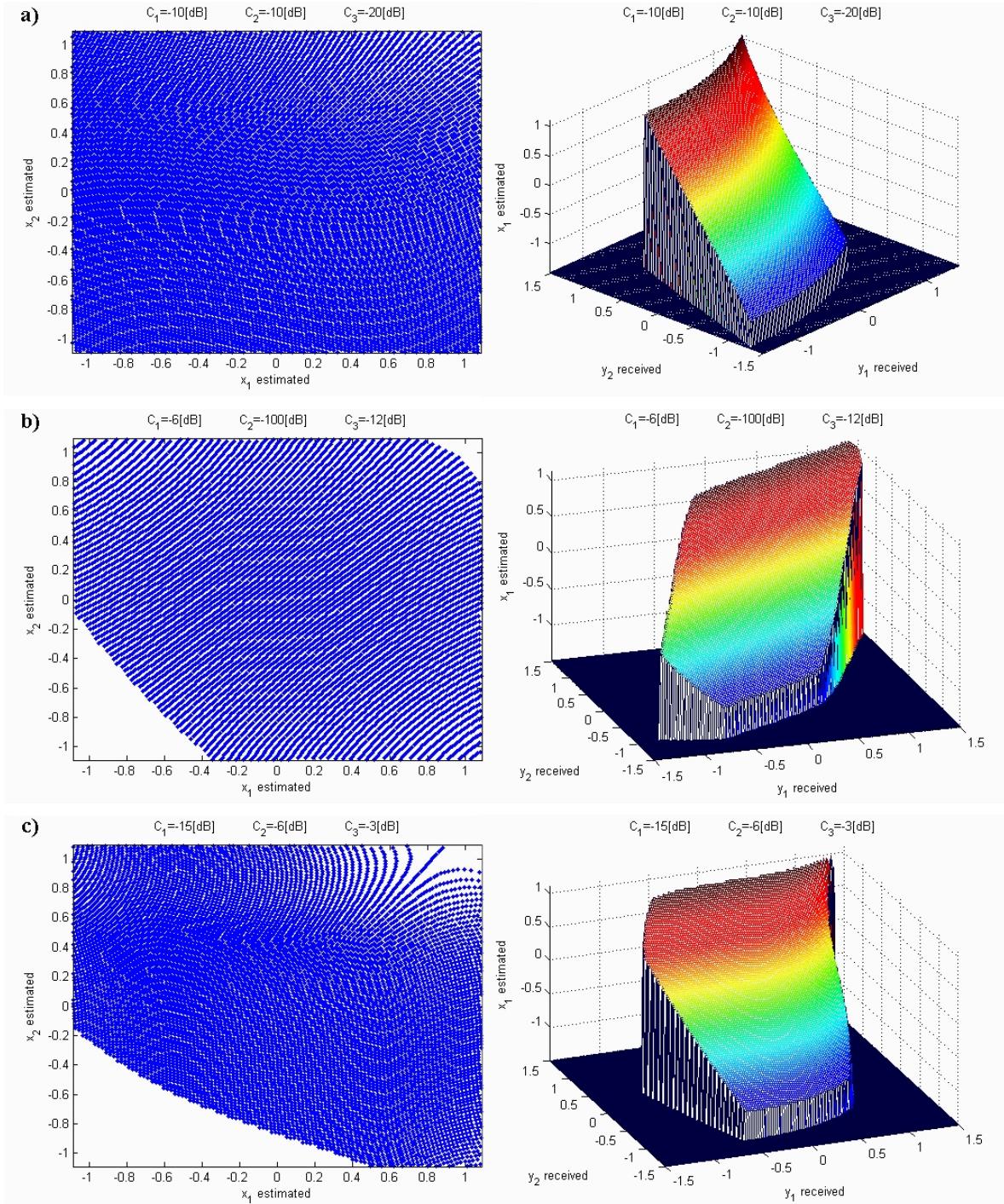


Figure 45: Using Sturm functions and Newton method for the unique solution search

constellation area, as is shown in Figure 45a and 45b. With coupling distortion parameters set to $[-6, -100, -12]dB$, the large blank triangle at lower left corner and a small triangle at upper right corner mark the areas without unique solution. This is why the image on the right appears horizontally sliced as the unique solution can be found only apart from that border. Similar situation is encountered with distortion parameters $[-15, -6, -3]dB$ where one compact area without unique solution in the lower left part of the image is presented while upper right preserves the unique solution only in narrow corridors. Comparing the three images on the right in Figure 45 we can see the change in the range of the functions describing coupling process as the surface defining the inverse function changes from smooth to wave-like area which additionally complicates the inverse function search.

Conclusion

From the mathematical point of view, the existence of the inverse function of simplified coupling function depends on all involved parameters $(x_1, x_2, C_{11}, C_{12}, C_{13}, C_{21}, C_{22}, C_{23})$, together with the signal constellation definition which determines the root space. Using the Sturm's method, it has been shown that the existence of one analytically expressed inverse function supporting all input pairs (x_1, x_2) is extremely low, which makes finding the global analytical solution practically impossible. It should also be emphasized that this ninth order problem was obtained without direct signal distortion and with coupling of only linear and cube order. Hence the inverse function can only get more complex. Apparently, the Sturm-Newton combination can numerically offer good inverse solution to the nonlinear coupling problem but only under lower coupling conditions. However, the drawback of all numerical approximations is that the complete algorithm has to be executed for each point making it rather power consuming. Further on, the sequence is known to converge only if the initial approximation is close enough to the root but no guarantee of the number of iterations can be made. Thus their application inside coupling cancellation module of the terminal phone is computationally not a reliable option.

5.2 Surface fitting approximations

In general, the optimization methods can either be exact or approximated. The first ones perform a systematic search for optimal solution under some criteria while the approximated

methods concentrate on finding a good enough approximation of the global minimum in a large search space, within a finite and limited number of trials. The shortcoming of the approximate methods, commonly termed heuristics, is that no theoretical guarantee for finding the optimal, or even feasible solutions, is provided. On the other hand the solution is not ambiguous and once the inverse surface is well defined only a small portion of the algorithm has to be executed to obtain the desired solution. This class of methods have been used in combinatorial optimization for finding optimal or near-optimal solution for more than a few decades. They have successfully been used for problems like the travelling salesman problem, school scheduling, vehicle routing etc. Since coupling problem is mathematically well described with signal domain, function range and function dependency strictly defined, it could be approached using some type of maximum likelihood model. Thus, the pursuit for analytical solution is avoided and the attention centers to obtaining inverse coupling function with mathematical surface approximations.

Surface fitting approximations are based on finding a surface that matches a series of constraints by minimizing the difference between reference and constructed or inverse surface. In other words, a set of predefined points, denominated calibration data, are available to the observed process (in our case coupling) forming the reference surface. The corresponding obtained values form an inverse function domain which is used for building the approximated inverse surface. Eventually, point-to-point dependency of submitted-obtained, or sent-received, calibration data is used to define an approximate inverse higher level dependency of all received points.

In order to develop a surface fitting algorithm six important parameters should be defined, namely:

- the direct surface
- the inverse surface
- the number of calibration data points
- mathematical surface model
- model parameter search algorithm
- the error calculation function.

The first three parameters go hand in hand, as the range of the direct surface is also the domain of the inverse one, and both domains depend directly on the number of calibration points. The calibration data should be well defined covering the whole direct function domain. For discrete signal coupling systems, the calibration points would correspond to the constellation specific points of the two antennas. In other words the whole domain of two 16QAM constellations is covered with at least 256 (16×16) data points, while 64QAM constellation requires a minimum of 4096 (64×64) points as the influence of all point combinations should be analyzed. The number of calibration points can vary, and actually for denser constellations like 64QAM it is not essential to use all signal pairs if the remaining calibration points are well spread over the constellation domain. Eventually, the required number of calibration points will be analyzed in detail once the surface model and simulation system are developed.

Definition of the Fitting Surface model

The inverse coupling fitting surface requirements include easy implementation, low processing demands and above all the existence of the unique solution for all data of the inverse function domain. In general, fitting surface model should be developed and adopted to the corresponding physical process or mathematical problem. For example, the trajectory of the flying balloon influenced by gravity would follow a parabolic path. Thus matching trajectory data points to a parabolic curve would make sense. On the other hand, the trajectory of the anchored ship on the sea depends on the effects of the Moon, and the ship goes up and down in a sinusoidal pattern. Hence, the data points should be matched to a sine wave, or the sum of two sine waves of different periods, if the effects of both the Moon and Sun are considered.

Inside our nonlinear coupling case, the inverse coupling function of the higher order appears. Mathematically speaking, low-order polynomials tend to be smooth and high order polynomial curves like the one of the ninth degree, tend to be "lumpy" with several local minimum and maximums making the surface approximation severe. As the surface is not smooth many degrees of freedom are needed. This conclusion is confirmed if \hat{x}_1 is expressed directly from the eq. (67):

$$\begin{aligned}
& (c_{13}c_{23}^3) \hat{x}_1^9 + (3c_{13}c_{22}c_{23}^2) \hat{x}_1^8 + (3c_{13}c_{21}c_{23}^2 + 3c_{13}c_{22}^2c_{23}) \hat{x}_1^7 \\
& + (-c_{12}c_{23}^2 + 6c_{13}c_{21}c_{22}c_{23} + c_{13}c_{22}^3 - 3c_{13}c_{23}^2y_2) \hat{x}_1^6 \\
& + (-2c_{12}c_{22}c_{23} + 3c_{13}c_{21}c_{22}^2 + 3c_{13}c_{21}^2c_{23} - 6c_{13}c_{22}c_{23}y_2) \hat{x}_1^5 \\
& + (-2c_{12}c_{21}c_{23} - c_{12}c_{22}^2 + 3c_{13}c_{21}^2c_{22} - 3c_{13}c_{22}^2y_2 - 6c_{21}c_{23}c_{13}y_2) \hat{x}_1^4 \\
& + (c_{11}c_{23} - 2c_{12}c_{21}c_{22} + 2c_{12}c_{23}y_2 + c_{13}c_{21}^3 - 6c_{13}c_{21}c_{22}y_2 + 3c_{13}c_{23}y_2^2) \hat{x}_1^3 \\
& + (c_{11}c_{22} - c_{12}c_{21}^2 + 2c_{12}c_{22}y_2 - 3c_{13}c_{21}^2y_2 + 3c_{13}c_{22}y_2^2) \hat{x}_1^2 \\
& + (c_{11}c_{21} + 2c_{12}c_{21}y_2 + 3c_{13}c_{21}y_2^2 - 1) \hat{x}_1^2 + (y_1 - c_{11}y_2 - c_{12}y_2^2 - c_{13}y_2^3) = 0
\end{aligned}$$

where a set of higher order polynomial roots appear as a solution.

At this stage, the straightforward modelling approach would be to construct the surface out of polynomial roots of different orders. However, root calculation, apart from not offering the unique solution, is computationally extremely demanding process and therefore this kind of surface model is discarded from the beginning. On the other hand, other types of curves, offering many degrees of freedom such as the conic ones might offer a solid basis for the surface model construction. Mathematically this kind of curves is formed by intersecting a circular conical surface with a plane one. Their advantage is that inside one general function form different structures like circular, elliptical, parabolic, and hyperbolic arcs are described depending on the function parameters. Furthermore, the function is expressed in a simple way and it offers a unique solution to each input pair. Actually, in the Cartesian coordinate system, the graph of a quadratic equation of two variables is always a conic curve, and all conic sections arise in this way:

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0 \quad (74)$$

with A, B, C not all zero. Then with $D = E = F = 0$

- if $B^2 - 4AC < 0$, the equation represents an ellipse ($\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$, $\frac{x^2}{b^2} + \frac{y^2}{a^2} = 1$);

- if $A = C$ and $B = 0$, the equation represents a circle ($x^2 + y^2 = a^2$);
- if $B^2 - 4AC = 0$, the equation represents a parabola ($y^2 = 4ax$);
- if $B^2 - 4AC > 0$, the equation represents a hyperbola ($\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$, $\frac{x^2}{a^2} - \frac{y^2}{b^2} = -1$);
- if $A + C = 0$, the equation represents a rectangular hyperbola ($xy = c^2$).

Eventually, using different indexing and parameter denomination, the conic curve in eq. (74) can be rewritten as:

$$k_{20}x^2 + k_{11}xy + k_{02}y^2 + k_{10}x + k_{01}y + k_{00} = 0 \quad (75)$$

or in a more generalized form as:

$$\sum_{i=0}^{i \leq 2} \sum_{j=0}^{j \leq 2} k_{ij}x^i y^j = 0, \quad k_{12}, k_{21}, k_{22} = 0 \quad (76)$$

It is interesting to deduce that all two dimensional conic curves have been constructed without the higher order multiplications, and with just one simple linear multiplication. However, if we were to construct a surface model using strict conic curve definition the N_0 would be equal to two, and the approximation surface would be constructed based on six parameters k_{ij} , as the parameters k_{12}, k_{21}, k_{22} are set to zero. Since the inverse nonlinear coupling cancellation function is a nine dimensional case, only two degrees of freedom per variable and just six overall degrees of freedom do not guarantee good enough approximation. Three additional levels of freedom can be achieved by introducing parameters k_{12}, k_{21}, k_{22} which are not necessary set to zero inside the inverse coupling approximation. Furthermore, it is logical that the ideal approximation model of the ninth order space should at least be of the same order, thus the introduction of the polynomials of the higher order is proposed.

Adjusting the eq. (76) to our coupling cancellation problem, the variables x and y would correspond to y_1 and y_2 , and the approximated solution of the equation instead of zero becomes \hat{x}_1 .

$$\hat{x}_1 = \sum_{i=0}^{i \leq N_0} \sum_{j=0}^{j \leq N_0} k_{ij} y_1^i y_2^j \quad (77)$$

where N_0 represent the degrees of freedom of variables y_1 and y_2 . ($N_0 = 2$ in eq. (76))

According to the eq.(77), the increase of the degrees of freedom of the received variables N_0 , is followed by the exponential increase in the number of surface model parameters. Hence, by setting $N_0 = 9$, a staggering 81 parameters should be established in order to define the corresponding surface model. Theoretically, using this number of degrees of freedom would make the approximation more accurate but at the expense of the unjustified exhaustive solution search. Having in mind that complex curves can be formed without higher order multiplications, the complex surfaces should also follow the same path. Hence, the proposed approximation model is defined as a sum of two independent higher order polynomials with one simple linear multiplication [Ara07b]:

$$\hat{x}_1 = \sum_{i=1}^{i \leq N_1} a_i y_1^i + \sum_{j=1}^{j \leq N_2} b_j y_2^j + k_1 y_1 y_2 + k_2 \quad (78)$$

where N_1 and N_2 represent the degrees of freedom of variables y_1 and y_2 respectively, while variables a_i, b_j, k_1, k_2 are the decoupling coefficients, determined inside the surface approximation search. The ranks of polynomials, controlled with (N_1, N_2) , are determined as a trade off between the acceptable system performance and the required computational power. By setting the approximation precision to $(N_1, N_2) = (9, 9)$, 81 coefficients are reduced to 20 ($k_1, k_2, a_i, b_j; i, j = 1..9$) that have to be determined in order to fully describe the corresponding nonlinear coupling cancellation fitting surface.

In order to be able to adopt the surface search method to general decoupling problem the data on which the search is carried out is restrained to $\langle -1, 1 \rangle$. This is achieved by implementing the normalization according to the maximum received symbol level taking into account all calibration symbols on both antennas, as shown in the following equation:

$$y_{1k}^N = \frac{y_{1k}}{\text{Max}(|y_{1k}|, |y_{2k}|)} \quad k = 1..N \quad (79)$$

$$y_{12k}^N = \frac{y_{12k}}{\text{Max}(|y_{1k}|, |y_{2k}|)}$$

where y_{1k}^N and y_{2k}^N are the k-th normalized received symbols on the corresponding antennas, $\text{Max}(|y_{1k}|, |y_{2k}|)$ is the maximum received absolute value level, and N is the number of calibration signals. Eventually, since the approximation surface is constructed based on

the normalized levels, the true value of reconstructed symbol is obtained by expanding the reconstructed normalized symbol \hat{x}_{1k}^N to the received dynamic range:

$$\hat{x}_{1k} = \hat{x}_{1k}^N \text{Max}(|y_{1k}|, |y_{2k}|) \quad (80)$$

Hence, the real decoupling approximation surface model is defined as:

$$\hat{x}_{1k} = \text{Max}(|y_{1k}|, |y_{2k}|) \left\{ \sum_{i=1}^{i \leq N_1} a_i (y_{1k}^N)^i + \sum_{j=1}^{j \leq N_2} b_j (y_{2k}^N)^j + k_1 y_{1k}^N y_{2k}^N + k_2 \right\} \quad (81)$$

With the introduction of normalization, the search domain is well defined between $\langle -1, 1 \rangle$, and the search depth is easily controlled with symbol resolution factor. However, the search process is numerically carried out on the same data area regardless of the coupling strength or system transmission parameters. This proposal allows the adaptation of implemented search methods to general decoupling problem offering at the same time constant precision and efficient optimization of search process.

If coupling strength information is known inside the receiver than, after \hat{x}_1 is determined, \hat{x}_2 can be calculated according to eq. (65):

$$\hat{x}_2 = y_2 - c_{21}\hat{x}_1 + c_{23}\hat{x}_1^3 \quad (82)$$

However, as the information of direct coupling function is not available inside the mobile terminal, the same form of fitting surface with the adequate parameters can also be constructed for \hat{x}_2 :

$$\hat{x}_{2k} = \text{Max}(|y_{1k}|, |y_{2k}|) \left\{ \sum_{i=1}^{i \leq N_1} d_i (y_{2k}^N)^i + \sum_{j=1}^{j \leq N_2} g_j (y_{1k}^N)^j + h_1 y_{1k}^N y_{2k}^N + h_2 \right\} \quad (83)$$

This way, the decoupling process can be seen as a simple software module used for the digital signal enhancement where the processing load of the inverse coupling function is basically concentrated in the parameter search of the approximation fitting surface. Once all design parameters of the inverse coupling surface are defined, the decoupling, based on eq.(81 and 83),

is implemented inside coupling cancellation module on all incoming pair of signals achieving signal enhancement with almost zero processing demands.

Cost Function

Inside signal coupling module, the reference surface is defined with a pre-established set of calibration symbols ($x_{1i}, x_{2i}, i = 1..N$) and their corresponding values after transmission, coupling and A/D conversion marked with ($y_{1i}, y_{2i}, i = 1..N$). On the other side, the decoupling, or inverse surface is obtained using the A/D conversion values ($y_{1i}, y_{2i}, i = 1..N$) as input data for reconstructing the calibration symbols ($\hat{x}_{1i}, \hat{x}_{2i}, i = 1..N$) according to decoupling approximation surface model in eq.(81) meaning that each set of parameters constructs a different surface. The chosen model function has adjustable parameters ($k_1, k_2, a_i, b_i, i = 1..N$) defining the surface. The objective of decoupling module consists in adjusting these parameters to best fit a calibration data set according to a defined error criterion or so called cost function. The cost function calculation has to be carried out for each potential surface candidate and carries the mayor part of the decoupling processing load. Thus the selection of the cost function should be made as trade off between the processing load and system overall performance.

The typical cost function for this kind of problems is known as the ordinary least sum square method. This method was first described by Carl Friedrich Gauss around 1794 and today forms part of most statistical software packages. It determines the values of unknown quantities in a statistical model by minimizing the sum of the squared residuals (the difference between the predicted and observed values):

$$Q_1 = \sum_{i=1}^N (x_{1i} - \hat{x}_{1i})^2 \quad (84)$$

where N is the number of symbols used for calibration, x_{1i} is the i^{th} calibration symbol theoretical and \hat{x}_{1i} calculated value according to eq. (81) using the corresponding model adjustable parameters. The set of decoupling coefficients displaying minimum Q_1 value are then used in decoupling module as the inverse coupling surface function.

Another cost function similar to least square method is root mean square, also known as the quadratic mean. In mathematics this measure is known as a statistical measure of the magnitude of a varying quantity. On the other hand, physical scientists often use the term root-mean-square as a synonym for standard deviation when they refer to the square root of

the mean squared deviation of a signal from a given fitting model. The name comes from the fact that it is the square root of the mean of the squares of the values:

$$Q_2 = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_{1i} - \hat{x}_{1i})^2} \quad (85)$$

Again, a set of decoupling coefficients displaying minimum Q_2 value are seen as the best fit inverse coupling surface. As the square calculation of potential surfaces over all calibration points, might overload the processing power of some terminals, the sum of absolute differences can also be implemented as the cost function:

$$Q_3 = \sum_{i=1}^N |x_{1i} - \hat{x}_{1i}| \quad (86)$$

Weighted least squares, and weighted sum of absolute differences are methods similar to least squares and the sum of absolute differences respectively where unlike other cost functions all points are weighted equally. Instead, they are weighted in a way that points with a greater weight contribute more to the fit:

$$Q_4 = \sum_{i=1}^N w_i (x_{1i} - \hat{x}_{1i})^2 \quad (87)$$

$$Q_5 = \sum_{i=1}^N w_i |x_{1i} - \hat{x}_{1i}| \quad (88)$$

where w_i is error weight. Often, w_i is given as the inverse of the variance, giving points with a lower variance a greater statistical weight ($w_i = 1/\sigma_i^2$). Once more, a set of decoupling coefficients displaying minimum Q value are used for the inverse coupling surface construction. However, as this approach requires additional processing load it is not suitable for the implementation on the mobile terminal, and will not be treated inside this thesis.

Eventhough the surface obtaining minimum Q value is mathematically the best fit according to the corresponding cost function, it is not necessarily the best fit surface when the entire system performance is analyzed. Since each transmission system has its own special requirements the numerical values based on mathematical calculations might lead us the wrong way.

For example, the minimum Q function of the 64QAM transmission system is obtained with the surface approximating perfectly 63 positions and presenting significant error only at one constellation point. At the same time another surface that approximates equally well all 64 points presents higher cost function value. If the performance of the overall system exhibits better results with the second surface then with the first one, the implemented cost function type is not adequate type for the corresponding system. Since Q_1 and Q_2 are very similar having the only difference in one square root and one division, only the behaviour of Q_2 and Q_3 cost functions on the inverse coupling approximations will be analyzed. The cost function which suites best the system overall BER performance and search convergence speed will be used inside decoupling module search algorithms.

Parameter search

With the introduction of the approximation methods, the coupling cancellation function search is transferred into minimization the inverse surface cost function which directly depends on the surface model parameters. The precision of the approximation depends greatly on the number of model parameters and on the parameter space where the search is carried out. An exhaustive exploration of the search domain becomes rapidly impossible as the number of design variables increases, even with very small number of sampling points per variable. In fact, if M values are chosen for each of the N design variables, the number of possible configuration to be explored is M^N making it clearly impractical to explore all the existing configurations. Having in mind that the processing power of mobile terminals is limited, some kind of optimized, easy-to-use, search algorithm should be implemented. Classical instances of heuristic approximate methods like simulated annealing, tabu search, genetic algorithms, probabilistic hill climbing, classical perturbations methods etc. [Wal99, Aar97] offer the solution to the search problem.

1	accuracy
2	speed
3	simplicity
4	independent of starting point
5	robustness

Table 3: The priorities of nonlinear RF front-end algorithm

All heuristic approximation methods are widely used in optimization because of their simplicity

and robustness, however none can be denoted as the best one for all the applications. The methods differ in dependence of starting point value, speed of finding the optimum solution, accuracy, processing load etc. The cost function of the inverse problem of a realistic device can be a function with several local minima's that are very different, in value and location, from the global one, complicating the minimum search process. The list of priorities for the nonlinear RF front-end coupling cancellation algorithm is shown in Table (3).

Since accuracy has high priority, the center of the attention is directed towards techniques based on Simulated Annealing as this class of methods can statistically guarantee finding an optimal solution. Among other good properties, it can process the cost functions possessing arbitrary degrees of nonlinearities and discontinuities. Mathematical proofs state that this procedure converges to the global minimum, but no estimate can be made on the number of trials needed to reach it. In other words, the optimal solution is found at the expense of the speed and that is why this algorithm is often pushed away by faster algorithms sacrificing the accuracy. As speed property is also positioned high on the property list, it is clear from the start that this method can not be implemented inside the mobile terminals as all excellent search properties are cast aside by heavy processing load and slow convergence. In fact, the purpose of this method is to act as a reference measurement in the first place regarding the system decoupling accuracy. This algorithm will also be used for determining the most appropriate cost function while the size of the required calibration points will be determined for each search algorithm individually. Eventually, once all the system parameter are defined, the search based on Simulated Annealing, Very Fast Simulated Annealing algorithm, guided multi-level search and their combination will be conducted and analyzed. Guided multi-level search method is completely new deterministic method, developed inside this thesis for fast and accurate nonlinear coupling cancellation.

Calibration

With the introduction of the fitting surface approach the solution to the inverse coupling problem can be divided into two steps: calibration process and point-to-point real-time decoupling. The calibration process includes the transmission of the calibration data used for the reference surface construction and the search for the inverse surface based on some search optimized algorithm.

The second step includes the real-time decoupling based on the obtained inverse coupling

surface. Even though the second stage is done at signal to signal bases, by developing simple inverse surface model, the majority of the processing load is carried inside the calibration process. Thus the overall system processing requirements depend on the number of calibration signal, heuristic algorithm complexity and above all on the calibration process repetition rate.

The calibration approach is often found inside the transmission systems with multiple antennas, and especially when direct downconverter methods are implemented. Actually, the I/Q channel phase and amplitude mismatches are minimized with the calibration path included into the RF design of the future mobile terminals in [Bar04, Fer04]. Inside this proposal the RF receiver is designed with a double down conversion: from RF-band ($\approx 5GHz$) to L-band ($\approx 2GHz$) and than a direct conversion to base band. Than, by choosing the same L-band for both the transmission and receiver chains the calibration path inside a mobile terminal receiver can be designed as shown in Figure 46.

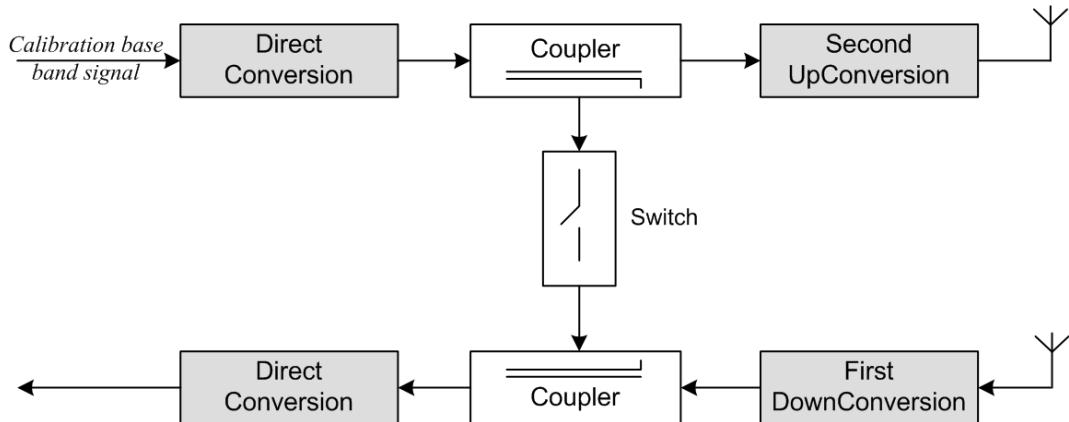


Figure 46: Block diagram of the calibration path

Once a calibration base band signal is converted up to L-band, using the direct upconverter, the calibration path is enabled by the switch, and the signal is passed to the direct down-converter. The test signal is processed digitally at the A/D output as to correct as much as possible the I/Q impairments. In this operation mode, second conversion stage is disabled so that the RF module is neither transmitting nor receiving. The authors propose this calibration to be carried out periodically on each transceiver chain of the MIMO RF front-end, but no estimate on the calibration duration or repetition rate is given. However, the importance of this approach is that the same calibration path can be used for the reference surface construction, with

the calibration process not covering only the coupling appearing in the second up-conversion transmission path.

The influence of the entire RF front-end on the signal coupling can be evaluated inside the reverberation chamber. This in fact, is the approach proposed in [Kil04] for the measurements of the multiple antennas radiation efficiency, diversity gain, correlation and mutual coupling. The measurements obtained with this chamber are low-time consuming and almost environment independent. However, they tend to give too optimistic results regarding capacity when compared with more realistic measurements [Gar07] while the accuracy of the coupling measurements is still to be analyzed. Nevertheless, in the absence of the calibration path inside the mobile terminal the reverberation chamber can be used for coupling calibration at the industrial level.

Naturally, the most accurate calibration is achieved inside the real system environment, when the calibration signals are transmitted from the mobile base station and received at the terminal RF front-end. Theoretically, this is the ideal approach from the physical point of view as the calibration is done inside the actual system environment. However, it is also the most complex one as higher system transmission levels need to be changed. This is due to the fact that the calibration signals are sent from the base station, and thus should be sent to every mobile terminal inside the connection procedure which would require the calibration process standardization. In this approach, fast detection of *the good enough* inverse coupling surface approximation would be of the highest importance. Having in mind the complexity of the inverse surface parameter search without the starting point close to the global minimum, the real time requirements would eventually lead to poor decoupling module performance in many cases.

I will not enter into details on how to build a proper reverberation chamber or implement the calibration path inside the mobile terminal as those topics occupy wide investigation area, and could present phd thesis on their own. From the point of view of this thesis, it is important to emphasize that it is assumed that the calibration at physical level can be done and that no technological constraints are present. Physical calibration can be done at three different stages:

1. level: industrial calibration (inside reverberation chamber)
2. level: terminal calibration (using terminal calibration path)

3. level: real-time calibration (in real system environment)

Each level calibration can be executed independently or one after another using the solution of the previous level as a fitting surface starting point. Depending on the required accuracy only one calibration stage can be sufficient for some coupling cancellation environments.

The coupling phenomena depends on many variables, starting with the selection of electronic elements, throughout the RF front-end design, up to the working conditions of the system which include system bandwidth, intermediate frequency, channel parameters etc. Based on this fact inside the industrial calibration stage, the coupling as a consequence of RF front-end design is cancelled. This can be seen as the first level rough inverse function approximation. The second level parameter search should cancel the coupling specific to the selection of electronic elements and corresponding terminal construction implementation. After these two calibration stages, the inverse coupling surface is well defined with global minimum points area located and locked. The overall signal environment can not drastically change the already experienced coupling and so the inverse surface is already close to its best fit surface. Thus the search at the third level has no need to wonder around the surface looking for the minimum cost function and the parameters can be adjusted inside real time constraints.

I will assume that the only calibration type is available at the time. The parameter search algorithm will be developed and implemented on this bases. Using the system simulations, the size of the calibration data will be analyzed as a function of the required decoupling accuracy, given signal constellation and generated processing load. The calibration will be analyzed in noiseless environment corresponding to the industrial calibration and with AWGN noise corresponding to second and third level calibration.

5.3 System description

A proposed RF front-end coupling model, together with the nonlinear coupling cancellation module, can be seen in Figure 47. The module is positioned between the outputs of A/D converter and any further digital signal processing. This location is chosen since it least dependent on the transmission signal type as module inputs are eventually samples taken at the sampling rate of A/D converter. Inside the module, signal enhancement is achieved based

on the approximation of the inverse nonlinear coupling function with polynomial of up to the 9th degree. Hence, with maximum value of N_1 and N_2 restricted to 9, according to eq. (81):

$$\hat{x}_1 = \sum_{i=1}^{i \leq N_1} a_i y_1^i + \sum_{j=1}^{j \leq N_2} b_j y_2^j + k_1 y_1 y_2 + k_2$$

up to 20 parameters have to be determined in order to define the approximated decoupling function. Since the domain is well defined with data between $(-1, 1)$ the search process can be efficiently optimized and after the best fit inverse coupling surface is located, signal enhancement is achieved with almost zero processing demands.

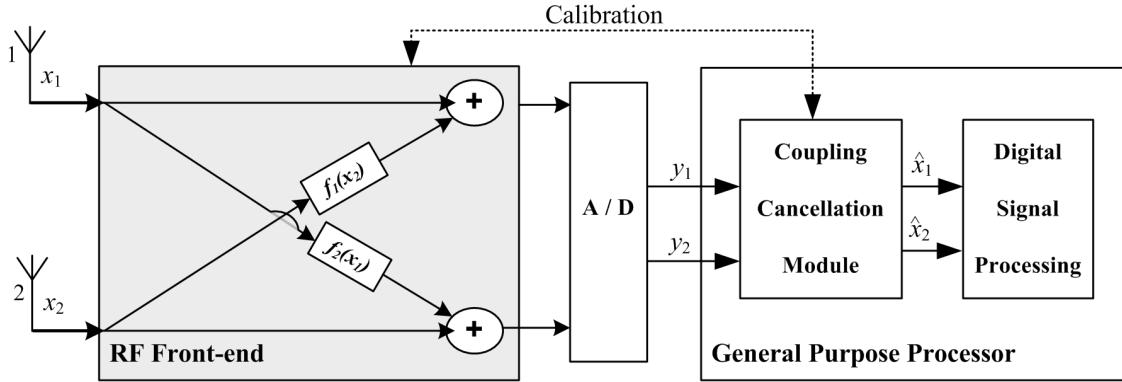


Figure 47: Simple scheme of coupling cancellation module position

In order to simulate the behaviour of nonlinear RF front-end coupling distortion a system model in MatLab is developed. The complete digitally implemented structure of the MatLab system model is depicted in Figure 48. The system consists of two independent antennas in the transmitter and two in the receiver with I/Q downconversion processing. Randomly generated bits are passed to modulation module which supports binary, 4-PSK, 16QAM and 64QAM constellations. At the modulation output a complex value is generated with real part corresponding to in-phase branch (I), and imaginary part to the in quadrature branch (Q). The signal is then oversampled with factor M and multiplied with the corresponding complex oscillator. Afterwards, AWGN noise is superimposed to the sum of the oscillator outputs and than the signal of two antennas is passed through the coupling module.

The coupling module is based on coupling functions defined in eq.(62) and (63) as:

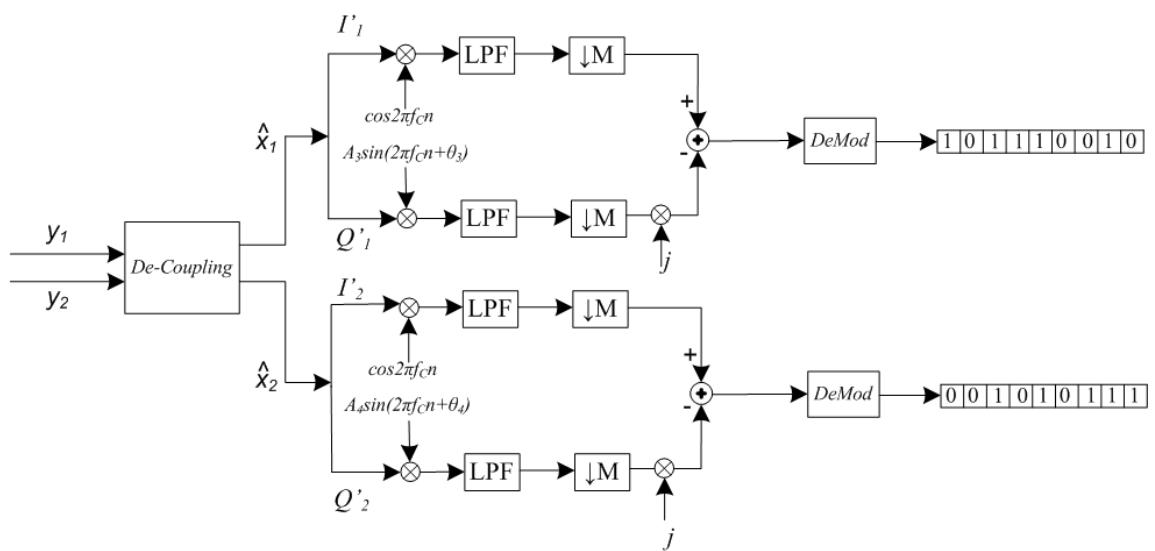
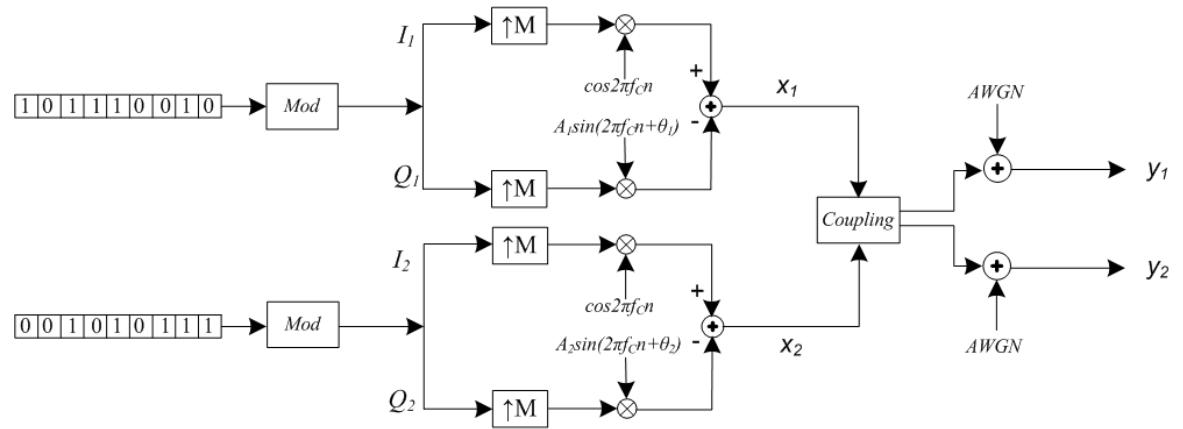


Figure 48: The structure of digital simulation model developed in MatLab

$$y_1 = x_1 + c_{11}x_2 + c_{12}x_2^2 + c_{13}x_2^3$$

$$y_2 = x_2 + c_{21}x_1 + c_{22}x_1^2 + c_{23}x_1^3$$

The parameters describing the coupling module are two set of coupling distortion parameters expressed in [dB] ($[C_{11}, C_{12}, C_{13}], [C_{21}, C_{22}, C_{23}]$).

At the receiver side two received signal are passed through decoupling module, and then to the corresponding quadrature downconverters. The decoupling module is located before low pass filtering which prevents neglecting the influence of second order coupling on decoupling results.

Since practical receiver implies imbalances in in-phase and quadrature branches, amplitude and phase mismatches of I/Q signals are implemented as imbalanced complex oscillator:

$$CLO = \cos(2\pi f_0 t) + jA \sin(2\pi f_0 t + \phi) \quad (89)$$

The imbalance factor is by default set to zero, and is taken into account only at the last stage of the analysis when different decoupling algorithm performances are analyzed. The imbalanced complex oscillator is followed by low-pass filtering and decimation with the same factor M . Finally, signal demodulation takes place, and the output bits are compared to the sent ones in order to calculate the bit error rate. The correct functioning of the system model was confirmed comparing the obtained simulation results without coupling and I/Q imbalances, with the theoretical values of the different constellations from [Lu99] as those values set the system performance limits.

As expected, in presence of coupling the system performance deteriorates as seen on Figure 49, where system with and without antenna coupling are shown. The simulations correspond to 16QAM Gray constellation. Since the existance of second order coupling is irrelevant for the receiver based of I/Q downconverter, the square coupling can be neglected when the performance of entire system is analyzed. Thus, the simulations are made for different coupling parameters expressed only with two physical factors ($[C_1, C_3]dB$), which are set to the same value on both antennas. The curve without coupling corresponds to the theoretical system performance under the influence of simple AWGN noise. It can be observed that nonlinear

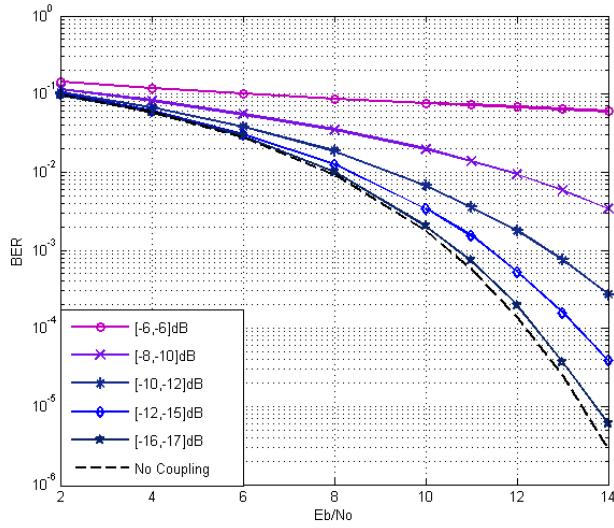


Figure 49: 16QAM system under different coupling interference levels $[C_1, C_3]$

coupling of $[-10, -12]dB$ leads to approximately $2dB$ loss at $BER=10^{-3}$, and as coupling strength increases this loss is rapidly increased.

The next set of simulation, shown in Figure 50, is performed for denser signal constellation, 64QAM, in order to see the performance of constellation that is more susceptible to errors.

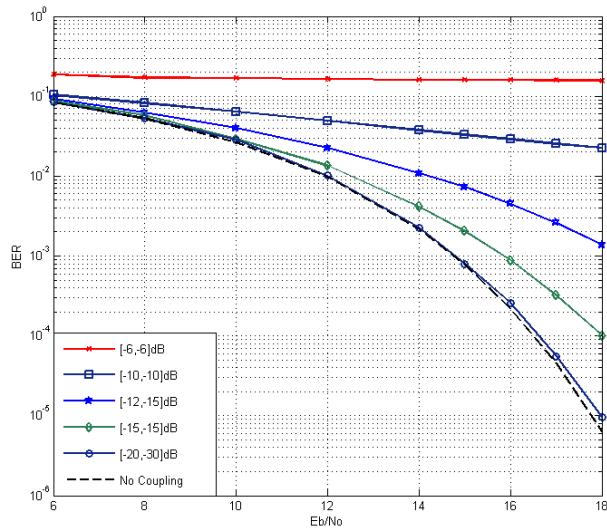


Figure 50: 64QAM system under different coupling interference levels $[C_1, C_3]$

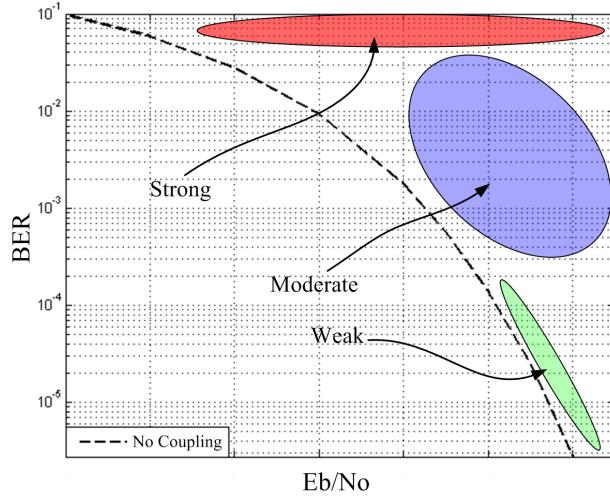


Figure 51: Coupling strength classification depending on the lost system performance

The results show stronger influence of coupling on denser constellation since the same coupling level (e.g. $[-12, -15]dB$) at $BER=10^{-3}$, leads to almost $2dB$ lost in 64QAM system while in 16QAM the loss is restrained to less than $1dB$. As the signal constellation gets denser the system performance is more vulnerable, and very low coupling strength like $[-20, -30]dB$ is able to degrade the system performance. Thus, even at this stage, the decoupling module is still a necessity as without it BER shows slower progress.

The simulations presented show the necessity of coupling cancellation module, as otherwise system performance is easily drawn to saturation. Analyzing the simulation figures, the influence of coupling depending on the system performance can be divided in 3 levels: strong, moderate and weak. The strict definition of these three levels can not be made as it depends on the constellation type. However, the borders can graphically be seen in Figure 51 where the presented curve shows theoretical performance of the corresponding system without coupling. Thus when referred to strong coupling, we are talking about the coupling that almost completely destroys the received signal and whose system performance almost shows flat line. Moderate coupling level follows the theoretical curve, but the system performance exhibits slow improvement, while the weak coupling system follows closely the ideal theoretical curve but the system performance is still visibly deteriorated. Even though weak coupling is expected to appear more often in real life situations, in the following, the search for the decoupling function will mainly be made for strong and medium coupling levels as they are more demanding cases. The decoupling module is expected to show excellent results for weak and moderate coupling.

However, as strong coupling includes wide range of hostile coupling environments, the correct functioning of decoupling module cannot be guaranteed for all cases. The limiting case for which the performance of system with decoupling module is still considered good enough for transmission will be set empirically through simulations.

6 Decoupling Module Search Algorithms

The proposed decoupling module presented in previous chapter is positioned between the outputs of A/D converter and any further digital signal processing. Inside the module, signal enhancement is achieved based on the surface approximation of the inverse nonlinear coupling process. Surface construction is based on the calibration set of signals and has no restriction on the physical coupling source. The received set of calibration signals is normalized so that the corresponding numerical values are located between $\langle -1, 1 \rangle$. Thanks to this normalization neither coupling strength nor coupling form are restricted in any way and the optimal search parameters obtained in this thesis can directly be used on any coupling type of problem. Even with optimized search parameters the search for the best fit surface approximation presents the most demanding part the decoupling module. In this chapter three different search approaches are presented.

The first approach is based on Simulated Annealing (SA) algorithm. This is a heuristic kind of method which offers high precision but requires large number of iterations to reach it. Even though this method can not be used under real time conditions it provides the upper limit of the decoupling precision module and actually confirms the adequate selection of the surface model. The second approach is based on the Improved Fast Simulated Annealing (IFSA) method which is here developed specifically for nonlinear decoupling problem. Inside IFSA, the heuristic methods are preserved, and the search progress is accelerated on the cost of the decoupling precision. The idea and the development of the third approach presents the authors contribution to the surface approximation methods. The approach is denominated Guided Multi-level (GML) search and ensures deterministic solution to the decoupling problem with good precision versus speed ratio and low processing demands.

In order to adapt to the nonlinear decoupling problem, each method is analyzed in detail and the corresponding adequate set of search parameters is proposed. These parameters are later used in chapter 7, when the performance of decoupling module search algorithms is presented and compared.

6.1 Simulated Annealing

The history of Simulated Annealing (SA) starts with the first contribution regarding the modified Monte Carlo (MC) procedure published by Metropolis et al. in 1953 and associated to Chemical Physics [Met53]. The aspiration of authors was to find a general method for calculating the properties of any substance composed of individual molecules interacting. In mathematical form, these interactions are modelled as several hundred dimension integral. However, since it is impractical to carry out the integration of such a dimensions, a well known Monte Carlo stochastic technique emerged as a possible solution search method. With Monte Carlo algorithm a function is estimated by randomly selecting integration points, and then taking some kind of average of the function values at those points. Logically, better solution is accomplished if more points are taken inside the interval in which the contribution of the integrand function is higher compared to the other intervals. However, finding precisely those intervals is just as difficult as solving the integral in the first place, and therefore the process loses the stochastic property. Seeing that a brute force Monte Carlo method was useless for interactions of more than a few particles, Metropolis proposed a practical solution in a form of modified (hybrid) Monte Carlo procedure. The idea is that instead of choosing randomly the particle configurations (points), and than weighting them with some probability function, the configurations are directly chosen according to the probability function and no posterior weighting is required. This is commonly known as Metropolis algorithm and has since then been used to solve difficult problems in statistical physics for over fifty years now.

A technique based on the Metropolis algorithm known as Simulated Annealing was mentioned for the first time some thirty years after the original paper. It was developed independently by S. Kirkpatrick, C. D. Gelatt and M. P. Vecchi in [Kir83], and by V. Cherný in [Che82]. From then on it has been widely used in several engineering branches, usually in applications dealing with an enormous number of possible configurations which are to be explored. The name and inspiration come from the annealing in metallurgy, as the technique mimics the nature of the metal which is heated and than cooled slowly in order to reach its absolute minimum energy state. The heat causes the atoms to become unstuck from their initial positions (a local minimum of the internal energy) and wander randomly through the states of higher energy. If the liquid is cooled too quickly the atoms will get stuck inside the area close to some local minimum, and eventually will only reach suboptimal energy state. By gradually cooling the metal, the particles have more chances of finding the configuration with lower internal energy than the initial one. The basic assumption of the SA method is that sometimes, in order to

avoid the trap of local minimum, points that lead to worse minimum solution are accepted. This is commonly known as *uphill* movement as its orientation is contrary to the minimum energy direction.

By analogy with this physical process, the function to be optimized can be seen as the energy of the fluid, while the design variables represent the positions of its molecules. The goal is to bring the system from the arbitrary *initial* state, to a state with the minimum possible energy. This is achieved by letting the molecules wander which corresponds to random perturbation of the design variable coefficients. Before the new perturbation is accepted the SA algorithm calculates the change in the error function of the new *neighboring* (candidate) state. If the error function is lower the perturbation is accepted. On the other hand, if the neighboring energy level is higher the new position is accepted according to the probability function $p = f(T)$, where parameter T is denominated as temperature in accordance with physical process. Therefore, T is the parameter which is gradually decreased as the iteration progresses. The probability function is chosen to satisfy the following requirements:

- When T is large the algorithm is unlikely to become trapped in local minimum as *uphill* movements are accepted with more probability
- as T goes to zero (cooling process) drastic solution changes are not allowed and the algorithm only accepts perturbations that lead towards lower energy levels.

During the last few decades a number of variations of original simulated annealing algorithm have been published. The comparison of different simulated annealing algorithms, with other mostly heuristic algorithms can be find in [Laa87, Joh87, Dra92, Ing92, Ing93, Ing96, Li02, Ing89]. Regardless of different SA variations, any algorithm based on physical simulated annealing has proven to be an effective global optimization method because of several important features:

- it can process the cost functions that possess arbitrary degrees of nonlinearities and discontinuities, as no restriction on the form of the cost function exists
- it can statistically guarantee finding an optimal solution
- it can be implemented easily when convergence speed is not relevant

Nevertheless, SA is not without its critics. The primary criticism is that it is too slow, especially nearby the optimal point, and that the implementation is computationally too demanding [Zer94, Dra92]. Some authors like [Ing96, Joh89] give their own calculations to demonstrate that SA can have very poor performance for global optimum search in some severe instances. Then there is another class of critics that the algorithm is too broadly based on physical intuition and is too short on mathematical bases, the implementation depends on the problem and sometimes it can be quite difficult fine tuning to a specific conditions. Still, none of the authors questions the precision, robustness and reliability of this method.

These three factors led to selecting the SA algorithm as a heuristic method for nonlinear decoupling problem. After correctly choosing the SA parameters, the decoupling solution obtained with this method is expected to offer precise solutions payed with slow search process and high processing demands. The election of SA parameters adopted to inverse nonlinear coupling search is done in the following section.

6.1.1 Adopting SA to Nonlinear Decoupling Problem

The functioning of the original Simulated Annealing method can be described using the pseudo code, divided into initialization and the simulated annealing part, presented in Figure 52. Inside the first section the temperature factor T is set to the initial temperature (T_0). The temporary minimum of the cost function (Q_{min}) is calculated for the initial set of decoupling coefficients ($Coeff_0$) and the algorithm takes this point as the starting position appointing their values to actual state configuration:

$$\begin{aligned} Q_{act} &= Q_0 \\ Coef_{act} &= Coef_{min} \end{aligned}$$

Afterwards, the algorithm enters the SA search loop until allowed processing time measured in number of iterations is exceeded ($Iter > I_{MAX}$). The selection of new coefficients (neighbours) is done inside the *New_Candidate* function, and then the cost function for the set of selected coefficients is calculated (Q_{new}). If the cost function ($Cost$) of the new neighbour presents lower value than temporary minimum cost function (Q_{min}), the perturbation is accepted. The new cost function value (Q_{new}) becomes the new minimum and the corresponding set of the coefficients is stored in $Coef_{min}$. In case of higher cost function value the perturbation

```

 $Coef_{min}=Coef_0$ 
 $Q_{min}=Cost\_Func(Coef_{min})$ 
 $Coef_{act}=Coef_{min}$ 
 $Q_{act}=Q_{min}$ 
 $T=T_0=1$ 
 $Iter=0$ 
 $k=set\_k()$ 

While ( $Iter < I_{MAX}$ )
    for 1 to  $N_{VCPL}$ 

         $Iter++$ 
         $Coef_{new} = New\_Candidate(Coef_{act})$ 
         $Q_{new} = Cost(Coef_{new})$ 

        if ( $Q_{new} < Q_{min}$ )
             $Q_{min}=Q_{new}$ 
             $Q_{act}=Q_{new}$ 
             $Coef_{min}=Coef_{new}$ 
             $Coef_{act}=Coef_{new}$ 
        else
            if random(0-1) < Trans_Prob(k, T, ΔQ)
                 $Q_{act}=Q_{new}$ 
                 $Coef_{act}=Coef_{new}$ 
            end if
        end if
    end for
     $T=Ann\_Schedule(T)$ 
end while

```

Figure 52: Pseudo code of Simulated Annealing method

is accepted only if the randomly selected number between zero and one is smaller than the $Trans_Prob$ function of the actual temperature T . However, in this case only the actual position is updated and the global minimum cost configuration is left unchanged. The search at one temperature level is limited with a number of visited candidates per level defined with N_{VCPL} parameter. Afterwards, the temperature T is decreased according to the annealing schedule function ($Ann_Schedule$) and the quest is continued.

According to the formulation of SA algorithm in Figure 52, when applying the SA method to a specific problem, one must specify:

- the state space or the coefficient resolution
 - symbol precision (Res_{symbol})
 - coefficient precision (Res_{coeff})
- the initial coefficients values or the starting point ($Coeff_0$)
- the neighbour selection method ($New_Candidate$)
 - coefficient maximum displacement (R_0)
 - number of visited candidates per temperature level (N_{VCPL})
- the form of the cost function ($Cost_Func$)
- the probability transition function ($Trans_Prob$)
 - initial (maximum) probability of the acceptance of the worst move (p_{max})
- the annealing schedule function ($Ann_Schedule$)
 - time reduction factor (α)
- maximum allowed number of iterations (I_{MAX})
 - minimum probability of worse move during the search (p_{min})
 - iterations corresponding to true SA search (I_{SA})
 - number of greedy downhill iterations (I_{DH})

The choice of SA parameters can have a significant impact on the method effectiveness. Unfortunately, there is no selection good enough for all problems, and there is no general way to find the best set of parameters for a given problem. Regarding the nonlinear inverse coupling, the majority of these factors have to be analyzed empirically. As the empirical search for this large set of parameters, which adopt best to the decoupling problem, is as complicated as the decoupling itself, some more generic parameters (e.g. the annealing schedule or probability transition function) will be set based on the experience of other authors like [Ros92, Ing93, Ing96, Men97, Joh89, Li02]. This way the system search will be over-dimensioned favouring the precision on the cost of speed. This set of parameters will be used

for locating the parameters more related to specific problem (e.g. initial temperature or number of visited neighbours per level). Once all parameters are set, fine tuning of SA parameters can be implemented with emphases on convergence speed while keeping the precision level constant.

The state space

This parameter is defined with the number of design variables, their discrete domain resolution and the overall symbol precision. Starting from a continuous domain the search coefficients are transformed into discrete domain by setting the distance between the two sample points. For example, by setting the resolution to 0.1 the coefficients can only obtain resolution multiples values like 0.2, -0.1, 0.8 etc. By setting the resolution to more dense value, like 0.0001, exactly 1000 times more possible solutions are available for each coefficient inside the SA search. Expanding the solution search area naturally leads to better solution precision, but on the expance on the processing time. Furthermore, if the resolution is too dense the SA progress can slow down the search without significant accuracy improvements. Apart from the coefficients resolution, the state space also defines the resolution of the received and decoupled symbols. This precision should be higher than the coefficient precision. However, it depends entirely on the transmission system requirements and implemented A/D converter.

In the following simulations the SA method is applied to the 20-coefficient decoupling function symbol resolution set to $Res_{symb} = 0.00001$. Even though the digital low pass filters are defined up to the ninth digit, the coefficients of decoupling module are restrained the same resolution as applied on symbols, that is, $Res_{coef} = 0.00001$.

Coefficients initial values ($Coef_0$)

For the initialization point of this SA search model, it is supposed that no coupling takes place, and that the signal arrives in perfect conditions without any distortion. Thus, using the decoupling model with polynomials of up to the 9th degree according to eq. (81), all coefficients are set to zero, except a_1 which is set to one:

$$Coef_0 : \begin{cases} a = [1, 0, 0, 0, 0, 0, 0, 0, 0, 0] \\ b = [0, 0, 0, 0, 0, 0, 0, 0, 0, 0] \\ k = [0, 0] \end{cases} \quad (90)$$

In fact, to make the indexing simpler, the coefficients are expressed as 20 dimensional vector transforming the eq.(90) to:

$$Coef_0 = \left[1, \underbrace{0, 0}_\text{19 zeros} \right] \quad (91)$$

where first nine elements correspond to coefficients $a_i, i = 1..9$, second nine elements correspond to $b_i, i = 1..9$, and the last two correspond to k_1 and k_2 .

By setting coefficients initial values according to eq. (91) the signals that suffered larger distortions require longer search time as the solution is located far from the starting point. This way the algorithm favours the appearance of smaller coupling since we expect strong coupling distortion to rarely emerge under realistic conditions. Finding the decoupling solution to the signals that suffer large distortion would theoretically require more time as their solution is located further from the staring point. Nevertheless, if sufficient number of maximum iterations is available, the choice of initial coefficients value has no influence on the final decoupling solution as the whole solution search area is then thoroughly analyzed. More precisely, the optimum solution can be found although with more processing time required.

Neighbour selection method

The neighbour selection method is particularly critical part, as at each iteration the algorithm randomly chooses the next step out of the fixed number of options called neighbours (candidates). When selecting new candidates, it must be possible to go from the initial state to a *good enough* state by relatively short path while at the same time allowing the search to scan the area but never loosing the *good* path from the sight. This is achieved by introducing the parameter ξ which defines the maximum allowed coefficient displacement. Naturally, as each coefficient can be seen as physically independent particle, all coefficients can change their state inside the maximum allowed radius when moving into the new position. Then, if the actual

decoupling coefficient set ($Coeff_{act}$) is expressed as 20 dimensional vector the selection of the new candidate set ($Coeff_{new}$) can be described as:

$$Coeff_{new}(i) = \xi_i R_0 + Coef_{act}(i); i = 1, \dots, 20 \quad (92)$$

where ξ_i is the random number between (-1) and (1). If maximum allowed displacement denominated as R_0 is large, the SA search can easily explore wide search area, but needs more iterations to reach the optimal solution as the algorithm is easily *distracted* and moved away from the *good* path. With R_0 small, the algorithm can get stuck in local minimum and finish the search far from the optimal solution. As the behaviour of SA method on inverse coupling search is unknown, for the initial simulations the maximum allowed displacement is set to $R_0 = 0.05$. With this relatively large factor the SA convergence speed is neglected allowing the solution precision to be left intact if the allowed number of maximum iterations is large enough.

Two main flaws of this neighbour selection method lie inside the possibility of the algorithm of picking already visited neighbours, as no path history is taken into account, and inside the possibility of wandering too far and spending too much time on getting back on the good path. As history path control would complicate too much the original SA method it is not implemented. However, as the original search process is often needlessly too slow and time consuming the wandering area of the original SA algorithm is restrained. This includes setting the actual set of coefficients, at the beginning of each temperature level, to the position experiencing the minimum cost function so far. Indeed, physically it can not be done during the metal annealing process, but it is not always advisable always to imitate the original physical annealing process into the tiny detail. This change significantly accelerates the mathematical solution search as it acts as the reset button if the search is routed in wrong direction. On the other hand, this reset button favours the solution close to the starting point, while solutions located far from the $Coeff_0$ need even more iterations than the original code to be located. This is in accordance with expected realistic conditions where extreme coupling strength rarely occurs.

The number of new candidate searches executed on the same temperature level is defined in N_{VCPL} parameter, which stands for the *Number of Visited Candidates Per Level*. Afterwards, the change in temperature is realized according to the Annealing Schedule, and the new candidates are again selected N_{VCPL} number of times. This amount of visited neighbours

per temperature level is directly related to the specific problem. If N_{VCPL} is large, many coefficients sets are tested and the probability of finding the optimal solution is increased. However, at the same time, the SA search progress is slowed down, as the valuable processing time is lost on the worthless algorithm wandering. For the starting simulations this parameter is set to $N_{VCPL} = 1000$, to satisfy the precision requirements and will be reduced empirically through the following simulations.

Cost function

As emphasized in previous chapters the objective of Simulated Annealing search consists in adjusting decoupling coefficients to best fit a calibration data set according to the cost function. The calculation of cost function has to be carried out for each new candidate and presents the most demanding part of the decoupling processing load. The cost function typically implemented for the search of approximate inverse function is root mean square, also known as the quadratic mean:

$$Q_{new} = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_{1i} - \hat{x}_{1i})^2} \quad (93)$$

where N is the number of symbols used for calibration, x_{1i} is the sent i^{th} calibration symbol on first antenna, and \hat{x}_{1i} its corresponding value calculated using the coefficients of new candidate ($Coef_{new}$) according to eq. (81):

$$\hat{x}_1 = \sum_{i=1}^{i \leq 9} a_i y_1^i + \sum_{j=1}^{j \leq 9} b_j y_2^j + k_1 y_1 y_2 + k_2$$

The final form of the cost function will be determined once all other SA parameters are selected. Because there is no reason to assume that quadratic mean cost functions offers significantly worse solutions when compared to other types of cost function described in section 4.2, this method is used in the first set of simulations. Afterwards, the sum of absolute differences is also implemented as the cost function primarily in order to analyze the processing load versus the precision and convergence speed. The more appropriate cost function will than be used in simulations corresponding to decoupling module search algorithms.

Probability transition function

The key SA feature that prevents the system of becoming stuck in a local minimum is its probability to move to the new state even when this new state has worse characteristics than the current one. The probability of making the transition from the current state to a candidate new state is a function of the energies of the two states, and of a global time varying parameter T denominated as *temperature*. Even though the algorithm is open to the use of any probability function (p), the new configurations are usually accepted or rejected according to the Boltzmann probability distribution:

$$p(\Delta E) = e^{-\frac{\Delta E}{k_B T}} \quad (94)$$

where k_B is the Boltzmann constant that relates temperature to energy. However, when SA problem is not related directly to the physical energy this constant is replaced with problem specific constant k . Hence, modified Boltzmann probability function is used and the annealing schedule function is adjusted accordingly:

$$p(\Delta Q) = e^{-\frac{\Delta Q}{kT}} = e^{-\frac{Q_{new} - Q_{act}}{kT}} \quad (95)$$

where Q_{new} and Q_{act} are the cost function values associated with the actual state, and new candidate state respectively, while k is the constant dependent on the specific problem. It is evident that even if ΔQ is positive, the probability transition function returns the value between zero and one. Thus randomly selecting the number between $\langle 0, 1 \rangle$ the system will always have a chance to move from one configuration to another. However for a given ΔQ the transition probability is not constant throughout the annealing process as it also depends on the actual temperature. By setting the initial ital temperature T_0 to higher values the probability of search wandering is increased. For example, if $\Delta Q = 0.1$ and initial temperature $kT_0 = 1$ the perturbation will be accepted if the random variable falls between $\langle 0, 0.9 \rangle$ which correspond to 90%, while for $\Delta Q = 0.1$ and initial temperature $kT_0 = 0.1$ the probability falls to 36%. Nevertheless, as the temperature is constantly decreased according to the annealing schedule, the probability of *worse* move tends to zero. This is in accordance with the physical simulated annealing where no wondering around is possible at low temperatures.

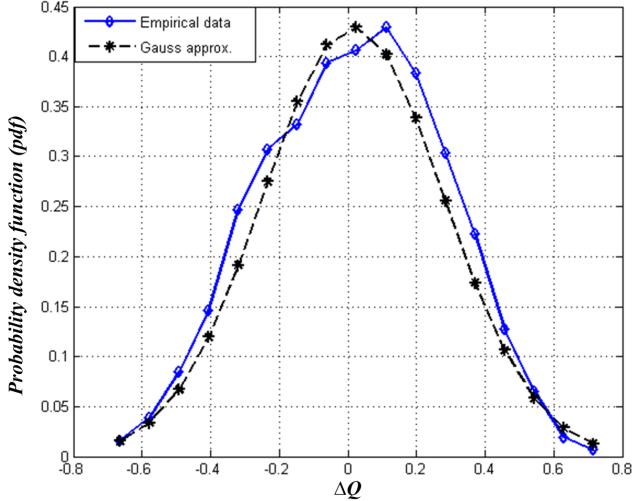


Figure 53: Approximation of the consecutive cost function differences with Gauss function

The selection of adequate constant k depends on the specific problem, and the initial probability of worse move acceptance defined as p_0 . As transition probability depends on the cost function difference, when defining the initial probability of worse move, we should take into account that the required probability should be achieved for wide range of energy differences. Actually, since new candidate selection method is based on random numbers (ξ_i) which follow the Gaussian probability function, ΔQ also appears to follow the same distribution. This is shown in Figure 53 where the histogram of energy differences of the 2000 consecutive candidates is approximated with Gaussian like distribution function. The candidates are selected with maximum allowed individual coefficient displacement of $R_0 = 0.05$. By approximating ΔQ distribution with Gaussian-like probability density function, three standard deviations represent the 99.7 percent of all possible cost function differences and the probability transition function can be adopted according to this value.

Hence, if 99% of new candidates have the initial probability of worse move set to p_{max} the initial probability transition function can be written as:

$$p_{max} = e^{-\frac{3\sigma_{\Delta Q}}{kT_0}} \quad (96)$$

As k relates directly actual temperature level to the cost function difference, it is typically set for $T_0 = 1$. Thus, according to eq. (96), numerical value of k is defined with initial probability

level p_{max} , standard deviation of the energy differences of two consecutive configurations:

$$k = -\frac{3\sigma_{\Delta Q}}{\ln(p_{max})} \quad (97)$$

By setting the initial probability of the worse move close to one, the initial search wander is allowed regardless of the energy difference between two candidates. This logically, slows down the search usually without any solution precision gained. Hence, the initial probability level is generally set to $p_{max} = 50\%$, and it is decreased according to the annealing schedule and modified Boltzmann probability function during the SA search process. The exact numerical value of k is calculated in `set_k` function before algorithm enters SA search loop. Inside the `set_k` function the approximation of Gauss distribution is obtained by implementing the actual neighbour selection method on 2000 new candidates.

Annealing schedule function

The annealing schedule function has no strict definition, as the function constraints restrict to allowing uphill surface exploration at the beginning, and then gradually restraining the search wondering favoring the movement in the direction of better solution only. Thus, the behavior of annealing schedule is specified by the user, with the obligation that the initial temperature T_0 is high enough to allow initial search wandering and that T equals zero towards the end of the given time budget. Inside physical simulated annealing this makes sense as the probability of the acceptance of the worse move is directly related to temperature. Thus, by requiring the probability of worst move acceptance of some amount p_0 , the annealing system temperature has to reach the level T_0 , where the desired temperature is expressed in Kelvin (K) and is directly related to Boltzmann constant $k_b = 1.3 \times 10^{-23} JK^{-1}$. As our system only simulates the physical process it uses modified Boltzmann function were constant k is set up to relate the required p_0 to a given T_0 expressed as a simple number without any direct physical meaning. In other words, inside mathematical interpretations the initial temperature level has lost its significance, and thus it is set to $T_0 = 1$ in order to simplify the calculation process.

A popular choice for the annealing function is the exponential schedule, where the temperature level, denoted as T or T_L , is decreased by a fixed factor α :

$$T_L = \alpha T_{L-1}; \quad T_0 = 1, \quad 0 < \alpha < 1, \quad L = 1, 2, 3\dots \quad (98)$$

where T_L and T_{L-1} are the new and the actual temperature respectively, and index L stands for the number of temperature level changes.

According to this annealing schedule, the system will initially wander inside a broad region containing good solutions and will ignore small energy differences. As the time passes it will drift towards low energy regions, and finally as the search time runs out it will only move in direction of lower energy. Hence, in order to define the annealing schedule, apart from the function form defined in eq. (98), the reduction factor α also has to be set inside the SA parameters. By setting α close to one, the search will show decreased progress. However, if enough amount of search time is available the precision level is preserved as α only affects the convergence speed. Since this parameter depends directly on the problem on which the SA method is applied on, the exact value will be obtained empirically.

Maximum number of iterations (I_{MAX})

The strength of SA method lies in its ability to statistically deliver a true global minimum. However, one of the drawbacks of this method is its slow delivery of the ideal solution once the global minimum area is located. Eventually, the localization of the global minimum area corresponds with the transformation of the true SA search, that manages uphill excursions, into the simple greedy downhill search where only configurations offering better solutions are accepted. As no indication of time required for reaching the optimal solution is available, the search process is typically suspended after the allowed search time is entirely spent. This time limit is defined through maximum number of iterations which consists of iterations corresponding to true SA search and iterations corresponding to downhill search. Even though SA search is implemented in many scientific areas, maximum number of iterations is always analyzed as one SA parameter. Still, it makes sense to split the original SA search process into two parts: one conserving the true spirit of SA search with real possibility of worse move, and other allowing only downhill movement. Consequently, the allowed number of iterations can also be divided into two parts as proposed in this thesis. This way, better control of the original SA search is accomplished as the number of true search iterations is easily extracted and better insight in SA search capabilities is obtained.

Hence, maximum number of iterations denoted as I_{MAX} is defined as:

$$I_{MAX} = I_{SA} + I_{DH} \quad (99)$$

where I_{SA} and I_{DH} are the number of true SA search iterations and downhill iterations respectively.

In the original SA search method maximum number of iterations is a function of coefficient precision, annealing schedule and transition probability function. Eventually, as ΔQ can be approximated with Gauss distribution, it would be logical to qualify the true SA search as finished when the probability of worse move of $3\sigma_{\Delta Q}$ is lower than minimum probability p_{min} expressed in percentage. Using common sense, minimum probability is defined as the case when inside 1000 worse movement inquiries only one has the realistic chance to be accepted. This case corresponds to $p_{min} = 0.1\%$. Consequently, when p_{min} is reached, the number of I_{SA} iterations is consumed and the implemented SA search enters simple downhill search mode. However, in order to preserve the minimum possibility of uphill escape, the second important change with respect to the original SA search consists in keeping the probability of worse move equal to p_{min} during the entire downhill search.

According to previous definition of p_{min} , true SA search is transformed in downhill search when the following equation is satisfied:

$$p_{min} = e^{-\frac{3\sigma_{\Delta Q}}{kT}} \quad (100)$$

At that stage, according to the annealing schedule in eq. (98) the actual temperature has been decreased L number of times:

$$T = T_L = \alpha^L \quad (101)$$

Therefore, the following holds:

$$p_{min} = e^{-\frac{3\sigma_{\Delta Q}}{k\alpha^L}} \quad (102)$$

During the annealing process, at each temperature level a number of visited candidates per level (N_{VCPL}) has been tested. Thus the actual number of iterations, corresponding to true SA search iterations can be expressed as:

$$I_{SA} = N_{VCPL} L \quad (103)$$

By extracting L from eq.(102), and inserting it in eq.(103) the following equation is obtained:

$$I_{SA} = N_{VCPL} \log_{\alpha} \left(-\frac{3\sigma_{\Delta Q}}{k \ln(p_{min})} \right) \quad (104)$$

According to eq.(97), as constant k is defined using the probability of the same Gaussian approximation, the following simplification is allowed:

$$I_{SA} = N_{VCPL} \log_{\alpha} \left[\frac{\ln(p_{max})}{\ln(p_{min})} \right] \quad (105)$$

Thus, the number of true SA search iterations is defined with temperature reduction factor α , maximal and minimal probability of worse perturbation acceptance, and a number of visited candidates per temperature level. However, as the initial probability of worse move acceptance, which corresponds to maximum probability p_{max} , is typically set to $p_{max} = 50\%$, and under previous definition $p_{min} = 0.1\%$, maximum to the number of I_{SA} iterations is approximated as:

$$I_{SA} = N_{VCPL} \log_{\alpha} (-1.7) \quad (106)$$

Since N_{VCPL} and α are closely related to the specific problem, the correct numerical value of I_{SA} parameter will be known only after those two parameters are empirically set to accurate values.

By defining maximum number of iterations in accordance with eq.(99), this SA parameter is set explicitly as I_{MAX} , or implicitly over I_{SA} and I_{DH} . When it is set explicitly, then first I_{SA} requirements are settled and the remaining number of iterations is left to I_{DH} .

Conclusion

In previous section the SA search algorithm together with its parameters was described in details. Even though, the majority of parameters should be selected based on the empirical observations, many of them have been set to typical values, or over-dimensioned ones favouring precision over convergence speed. Thus, the set of SA parameters obtained so far, can be regarded as theoretical parameter set since fine tuning will be made in the following section. Fine tuning is based on the various SA search simulations applied on strong and moderate

coupling conditions where the correct set of empirical values is established as a trade off between speed and precision.

<i>Typical Values</i>	<i>Standard Functions</i>	<i>Overdimensioned parameters</i>
$Res_{symb} = 0.00001$	$Coef_{new}(i) = \xi_i R_0 + Coef_{act}(i)$ $i = 1..20; \xi_i \in \langle -1, 1 \rangle$	$N_{VCPL} = 300$
$Res_{coeff} = 0.00001$	$Q = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_{1i} - \hat{x}_{1i})^2}$	$\alpha = 0.999$
$Coef_0(i) = \begin{cases} 1, & i = 1 \\ 0, & i = 2, \dots, 20 \end{cases}$	$p(\Delta Q) = e^{-\frac{\Delta Q}{kT}}$	$R_0 = 0.05$
$p_{max} = 50\%$ $p_{min} = 0.1\%$	$k = -\frac{3\sigma_{\Delta Q}}{\ln(p_{max})}$	$I_{MAX} = I_{SA}$ $I_{DH} = 0$
	$T_L = \alpha T_{L-1}$ $T_0 = 1, \alpha < 1, L = 1, 2, 3\dots$	
	$I_{SA} = N_{VCPL} \times \log_\alpha (-1.7)$	

Table 4: Theoretic set of SA search parameters before empirical analysis

The theoretical set of SA parameters before empirical analysis is presented in Table 4 where typical values, standard functions and over-dimensioned parameters are shown. In order to adopt the SA search to decoupling problem the over-dimensioned parameters are analyzed in the first place. Later, some of standard functions and typical values are changed to accomplished better speed results and retain the precision.

6.1.2 Fine tuning of SA parameters

The fine tuning of SA parameters is done for 64QAM signals as the most demanding constellation when compared to 16QAM or QPSK constellations. The results are based on 2000 calibration symbols transmitted under SNR of 100dB. This way the analysis is concentrated primarily on coupling rather than on noise distortions. The number of maximum iterations is set implicitly and is equal to I_{SA} . Number of true downhill iterations is set to zero ($I_{DH} = 0$) which sets the attention on true SA search rather than on downhill headless run. Since choosing a new candidate depends on the Gaussian probability function, the solution of SA search also carries probability factor. Thus, each presented simulations is in fact executed ten times, and the figures present the averaged values. The first set of simulations concentrates on finding an adequate pair of (N_{VCPL}, α) . Afterwards, the obtained adequate parameters are used for obtaining the appropriate maximum allowed displacement factor R_0 . Once these three factors are located, the emphasis is transferred to the cost function type. Two types of cost function are analyzed and one presenting best performance regarding the system BER performance, convergence speed and processing load is selected for the further simulations. Eventually, fine tuning is performed on factor corresponding to the required number of iterations as the analysis concentrates on I_{DH} factor. This factor does not influence directly the search process but defines the search finishing condition and as such is less important. Finally, the search based on adjusted SA search together with numerical solution to one inverse coupling approximation is presented.

$$\Rightarrow (N_{VCPL}, \alpha)$$

In theory, when keeping the number of visited candidates per level constant, the search is done more thoroughly if the temperature is decreased more slowly. Hence, better solution precision is achieved. On the other hand, if temperature reduction factor is constant, and the number of visited candidates per temperature level is increased, each level is scanned in more detail and, again, better results are expected. However, when both factor are set in favour of better precision the search is extremely slowed down and the precision improvement might not justify the processing load. In Figure 54, minimum cost function value is presented in dependence of temperature reduction factor and different number of visited candidates per level for moderate coupling distortion of $[-10, -12, -12]$ dB and maximum allowed displacement set to $R_0 = 0.05$. By setting the initial point to $Coef_0$ the cost function value corresponding to

approximately 0.069 is obtained. The simulations show that regardless of the N_{VCPL} number, the improvement of cost function minimum is unnoticeable for $\alpha = 0.3$ as it stays close to the initial cost function value. This occurs because the temperature is decreased too quickly. However, as α tends to one, the minimum of cost function is visibly decreasing except for the $N_{VCPL} = 2$ where the search scanning is not wide enough and the algorithm cannot progress adequately.

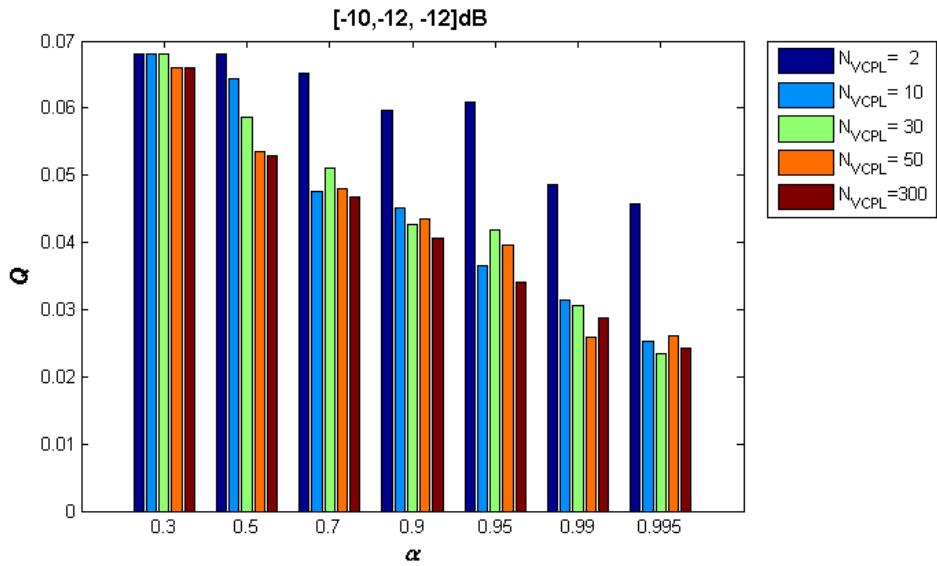


Figure 54: Q as a function of α and N_{VCPL} for $[-10, -12, -12]$ dB coupling distortion

The dependence of the cost function minimum on the number of visited points per level is seen more clearly in Figure 55, where bars of different colours correspond to different temperature reduction factors. With only two new candidates per level, the minimum cost function actually corresponds to the initial cost function value for $\alpha = 0.3$, $\alpha = 0.5$ and $\alpha = 0.7$. As α gets closer to 1, according to eq. (105) maximum number of iterations is increased from 4 for $\alpha = 0.3$ to 44 for $\alpha = 0.9$. Thus the probability of finding the better solution also increases, and as a result minimum cost function moves away from its starting value. However, even when α is increased the scan space around the actual minimum is always kept narrow. The search is easily stacked in local minimum and no progress can be made. This is seen in Figure 55 where for $N_{VCPL} = 2$ no significant progress is made even with $\alpha = 0.995$. According to these results, the number of visited points per level starts producing good results with $N_{VCPL} > 10$.

When calibration signals are submitted to strong coupling distortion like $[-8, -10, -8]$ dB the behaviour of the SA search precision follows the same interpretation as for moderate coupling.

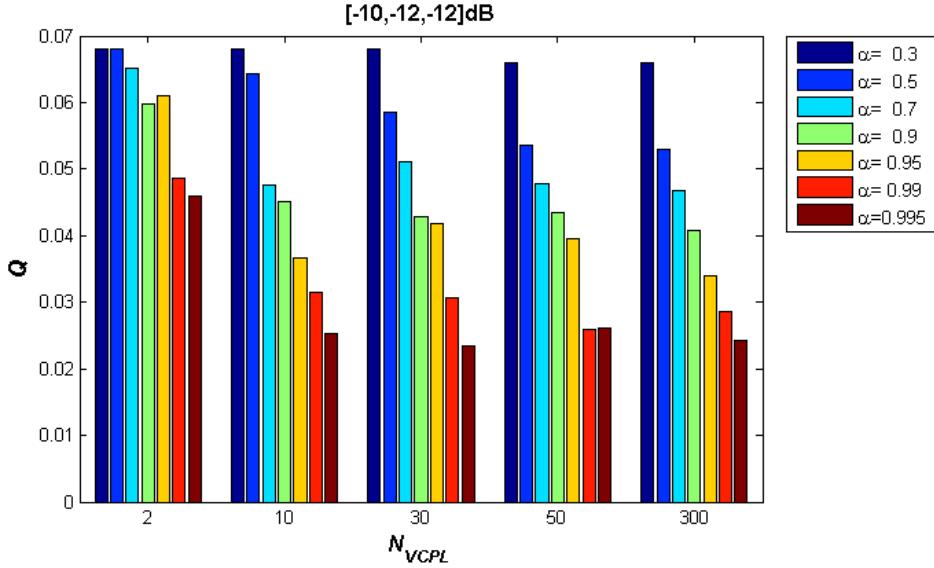


Figure 55: Q as a function of N_{VCPL} and α for $[-10, -12, -12]$ dB coupling distortion

The differences worth mentioning include the cost function minimum which, under same SA search parameters, is higher compared to the moderate coupling environment. Even the initial Q value is increased from 0.034 for moderate to 0.075 for strong coupling. This is expected since global minimum location in case of stronger coupling is situated further away from the starting point and larger number of iterations is required to reach it.

This is seen in Figure 56, where Q is presented as a function of α and N_{VCPL} , keeping the maximum allowed displacement set to $R_0 = 0.05$. In these simulations minimum N_{VCPL} number is increased from 2 to 5, but as coupling is also stronger, the SA search with $N_{VCPL} = 5$ permanently presents the solutions with lower precision. This means that minimum number of visited candidates per level should be higher than 5. On the other hand, the precision difference between simulations with $N_{VCPL} = 300$ and $N_{VCPL} = 50$ are rather small for both coupling environment. Furthermore, in some cases like $\alpha = 0.9$ for $[-8, -10, -8]$ dB distortion, the search with smaller number of visited candidates even obtained better precision results. As a conclusion the maximum number of visited candidates per level should also be lower than 50.

The temperature reduction factor showed poor performance in both set of simulations for $\alpha = 0.5$ which was expected since maximum number of iterations in this case is very low and the search can be described more as a lucky guess than any type of search algorithm. For α between 0.5 and 0.95 the search progress presented low consistency as the cost function

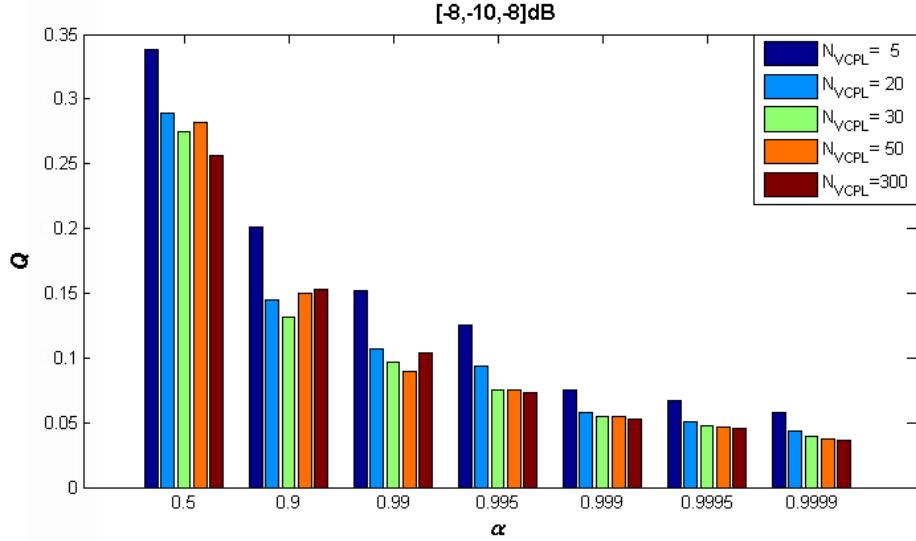


Figure 56: Q as a function of α and N_{VCPL} for $[-8, -10, -8]$ dB coupling distortion

minimum experienced big differences between two realizations of the same SA search. This shows that the temperature reduction is still reduced too quickly and that the final precision is still not guaranteed as search robustness is not achieved. This statement is confirmed by comparing the Q function obtained with different N_{VCPL} in this α range. For $\alpha = 0.95$ and coupling distortion set to $[-10, -12, -12]$ dB better precision is achieved with $N_{VCPL} = 10$ than with $N_{VCPL} = 50$. The same phenomena is presented in strong coupling environment where for $\alpha = 0.90$ better precision is obtained with $N_{VCPL} = 30$ than with $N_{VCPL} = 300$. Hence in order to avoid the incertanty, temperature reduction factor should be between 0.95 and 1 since this area shows signs of consistency with theoretical background. That is, as seen in figure 56, inside this area the precision is higher as α and N_{VCPL} are increased. On the other hand, as α enters the area between 0.999 and 1, the precision increase is relatively small regardless of the N_{VCPL} number but maximum the number of iterations is largely increased (from 14000 for $\alpha = 0.995$ to 690000 for $\alpha = 0.9999$ and $N_{VCPL} = 30$). This can be interpreted as indication of upper precision limit of temperature reduction factor making the factors $\langle 0.999, 1 \rangle$ not recommendable for the SA search of inverse coupling function if the required approximation precision can be obtained by adjusting other SA parameters. According to the presented simulation results, the SA search is adopted adequately to inverse coupling approximation with the number of visited candidates per level set to $N_{VCPL} = 30$, and temperature reduction factor set to $\alpha = 0.995$. These parameters will be used in following simulations.

$\Rightarrow R_0$

Since the number of visited points per level is defined with N_{VCPL} , and maximum displacement of each point is restrained to R_0 , then maximum search area at one temperature level of each factor is restricted to $R_0 \times N_{VCPL}$. According to the modification of SA algorithm, this area represents the maximum displacement from the actual minimum configuration during the entire search process. The SA modification corresponds to changing the actual configuration, at the beginning of each temperature level, to the one obtaining the minimum cost function for the moment. Inside the first set of simulations maximum displacement factor was over-dimensioned to $R_0 = 0.05$ in order to allow the wide area search and avoid the local minimum traps.

By setting the number of visited candidates per level to $N_{VCPL} = 30$, maximum displacement is set to 1.50 (30×0.05). In fact, since expected values of all approximated surface coefficients are between $\langle -1, 1 \rangle$, except for the first coefficient which is expected between $\langle 0, 2 \rangle$, this space is actually larger than the whole expected solution space. Hence, any potential local minimum is easily avoided as SA algorithm can not get stucked during the search. However, this also means that the search is not conducted adequately as no clear search path can be established. By adequately setting the maximum displacement factor the search should wander less and present better precision results. The analysis for adequate maximum displacement factor should be made for strong coupling first as in this case the ideal solution is located further from the starting point and any restriction on R_0 parameter should be noticed earlier. The conducted search is based on typical SA algorithm with $N_{VCPL} = 30$ and $\alpha = 0.995$.

The behaviour of cost function in dependence of R_0 under strong coupling conditions of $[-8, -12, -10]$ dB is presented in Figure 57. The figure can be divided into three clearly distinguished areas. The left side of the figure, with R_0 lower than 10^{-5} corresponds to algorithm where no wandering around is allowed. The search progress is slow and any encountered local minimum is presented as the final solution. The ideal solution is only found if the search is executed close to it. On the other hand, the right side of the figure represents the R_0 values larger than 0.5. In this area wandering is allowed during the whole search process and as a consequence no clear search path is established. The search tends to move too quickly and without any exploration around the actual position. As search is randomly wandering around, finding the ideal solution in this area is highly uncertain. The obtained solutions are still slightly better on the left side of the image than on the right side ($Q = 0.09$ vs $Q = 0.11$) as the ideal solution is obviously found close to the search starting point.

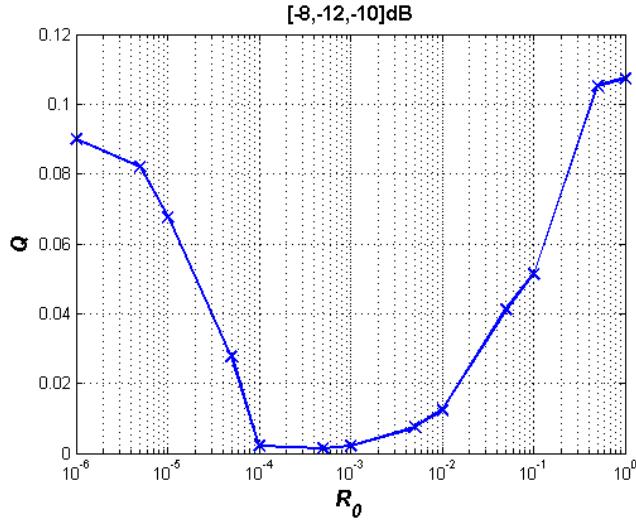


Figure 57: The behaviour of cost function in dependence of R_0 for strong coupling

The central part of the image coincides with the meaningful SA search where wandering around and detail space exploration are in balance. Here, the search method does follow some path and the traps of local minimums are successfully avoided. As a result the provided cost function is decreased by two orders of magnitude from approximately $Q = 0.11$ obtained in the first two image parts, to $Q = 0.0014$ for $R_0 = 0.0005$. Based on the presented Q values, the adequate maximum allowed displacement for strong coupling is between $\langle 0.0001, 0.01 \rangle$.

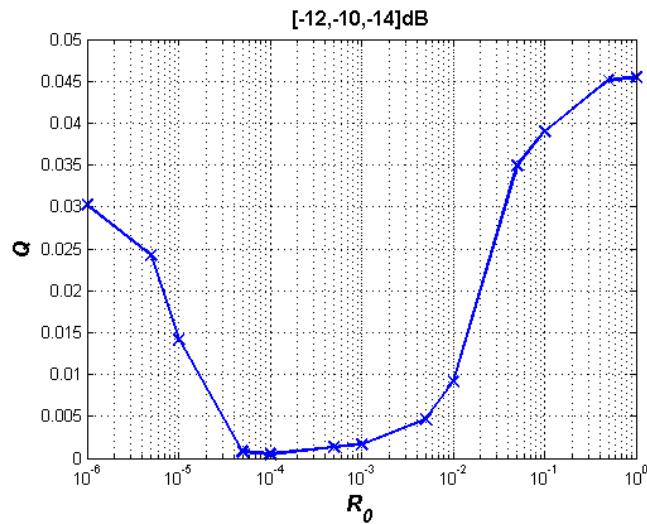


Figure 58: The behaviour of cost function in dependence of R_0 for moderate coupling

In Figure 58, the analysis of adequate R_0 factor is conducted in moderate coupling system of $[-12, -10, -14]$ dB. The figure in fact looks like a copy of maximum displacement behaviour under strong coupling conditions conserving the form of written letter “U” with three different solution areas. The only difference is found in the altitude of the image corners which are reduced on the left image side from $Q = 0.11$ under strong coupling to approximately $Q = 0.045$ under moderate coupling conditions, and from $Q = 0.09$ to $Q = 0.03$ on the right side. This is explained with the search starting positions which favours lower coupling and is thus located closer to the ideal solution. As increased wandering contributes rather little to Q value improvement when coupling distortions are not strong the left side of the image again presents slightly better results. The best Q value is experienced for R_0 between $\langle 0.0005, 0.005 \rangle$ which almost matches the area obtaining minimum cost value under strong coupling. Surprisingly, the bottom of “U” is numerically rather close to the bottom obtained as the solution to stronger coupling distortion.

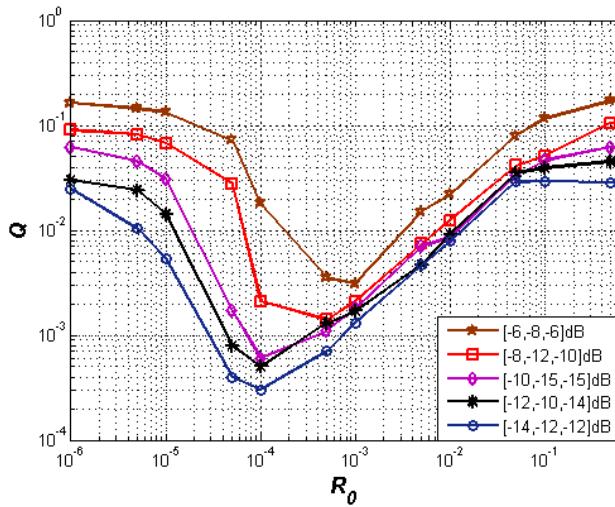


Figure 59: Q in dependence of R_0 under different coupling conditions

This Q function minimum positions are seen more clearly in Figure 59, where the value of cost function is shown as a function of maximum allowed displacement under different coupling conditions in a logarithmic scale. This figure in fact confirms that under adequate selection of SA search parameters the quality of the solution is preserved regardless of the suffered distortion. The obtained minimum cost function varies from $Q = 0.0007$ for $[-14, -12, -12]$ dB to $Q = 0.003$ for $[-6, -8, -6]$ dB while the optimal R_0 selection varies from 0.0001 to 0.001 under the same coupling conditions. This reaffirms that lower coupling conditions are managed

better using smaller R_0 , as their optimal solution is close to the search starting point. As a conclusion, and in accordance with Figure 59, maximum allowed displacement is set to $R_0 = 0.0005$ as a part of the adjusted SA search parameters.

$\Rightarrow Q$

The correct functioning of the cost function is confirmed when cost function decrease corresponds to the system performance enhancement. This test consists of two steps. In first step, SA search is implemented on the user defined coupling strength. Unlike standard SA search which produce the inverse surface coefficients at the end, this time a corresponding set of actual coefficients, together with their cost function value, are stored and at the end of each temperature level with maximum number of iterations limited to $I_{MAX} = 20000$. These stored values are used in second step for determining the relation between BER and the corresponding cost function level. In order to see the behaviour of BER in dependence of Q function, system environment is set to fixed noise level and the simulation of real system under AWGN noise is carried out. As signal transmission under BER of 10^{-4} is usually defined as sufficiently good for data traffic, the noise level permitting this BER under ideal conditions is used for the system environment set up. According to [Lu99], this BER is established at $E_b/N_0=16.5\text{dB}$ when 64QAM constellation is used.

The first analysis is based on quadratic mean square function defined in eq.(85) as:

$$Q_2 = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_{1i} - \hat{x}_{1i})^2}$$

where N is the number of symbols used for calibration, x_{1i} is the sent i^{th} calibration symbol on first antenna, and \hat{x}_{1i} its corresponding value calculated with coefficients obtained in first simulation step. Figure 60 shows the behaviour of BER performance as a function of Q_2 , under moderate coupling of $[-10, -10, -12]\text{dB}$. Under these conditions, BER first experiences linear decrease and than enters the saturation area close to ideal BER level of 10^{-4} . The linear BER improvement is clearly seen in the first part of curve as BER drops from 0.02 to 0.0002 following Q_2 which is decreased from 0.068 to 0.011. This enhancement is accomplished at almost constant pace, with very small deviations. As BER approaches its ideal value the saturation is bound to occur since system performance can not be improved beyond its limiting

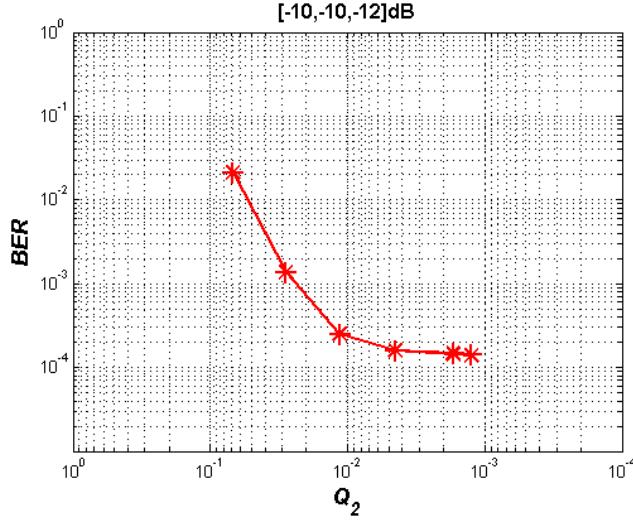


Figure 60: The behaviour of BER as a function of Q_2 under strong coupling

value. Thus, even though cost function is reduced and SA search produces various improved surface approximations BER performance is kept at the same level. In this case, when Q_2 is lower than 0.005 BER can be described with almost flat line. The robustness of cost function is here confirmed with small BER deviations around this flat line for various Q_2 values between 0.005 and 0.001. If excellent decoupling conditions are described with $BER < 2 \times 10^{-4}$ than when Q_2 is lower than 5×10^{-3} the corresponding cost function can be regarded as sufficiently low to guarantee excellent decoupling conditions and SA search can be abandoned.

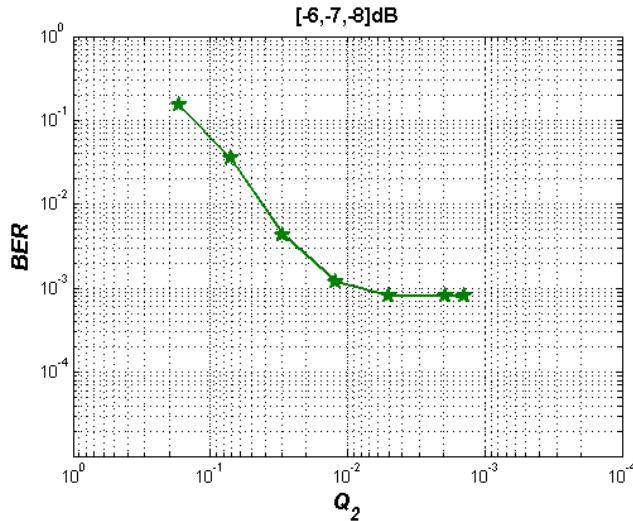


Figure 61: The behaviour of BER as a function of Q_2 under moderate coupling

The behaviour of quadratic mean square function under strong coupling of $[-6, -7, -8]$ dB is presented in Figure 61. Apparently the dependency is very similar to the previous case as it consists of linear reduction and saturation part. However, even though the saturation appears for same value of cost function of approximately $Q_2 = 0.01$, the corresponding BER performance is decreased from 0.0001 to 0.0008. The saturation occurs approximately at the same level of Q_2 , as under moderate coupling, which is close to $Q_2 = 0.005$. Apparently, at this stage of Q_2 , it makes no sense to continue with the search process as no significant improvement can be expected. However, under strong coupling conditions the system is drawn into saturation before reaching BER close to theoretical value which makes impossible entire signal decoupling.

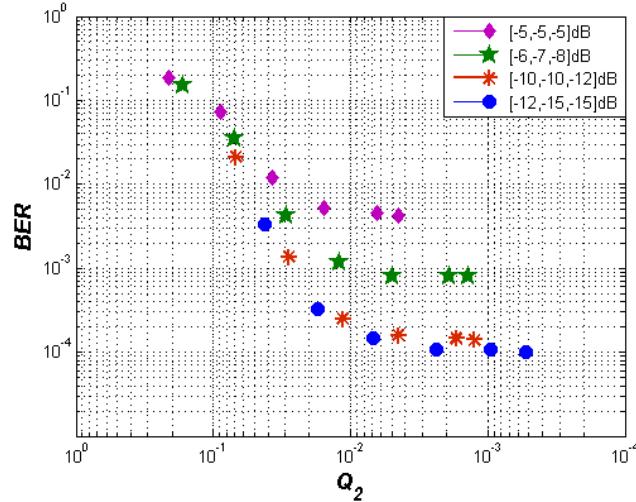


Figure 62: BER as a function of Q_2 under different coupling levels

The saturation area of BER performance is seen more clearly in Figure 62 which consists of BER- Q_2 dependency obtained under four different coupling levels. The figure shows that SA search based on Q_2 undoubtedly leads towards better BER performance. Apparently as coupling is increased the saturation level of BER also changes from $BER = 0.0002$ for $[-10, -10, -12]$ dB to $BER = 0.004$ for $[-5, -5, -5]$ dB. As this level of BER can not be enhanced any further it can be deduced that quadratic mean square function can not offer good BER performance under all coupling conditions.

The saturation doesn't appear always at the same level of Q_2 , as it starts to appear when Q_2 reaches a value of 0.01 under $[-5, -5, -5]$ dB coupling, and it appears only beyond $Q_2 = 0.005$ under $[-12, -15, -15]$ dB. Under moderate or low coupling conditions the saturation area

might appear after $Q_2 = 0.005$ but since this level appears inside the zone which corresponds to excellent decoupling conditions the search can be stopped when 0.005 is reached. If the search is executed under strong coupling conditions, the saturation area starts before $Q_2 = 0.005$ thus the search can be stopped at this saturation point without heavy consequences. Thus, in order to reduce the processing load all search algorithms based on quadratic mean square function can be stopped at level $Q_2 = 0.005$. However, since saturation area still exhibits small downhill inclination with regard to BER, by favouring precision the processing load of SA approach is not constrained to this cost function level.

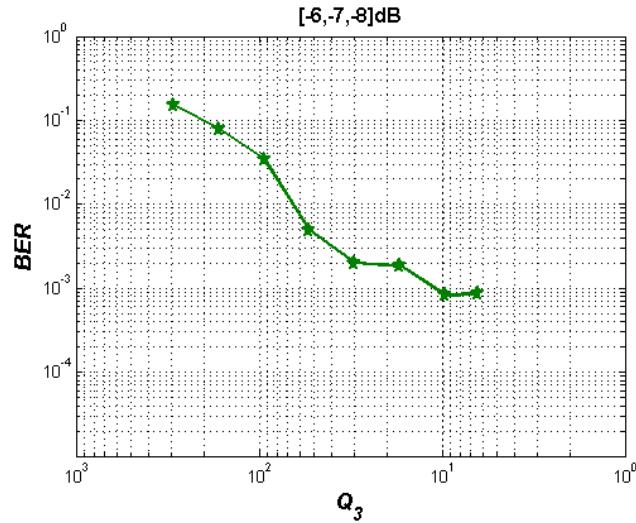


Figure 63: The behaviour of BER as a function of Q_3 under strong coupling

The second part of simulations consists of cost function based on the sum of absolute differences, as defined in eq.(86):

$$Q_3 = \sum_{i=1}^N |x_{1i} - \hat{x}_{1i}|$$

where the same symbol indexing is used as for the quadratic mean square function. From the point of view of processing requirements, Q_3 is far more attractive than Q_2 as in Q_2 each iteration is followed by N squares, and one square root calculation. The mathematical difference between the two presented cost functions lies in weighing factor, as Q_2 gives more importance to larger amplitude differences than Q_3 . Figure 63 shows the behaviour of BER performance based on Q_3 cost function under strong coupling conditions of $[-6, -7, -8]$ dB. At the beginning the dependence is linear with the angle of inclination lower than achieved with

Q_2 which suggest that better guidance is obtained with Q_2 in this area. Later, in the central part of the curve the inclination similar to Q_2 is obtained but it is not kept for long as clear inclination tendency can not be established. Much like Q_2 , the performance of the sum of absolute differences is limited to $BER = 10^{-3}$ meaning that precision of final solution is preserved when compared to Q_2 . Even though the progress path of Q_3 is not well defined this cost function does lead the SA search algorithm towards the same solution obtained with more demanding Q_2 search and can consequently be used under strong coupling conditions.

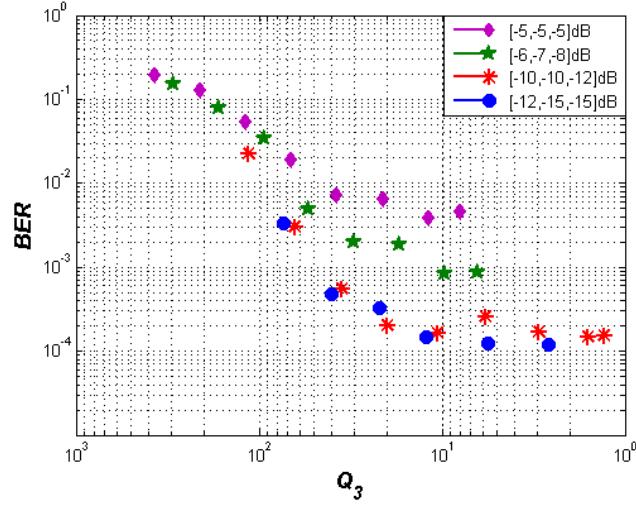


Figure 64: BER as a function of Q_3 under different coupling levels

The system performance of Q_3 cost function is considerably changed when moderate coupling conditions are analyzed. This is presented in Figure 64 where BER performance in dependence of Q_3 cost function under two strong and two moderate coupling level is depicted. Under moderate coupling levels BER falls rapidly at the beginning in both cases, but each coupling environment presents one digression from the established tendency. For example under $[-10, -10, -12]$ dB coupling BER obtained with $Q_3 = 6$ exhibits worse performance than the one obtained with $Q_3 = 20$. Similar situations are found under $[-12, -15, -15]$ dB coupling where BER between $Q_3 = 40$ and $Q_3 = 20$ is very similar but the difference between $Q_3 = 20$ and $Q_3 = 13$ falls rapidly. This indicates that even though the sum of absolute differences eventually does lead towards better solution, its robustness is quite low and smaller Q_3 values don't guarantee better solutions. This, sometimes erratic BER performance behaviour lead to decreased confidence.

As a conclusion, the simulations show that both cost functions lead the search towards better

solutions. The search based on quadratic mean square function exhibits higher robustness and its BER performance presents better consistency under different coupling situations. Low processing load of Q_3 does not compensate for the inconsistency and quality loss under moderate coupling when compared to Q_2 . Thus quadratic mean square function is selected to be implemented inside the decoupling module as overall better solution.

$$\Rightarrow (I_{MAX})$$

Maximum number of permitted iterations consists of true SA search iterations (I_{SA}) defined with SA search parameters and of plain downhill iterations (I_{DH}) which do not influence the search process but do affect the precision of obtained solutions. As defined in eq.(99) maximum number of iterations is defined as a sum of these two parameters:

$$I_{MAX} = I_{SA} + I_{DH}$$

According to eq.(106) SA search iterations are calculated as:

$$I_{SA} = N_{VCPL} \log_{\alpha} (-1.7)$$

Since N_{VCPL} and α are adopted to decoupling problem for $N_{VCPL} = 30$ and $\alpha = 0.995$, the number of true search iterations is set to $I_{SA} = 13770$. The number of downhill iterations can not be calculated in deterministic manner and is selected through simulations. Theoretically, if SA search is carried out correctly it should transform into the downhill algorithm in area close to the ideal solution. This means that the decoupling solution should than be located fast and in small number of iterations. This is seen in Figure 65, where Q is depicted as a function of the number of iterations under strong coupling conditions of $[-6, -8, -6]$ dB. The simulation is carried out using the SA search parameters whose values are established in previous sections.

The results of cost function show that the initial cost function value is significantly improved during the true SA search as it is decreased from $Q_0 = 0.15$, to $Q_{13770} = 0.004$. Afterwards, the downhill search started and after approximately 6000 downhill iterations the SA search presents stable solution around $Q_{SA} = 0.0015$. Nevertheless, if downhill search is still continued, the additional solution improvements are obtained which is confirmed with small downhill inclination of Q function after 20000 iterations. However, in this case the required number

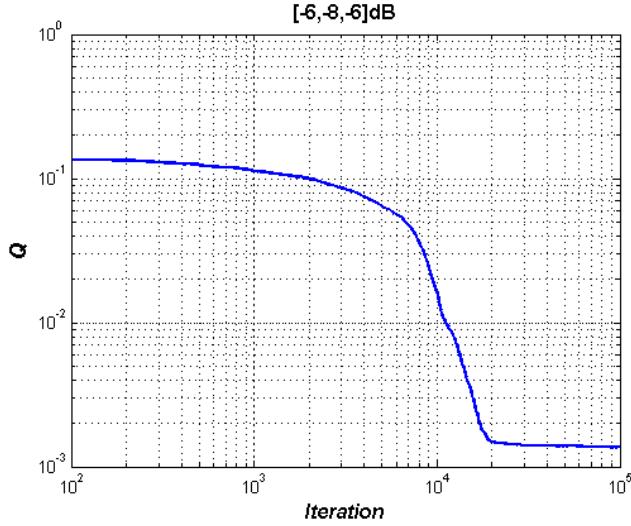


Figure 65: Cost function in dependence of the number of iterations under [-6,-8,-6]dB

of iterations for the solution improvement is measured in tens of thousands of iterations and the obtained improvement does not justify the processing load. For example, the cost function is reduced from $Q_{20000} = 0.0015$ to $Q_{100000} = 0.0014$ based on 80000 iterations, which means that the number of I_{DH} iterations is almost 5 times higher than the number of I_{SA} iterations. Furthermore, as seen in previous section the reduction of cost function beyond $Q = 0.005$ doesn't lead to better system BER performance and the search should stop at this

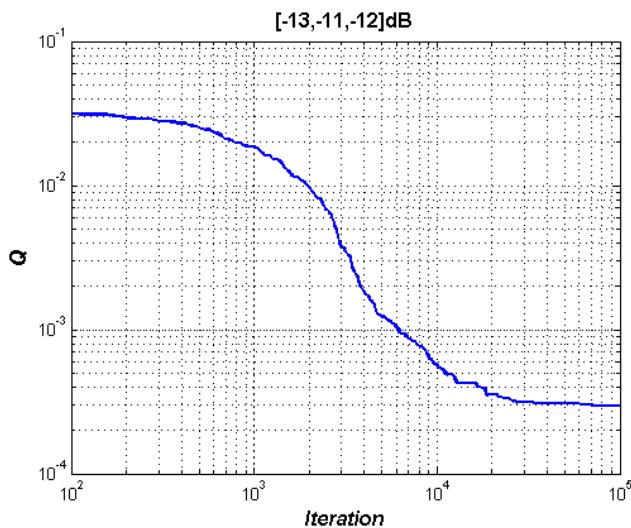


Figure 66: Cost function in dependence of the number of iterations under [-6,-8,-6]dB

level. Thus, in this case only a few thousand of downhill iterations is required to obtain a decoupling system with excellent BER performance. The cost function dependence changes rather little if maximum number of iterations is analyzed under moderate coupling conditions as shown in Figure 66.

In this case, the true SA search leaves the algorithm even closer to the ideal solution and downhill search can not contribute significantly to solution improvement. The true SA search drives the cost function from initial value $Q_0 = 0.03$ to $Q_{13770} = 0.0004$. According to previous analysis of quadratic mean square cost-function, the excellent improvement of system BER is achieved under moderate coupling for cost function lower than 0.005. Actually, in this case the downhill search cannot offer any BER improvement as the cost function has already entered the saturation area therefore downhill search is eventually redundant.

As SA search should produce the most accurate decoupling solutions under all coupling levels the downhill search should also be included since it produces better solution results under strong coupling conditions. Thus the SA search implemented inside the decoupling module is restrained to $I_{SA} = 13770$ and $I_{DH} = 2230$ iterations, which sets the maximum number of permitted iterations to $I_{MAX} = 16000$.

⇒SA Search implementation example

The SA search is conducted under moderate coupling conditions of $[-10, -12, -10]$ dB using the SA search parameters adopted to nonlinear decoupling problem. The behaviour of adopted SA search algorithm is shown in Figure 67, where cost function is presented in dependence of number of iterations for 10 different SA search executions. The image shows that the search path in first 500 iterations is similar in all ten instance. The curves than starts to spread, and are grouped again at the end. Nevertheless, the cost function solutions are kept in the same orders of magnitude throughout the entire search process. The small differences between the temporary cost function progress, and stable final solution precision confirm that the SA parameters have been selected adequately as the search successfully avoids local minimum traps and follows the correct search path. Furthermore, since BER performance enters saturation for $Q_2 < 0.005$ the presented search can be stopped at 4000 iterations preserving the decoupling precision.

Eventhough the obtained cost function value is similar for all ten instances of conducted SA search, the presented decoupling coefficients numerically offer completely different solutions.

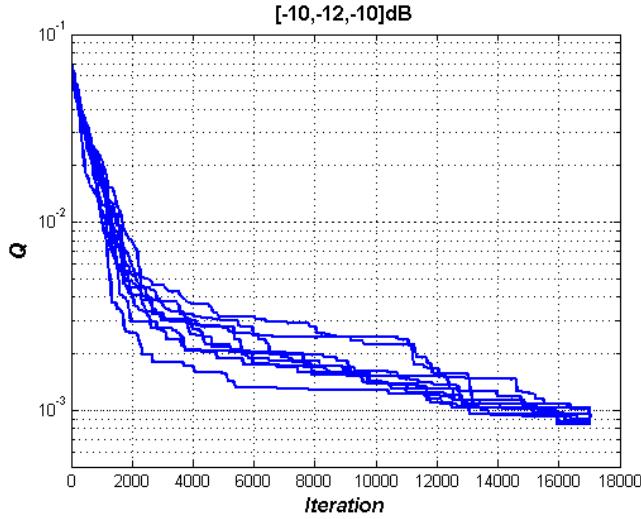


Figure 67: Ten instances of adopted SA search algorithm under moderate coupling

Two different sets of 20 decoupling coefficients, obtained as a result of two different SA searches under same coupling conditions, are presented in Figure 68. The difference between the cost functions in this case is almost zero as $Q_a = 0.0009$ and $Q_b = 0.0008$, which means that the constructed inverse coupling surfaces are much alike. However, they are based on different building material as for example a_7 is very close to zero in first solution while in second one is $a_7 = 0.015$. On the other hand a_5 and b_3 show opposite behaviour, as they present higher importance in second than in first solution. Furthermore, the amplitudes of decoupling coefficients don't exhibit the tendency of reduction as the exponential order is increased since, for example, $a_8 \gg a_3$ or $b_7 \gg b_3$. This indicates that the surface simplification based on

a₁	a₂	a₃	a₄	a₅	a₆	a₇	a₈	a₉
1,00840	0,00468	0,00433	0,00218	0,00245	-0,02087	0,00018	0,01541	-0,00384
b₁	b₂	b₃	b₄	b₅	b₆	b₇	b₈	b₉
-0,09903	-0,01802	-0,01042	-0,00752	0,00664	0,02397	-0,01103	-0,01454	0,00752
k₁	k₂							
0,00237	0,00002							

a₁	a₂	a₃	a₄	a₅	a₆	a₇	a₈	a₉
1,01160	-0,00424	0,00040	0,00448	-0,01285	0,01531	0,01581	-0,01373	-0,00417
b₁	b₂	b₃	b₄	b₅	b₆	b₇	b₈	b₉
-0,10036	-0,01088	-0,00315	-0,00896	-0,01443	-0,00261	0,01502	0,00645	-0,00351
k₁	k₂							
0,00298	-0,00001							

Figure 68: Two sets of decoupling coefficients for [-10,-12,-10]dB coupling environment

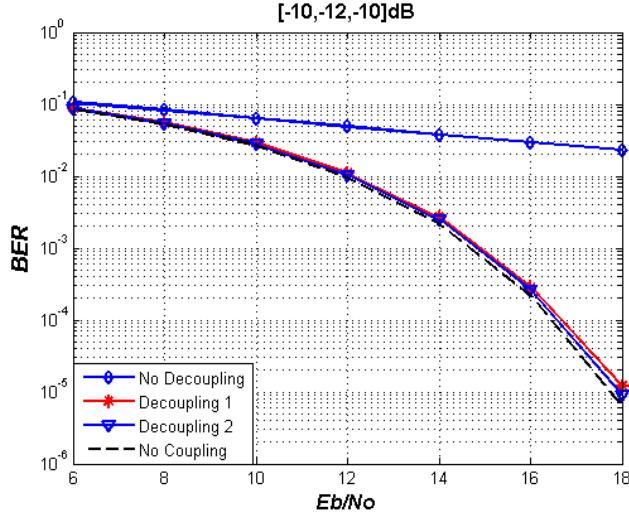


Figure 69: Performance of SA based decoupling module under moderate coupling

simple reduction of order of polynomials might not achieve the same decoupling efficiency. Nevertheless, this kind of simplification will be analyzed in next chapter.

Still, the system performance of decoupling module based on the obtained coefficients traces almost the same curve inside E_B/N_0 -BER graph as shown in Figure 69. The upper curve on the image presents the system behaviour under coupling of $[-10, -12, -10]$ dB and no decoupling module. The lowest curve shows the theoretical system performance without any distortion and only with AWGN. The two curves in the middle, exhibit the system performance with the decoupling module.

Since the processing load of SA algorithm is not restrained to $Q_2 < 0.005$, the performance of system with decoupling module is eventually very close to the theoretical system performance. This confirms that the cost function and inverse surface model are both selected adequately and that SA search produces accurate results.

⇒Conclusion

The presented SA search method is based on the physical simulated annealing process from metallurgy. When compared to the original SA search method the proposed algorithm restrains the wandering space at every temperature level. This way the search progress is accelerated

and the solutions close to the initial search point are favoured which is in accordance with the expected coupling environment.

The proposed SA search method is adopted to nonlinear coupling problem based on empirical analysis performed through different set of simulations. The adopted set of SA search parameters is presented in Table 5. The table consist of two columns with standard values based on

<i>Standard Values</i>	<i>Values based on Empirical Decoupling Analysis</i>
$Res_{symb} = 0.00001$	$N_{VCPL} = 30$
$Res_{coeff} = 0.00001$	$\alpha = 0.995$
$Coeff_0(i) = \begin{cases} 1, & i = 1 \\ 0, & i = 2, \dots, 20 \end{cases}$	$R_0 = 0.0005$
$I_{SA} = N_{VCPL} \times \log_{\alpha} (-1.7)$	$Q = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_{1i} - \hat{x}_{1i})^2}$
$Coeff_{new}(i) = \xi_i R_0 + Coef_{act}(i)$ $i = 1..20; \xi_i \in \langle -1, 1 \rangle$	$I_{MAX} = 16000$ $I_{SA} = 13370$ $I_{DH} = 2230$
$p(\Delta Q) = e^{-\frac{\Delta Q}{kT}}$ $k = -\frac{3\sigma_{\Delta Q}}{\ln(p_{max})}$	$T_L = \alpha T_{L-1}$ $T_0 = 1, \alpha < 1, L = 1, 2, 3\dots$
$p_{max} = 50\%$ $p_{min} = 0.1\%$	

Table 5: Set of SA search parameters after empirical analysis

theoretical assumptions occupying the left column, and values based on empirical decoupling analysis in the right column.

The performance of decoupling module based on the proposed SA search method shows stable behaviour and excellent decoupling results. Eventually, as SA search should guarantee solution precision under all coupling conditions, the number of iterations is large. Since it requires high processing power, this approach can not be used in slow and battery dependant devices like mobile phones and will be presented only as the precision reference to other two decoupling search approaches.

6.2 Improved Fast Simulate Annealing (IFSA)

The strength of SA approach lies in its ability to statistically find true global minimum regardless of the problem type or form. However, the fact is that SA search is very slow when compared to other heuristic methods and is often avoided as more importance is given to speed than to precision. Hence, the original SA algorithm is rarely used, but it is worth mentioning some of the known SA enhancements that have been proposed in numerous articles [Ing89, Ros92, Ren96, Men97, Fin04]:

- Quantum Annealing (QA)
- Simulated Quenching (SQ)
- Mean Field Annealing (MFA)
- Fast Simulated Annealing (FSA)
- Parallel Recombinative Simulated Annealing (PRSA)
- Adaptive Simulated Annealing (ASA)

All of these methods follow the same pseudo code presented for original SA approach in Figure 52. However, in contrast to the original SA approach these methods have been characterized as fast search algorithms. The increase in speed is naturally paid off in precision, but as the precision is still kept at high level, the loss is acceptable compared to the speed gain.

The proposed Improved Fast SA (IFSA) approach , as the proper name indicates, belongs to fast search algorithms, but can be described more like a reformation than enhancement of the original SA algorithm since the changes include all three vital parts of SA method:

1. annealing schedule
2. probability transition function
3. maximum allowed displacement.

The IFSA implements the annealing schedule function that has been used in several SA enhanced methods like ASA or SQ, while the probability transition function actually corresponds to FSA, and hence the name Improved FSA approach. Still, up to now, all modified SA methods, keep maximum allowed displacement constant while in IFSA this parameter is also a function of temperature.

Eventually, all enhanced SA methods labeled as fast are based on steeper annealing schedule function when compared to traditional SA approach. This is logical since annealing schedule function carries the major weight in convergence speed while the rest of the search parameters are more associated to the precision and local minimum avoidance. As seen in eq. (98), the original SA approach is based on the exponential annealing schedule function where the temperature level T_L , is decreased by a fixed factor α during the entire search process:

$$T_L = \alpha T_{L-1}; \quad T_0 = 1, \quad 0 < \alpha < 1, \quad L = 1, 2, 3\dots$$

where T_L and T_{L-1} are the new and the actual temperature respectively, and index L stands for the number of temperature level changes. According to this annealing schedule, the temperature is decreased at constant pace during the entire search process. If this decrease rate is too fast, the search can get stuck in a local minimum and the algorithm loses much of its authority. On the contrary, if the decrease rate is too slow, the processing time is large and approach is usually replaced with some faster method. However, it makes sense to decrease the temperature more rapidly at the beginning when wandering is desirable, and then reduce the temperature carefully towards the end when SA turns into a greedy search algorithm. This is precisely the effect achieved with the following annealing schedule function used in IFSA approach:

$$T_L = \frac{T_\alpha}{\ln(L+1)}; \quad L = 1, 2, 3\dots \quad (107)$$

where L is the actual temperature level and T_α is the temperature constant whose numerical value is actually irrelevant to performance of IFSA search. In real physical process factor T_α controls the initial temperature and as such it influences the probability of uphill movement during the search as this probability is related to temperature with Boltzmann constant k_B . Since decoupling problem is not related directly to the physical energy this constant is replaced with problem specific constant k_c which is adjusted according to the initial temperature level as will be shown in the analysis of the uphill probability movement. In order to easily compare the SA to IFSA search T_α is chosen to equal the initial temperature level of SA search algorithm:

$$\begin{aligned} T_1^{SA} &= T_1^{IFSA} \\ \alpha T_0 = \frac{T_\alpha}{\ln(2)} &\Rightarrow T_\alpha = \alpha \ln(2) \end{aligned} \quad (108)$$

with $T_0 = 1$ and $\alpha = 0.995$ which corresponds to the temperature reduction factor of the adjusted SA search the initial temperature level of IFSA method is set to $T_\alpha = 0.69$.

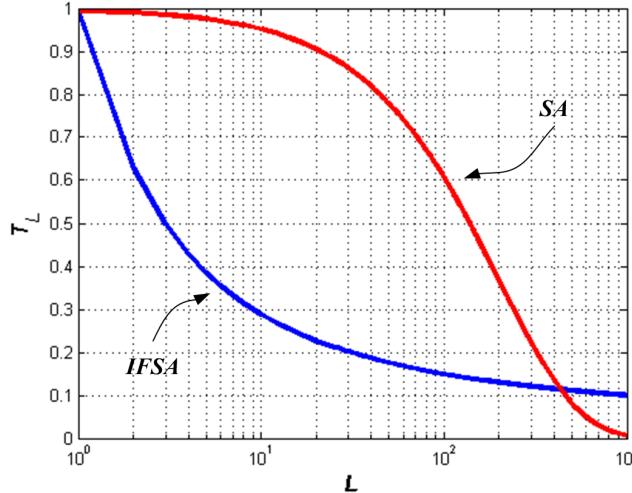


Figure 70: Comparison of original SA and IFSA annealing schedule functions

The difference between SA and IFSA annealing schedules based on adjusted parameters is seen in Figure 70. As can be seen in the figure, the temperature level of the original SA algorithm is kept high during the first ten levels. At this time, IFSA temperature is decreased to less than 30% of the initial temperature and shortly afterwards starts to decrease slowly, while SA temperature decrease now gets steeper. At the end, IFSA keeps temperature almost constant allowing the thorough search of actual neighbour space. On the other hand, the original SA search turns into a greedy algorithm without any uphill excursions allowing only perturbations

which lead to better solution. In short, when compared to original SA method IFSA annealing schedule is steeper at the beginning and kept at higher level towards the end of the search. This allows higher convergence speed at start up, and better analysis of close neighborhood area at lower temperature levels.

The second important difference regarding the SA and IFSA approach is based on the probability of uphill movement. According to eq. (95) the original SA method is based on Boltzmann probability density function:

$$p^{SA}(\Delta Q) = e^{-\frac{\Delta Q}{kT_L}}$$

where T_L is the actual temperature level, ΔQ is cost function difference between the actual and new candidate configuration, and k constant which directly relates these two variables. Numerically, k is set to value which, at first temperature level, guarantees the probability of transition defined as p_{max} to a 99% of new candidates:

$$k = -\frac{3\sigma_{\Delta Q}}{\ln(p_{max})}$$

The drawback of this distribution is its thin tail, which makes the configurations with substantial cost function difference rather unreachable under small number of iterations. However, it sounds like a good idea to permit this far excursions more often since wider search space is covered in smaller number of iterations and deeper local minimums are successfully avoided. Although the majority of SA enhancements is based on the typical Boltzmann probability density function, it was noted in [Szu87] that Cauchy distribution, which has a fatter body than Boltzmann distribution, permits easier access to larger set of candidates without any significant algorithm convergence loss. The modified Cauchy distribution they define and which is used in IFSA is expressed as:

$$p(\Delta Q) = \frac{T_L}{(\Delta Q^2 + k_c T_L^2)^{(D+1)/2}} \quad (109)$$

where k_c relates T_L to ΔQ and D defines the degrees of freedom, that is, the number of unknown parameters. Adapting D parameter to nonlinear decoupling problem is pretty straightforward. Decoupling approximation surface is constructed according to eq. (81):

$$\hat{x}_1 = \sum_{i=1}^{i \leq N_1} a_i y_1^i + \sum_{j=1}^{j \leq N_2} b_j y_2^j + k_1 y_1 y_2 + k_2$$

where a_i, b_j, k_1, k_2 are the decoupling coefficients, and N_1 and N_2 represent the degrees of freedom of variables y_1 and y_2 respectively. Then, by setting the approximation precision to $(N_1, N_2) = (9, 9)$, parameter D is equal to 20.

Modified Cauchy distribution constant k_c is chosen to satisfy the same requirement as in the Boltzmann function. In other words, maximum probability of transition p_{max} , which occurs at the first temperature level T_1 , should be satisfied for 99% of all possible cost function differences. Since ΔQ follows the Gaussian distribution 99% of all possible numerical values of ΔQ matches three standard deviations σ of the corresponding distribution. Hence, p_{max} can be expressed as:

$$p_{max} = \frac{T_1}{\left((3\sigma_{\Delta Q})^2 + k_c T_1^2 \right)^{(D+1)/2}} \quad (110)$$

Then, by extracting k_c from eq.(110) the following statement is obtained:

$$k_c = \sqrt{\frac{1}{T_1^2} \left(\sqrt[(D+1)/2]{\frac{T_1}{p_{max}}} - (3\sigma_{\Delta Q})^2 \right)} \quad (111)$$

The differences between these two probability transition approaches are seen easily in Figure 71 where transition probability functions of SA and IFSA approach are presented.

As expected the probability function based on modified Cauchy distribution is fatter than the one based on Boltzmann function allowing uphill movements with more ease. With Cauchy approach the transition probability decreases very slowly at low cost function differences. The Boltzmann approach is based on the opposite logic since at low cost function differences its probability transition experiences steepest fall. When ΔQ reaches 0.1, the transition probability of SA method falls to 55% while IFSA keeps the transition probability as high as 92%. Actually, the highest probability difference between the two approaches is noticed for slightly larger ΔQ values whose perturbations are accepted with almost twice as much probability in IFSA than in SA approach. The advantage of IFSA probability transition function lies inside this middle area which covers ΔQ from 0.1 to 0.6. With $\Delta Q = 0.2$, for example, transition

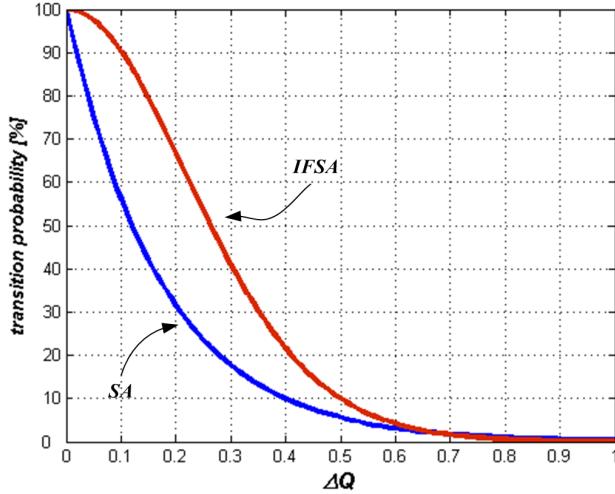


Figure 71: Transition probability functions of SA and IFSA approach

probability of IFSA marks approximately 65%, while SA is limited to approximately 32%. Apparently by including this middle class of ΔQ values more often into the search path, search robustness is increased but convergence speed doesn't suffer substantial loss. Once cost function difference surpasses the 0.6 border, both approaches present almost identical transition probability of less than a few percent as this kind of movement is rarely accepted.

Since probability transition function in IFSA presents high values not only for very small energy differences, it can be stated that IFSA favours the wide search space exploration during the entire search process. However, the depth of exploration depends not only on the transition probability function but also on the annealing schedule and maximum displacement factor R_0 . The annealing schedule function of IFSA is very steep at the beginning, suppressing the initial deep exploration rather fast and keeping the rest of the search concentrated on small energy differences around the actual point. Thus, so called *deep* exploration which easily permits uphill movements is available only at the initial time. Still, these initial search opportunities can not be exploited if R_0 factor is small as only the small part of the potential search space can be reached for the analysis. This small maximum displacement factor is adequate choice for the advanced search status when global minimum area is well located, and thorough search is required in order to find the best solution. By keeping R_0 constant during the entire search process, like in original SA approach, two opposite search requirement can no be met and consequently R_0 is chosen as their compromise. This is changed in IFSA approach and this change presents the third major difference between these two methods. In IFSA, maximum

<i>Inherited SA Values</i>	<i>New IFSA Functions</i>
$Res_{symb} = 0.00001$	$T_L = \frac{T_\alpha}{\ln(L+1)}$ $T_\alpha = 0.69, L = 1, 2, 3\dots$
$Res_{coeff} = 0.00001$	$p(\Delta Q) = \frac{T_L}{(\Delta Q^2 + k_c T_L^2)^{(D+1)/2}}$
$Coeff_0(i) = \begin{cases} 1, & i = 1 \\ 0, & i = 2, \dots, 20 \end{cases}$	$k_c = \sqrt{\frac{1}{T_1^2} \left(\frac{(D+1)/2}{\sqrt{\frac{T_1}{p_{max}}}} - (3\sigma_{\Delta Q})^2 \right)}$
$Coeff_{new}(i) = \xi_i R_0 + Coef_{act}(i)$ $i = 1..20; \xi_i \in \langle -1, 1 \rangle$	$R_0(L) = \delta R_0(L-1)$ $R_0(0) = r, L = 1, 2, \dots$
$Q = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_{1i} - \hat{x}_{1i})^2}$	$r = 0.1$ $\delta = 0.99$
$p_{max} = 50\%$	$N_{VCPL} = 50$

Table 6: Theoretic set of IFSA search parameters before empirical analysis

displacement factor is set to high value covering whole search space at the beginning, and it is gradually decreased as the search advances. The dependence is made according to the exponential function:

$$R_0(L) = \delta R_0(L-1), R_0(0) = r, L = 1, 2, \dots \quad (112)$$

where L stands for the corresponding temperature level, δ is the search space reduction factor, and r the initial maximum allowed displacement. The numerical values of δ and r are determined through empirical analysis.

Based on these three radical changes, the rest of SA search parameters also experienced some modifications as seen in Table 6. As IFSA approach is not oriented only on precision the maximum number of iterations is not limited to the probability, but the search is stopped when in five consecutive temperature levels no cost function improvement is obtained or cost function reaches the sufficient precision.

The parameters describing IFSA approach can be divided into the ones inherited from the SA method and the ones developed for the corresponding IFSA search. The inherited SA factors are from the beginning set to the adequate numerical values obtained in previous SA analysis while four new IFSA factors that have to be set empirically (r, δ, N_{VCPL}) are over-dimensioned in this set of parameters favouring precision over speed. However the final set of parameters, which is proposed after the simulation analysis, is based on precision-speed compromise. A set of these new parameters adopting IFSA to nonlinear decoupling problem will be defined in the following section.

6.2.1 Fine tuning of IFSA parameters

The fine tuning of IFSA parameters is done for 64QAM signals as the most demanding case between 16QAM, 64QAM and QPSK constellations. The results are again based on 2000 calibration symbols transmitted under SNR of 100dB and every presented result is the average of ten simulations under same operating conditions.

First set of simulations focuses on adequate pair of parameters for maximum allowed displacement definition (r, δ) . Afterwards, the obtained pair is used for analyzing the appropriate number of visited candidates per temperature level (N_{VCPL}). IFSA search is adopted to the nonlinear decoupling problem once these three factors are encountered as all other parameters are already well defined inside the SA analysis and the same conclusions are valid for IFSA. Finally, the search based on adjusted IFSA search together with numerical solution to one inverse coupling approximation is presented.

$$\Rightarrow(r, \delta)$$

Since probability transition function in IFSA presents high values not only for very small energy differences, it can be stated that IFSA favours the wide search space exploration during the

entire search process. However, the depth of exploration depends not only on the transition probability function but also on the annealing schedule and maximum displacement factor R_0 . The annealing schedule function of IFSA is very steep at the beginning, suppressing the deep exploration rather fast and keeping the rest of the search concentrated on small energy differences around the actual point. Thus, so called *deep* exploration, which easily permits uphill movements, is only available at the initial time. Still, these initial search opportunities can not be exploited if R_0 factor is small as only the small part of the potential search space can be reached for the analysis. However, small search space is adequate choice when the search is located close to the global minimum area as the ideal solution is located faster within smaller space. By keeping R_0 constant during the entire search process, like in original SA approach, the two opposite search requirement can no be met and consequently R_0 is chosen as their compromise. In IFSA approach maximum displacement factor (R_0) is at the beginning set to high value which should cover the whole search space ($R_0(0) = r$), and it is gradually decreased as the search advances according to eq.(112):

$$R_0(L) = \delta R_0(L - 1), R_0(0) = r, L = 1, 2, \dots$$

with L being the corresponding temperature level, δ is the search space reduction factor, and r is the initial maximum allowed displacement.

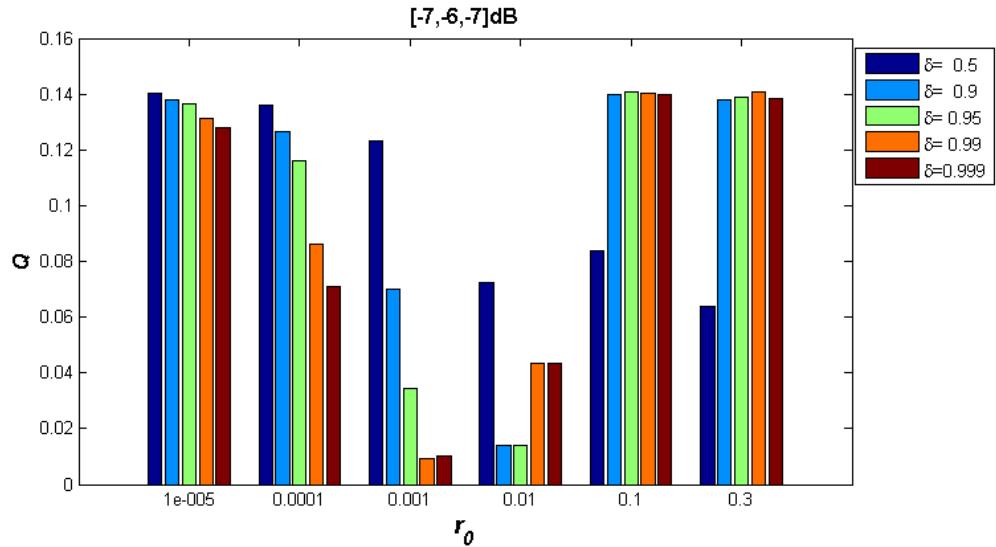


Figure 72: IFSA search based on different sets of (r, δ) factors under strong coupling

If the starting search space is large, and the reduction factor very close to one, the search will progress rather slowly. In this case the probability of worse movement decreases and even enters the downhill search but the search space is kept large all the time. As a consequence five consecutive temperature levels without the reduction of cost function occur far before the global minimum area is located. This is confirmed in Figure 72, where the precision of IFSA search based on different sets of (r, δ) factors under strong coupling is presented. As is seen in the right image side, when the starting search space is set to $r_0 = 0.1$ or $r_0 = 0.3$ the algorithm presents poor precision for all space reduction factors except for $\delta = 0.5$. Still, with reduction factor of 50% the search is easily drown into saturation as clear tendency path can hardly be established at such fast pace.

If the starting search space is too small, the search progress cannot move far from the starting point and the obtained solution results depend largely on the coupling strength. Under strong coupling the starting search space corresponding to $r_0 = 10^{-5}$ and $r_0 = 10^{-4}$ represents the selection of narrow starting search space. The obtained cost function precision is generally low, but as the reduction factor gets closer to one, it is increased in both cases. This is logical as if the small search space is additionally reduced fast the algorithm is easily left without any new candidates and no progress can be made.

Consequently, two starting search spaces, in the middle of the image, built around $r_0 = 0.001$ and $r_0 = 0.01$ offer the adequate selection. In fact the best precision under strong coupling is reached with the following four pair of factors: $(0.001, 0.99)$, $(0.001, 0.999)$, $(0.01, 0.90)$ and $(0.01, 0.95)$. The pairs are in accordance with the theoretical expectations as they ratify that in vicinity of adequate set of parameters smaller search space requires lower reduction factor to obtain the same precision level.

The same analysis is repeated for system under moderate coupling of $[13, -12, -10]$ dB in Figure 73. The behaviour of algorithm is similar to the one presented for strong coupling as the search corresponding to the large search space is located on the left image side, and the one corresponding to the narrow search space on the right image side. Naturally, the potentially adequate set of parameters is found in the center of the image. However, in this case the center is moved slightly to the right since the coupling strength is lower and the ideal solution is located closer to the initial search point. Thus, in case of moderate coupling smaller search space is favoured and the best precision is reached for the following pair of factors: $(0.0001, 0.99)$, $(0.0001, 0.999)$, $(0.001, 0.90)$, $(0.001, 0.95)$ and $(0.001, 0.99)$.

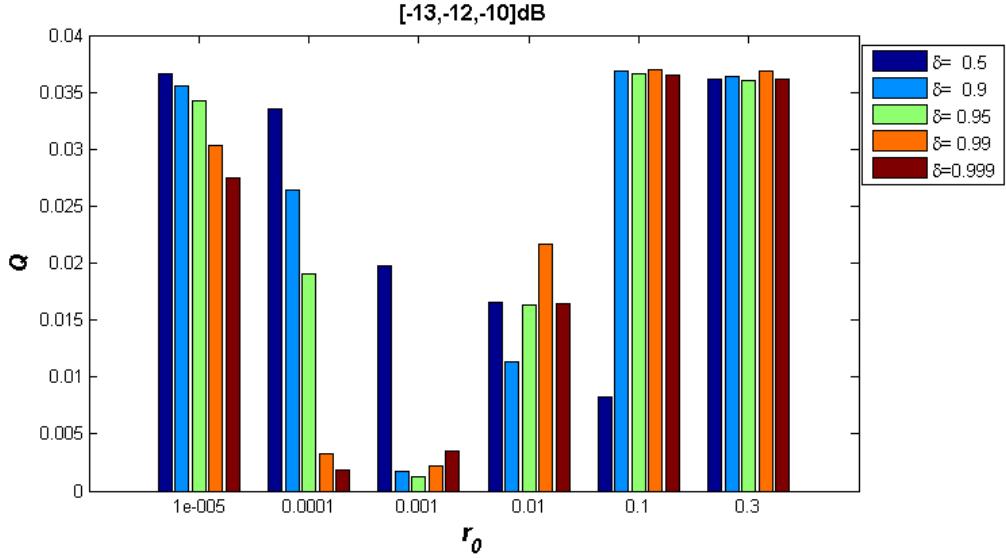


Figure 73: IFSA search based on different sets of (r, δ) factors under moderate coupling

Comparing the two sets of factors, the only pair emerging in both coupling types is $r_0 = 0.001$ and $\delta = 0.99$. Hence, these values are used as adequate parameters for IFSA search adopted to nonlinear decoupling problem.

$$\Rightarrow N_{VCPL}$$

The number of visited candidates per temperature level defines the number of energy configurations analyzed between two temperature changes. If this number is large, the search space is analyzed in detail which produces stable results but consumes a lot of the processing time. Nevertheless, if the number of visited configurations is small, the search progress is made based on unreliable data which leads to poor cost function results.

The performance of IFSA search method based on different N_{VCPL} number under strong coupling conditions of $[-8, -6, -6]$ dB is depicted in Figure 74. As expected, when N_{VCPL} number is too small, like for $N_{VCPL} = 5, N_{VCPL} = 3$ and especially for $N_{VCPL} = 1$, the search is easily drowned into a *blind ally*. Local minimum area is located well, but as all temperature levels are spent fast, no uphill movements are available, and no additional progress can be made despite the additional number of iterations.

On the other hand, if number of visited candidates is large and the temperature reduction function is not smooth enough the search can get stacked while in wandering phase and can

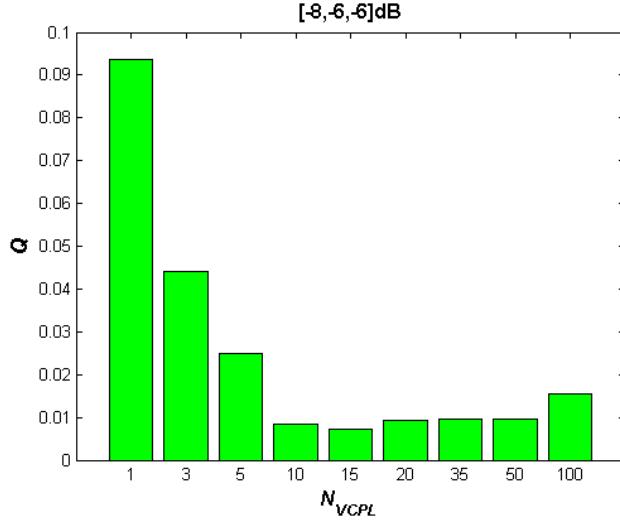


Figure 74: IFSA search with different N_{VCPL} factor under $[-8, -6, -6]$ dB coupling

only offer local minimum as final solution. The probability of producing local minimum as final solution increases as the coupling gets stronger. This occurs under strong coupling for $N_{VCPL} = 100$, where the obtained precision is visibly lower than, for example, the one obtained with $N_{VCPL} = 50$ or $N_{VCPL} = 20$. Eventually the best precision under strong coupling is reached with $N_{VCPL} = 10$ and $N_{VCPL} = 15$.

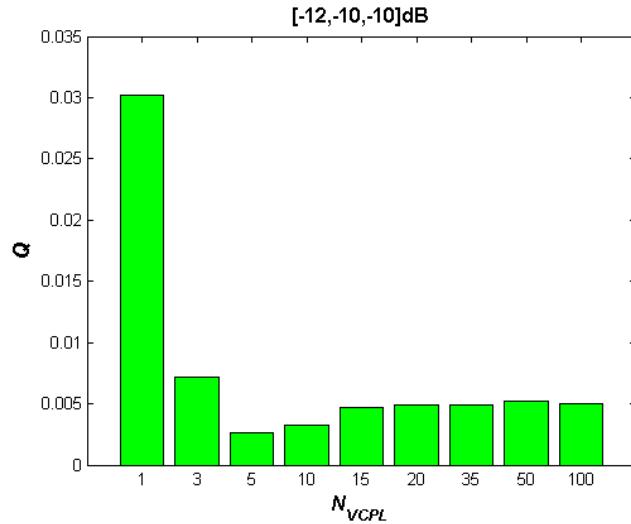


Figure 75: IFSA search with different N_{VCPL} factor under $[-12, -10, -10]$ dB coupling

When the analysis is centered on moderate coupling the behaviour of IFSA method changes

very little. This is seen in Figure 75, where cost function in dependence of N_{VCPL} under moderate coupling of $[-12, -10, -10]$ dB is presented. Like under strong coupling, the N_{VCPL} number is too small for $N_{VCPL} = 3$ and $N_{VCPL} = 1$, as the cost function indicates lower solution precision than obtained with larger number of visited points per level. On the other hand, for N_{VCPL} higher than 10 the obtained precision is kept constant, but the best cost function is obtained with $N_{VCPL} = 5$ and $N_{VCPL} = 10$.

According to the presented simulation results, the adequate selection of visited number of candidates per level lies between 5 and 15. Since strong coupling conditions discard $N_{VCPL} = 5$ as too small, and moderate coupling exhibits slightly better results with $N_{VCPL} = 10$ than with $N_{VCPL} = 15$, the adequate parameter adjusted to the nonlinear coupling problem is set to $N_{VCPL} = 10$.

⇒IFSA Search implementation example

In this section, the IFSA search with the adjusted set of IFSA search parameters is implemented on moderate coupling of $[-10, -12, -10]$ dB. The search is based on surface constructed with 2000 calibration symbols. The behaviour of IFSA algorithm is shown in Figure 76, where cost function is presented in dependence of number of iterations for 10 instances of IFSA search under same operating conditions.

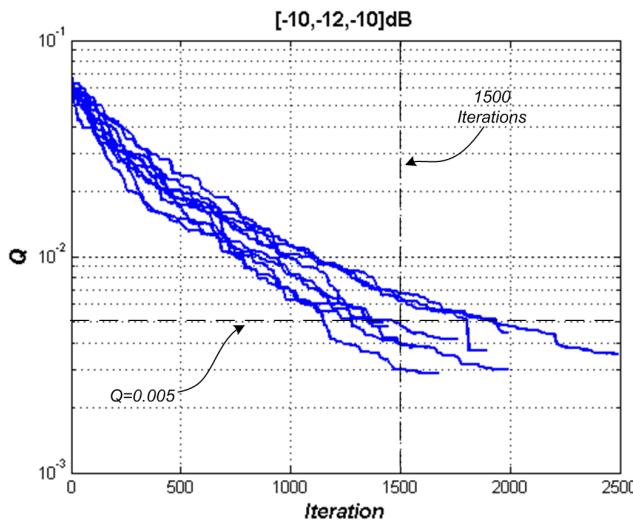


Figure 76: Ten instances of adjusted IFSA search under moderate coupling

According to IFSA approach the search is stopped if five consecutive temperature levels do not produce any improvement of cost function. As the image confirms this occurred in 9 instances before reaching 2500 iterations which, when compared with 17000 iterations of the original SA approach, presents the reduction of processing load of approximately 85%. Unlike SA search which produced 10 closely related curves, the IFSA curves are spread over wider area as the search progresses due to the higher influence of probability since precision is sacrificed in favour of smaller number of iterations. Furthermore, the IFSA cost function is also based on quadratic mean square function for which it has been demonstrated that smaller Q is not transformed into any significant BER enhancement for $Q < 0.005$. Thus, since the search is not oriented on the precision of cost function it can be stopped when cost function reaches the limit value of 0.005. Eventually, this occurs on average for 1500 number of iterations as shown in Figure 77 where the corresponding cost function value after 1500 iteration is presented for all ten search instances.

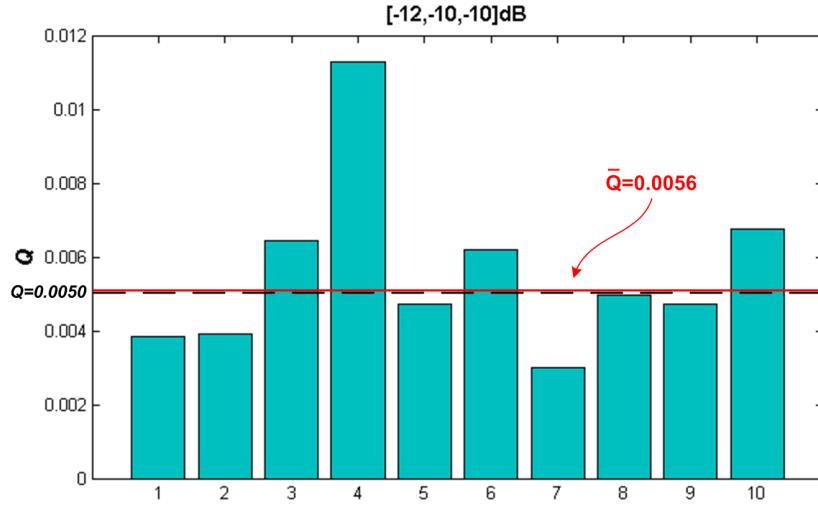


Figure 77: The precision of IFSA search after 1500 iterations

The image shows that all cost function solutions are kept in the same orders of magnitude. Six instances obtained result under $Q = 0.005$ while three search executions ended in area of $Q = 0.006$ and only one is discarded as intolerable since it reached $Q = 0.011$. The average of all ten executions is located at $\bar{Q} = 0.0056$ which indicates that the search under moderate coupling can be suspended when 1500 iterations are surpassed. If at this time, the presented cost function exhibits high value like 0.01 the search can be executed again. If only one out of ten searches is repeated, according to this approach, the average number of iteration

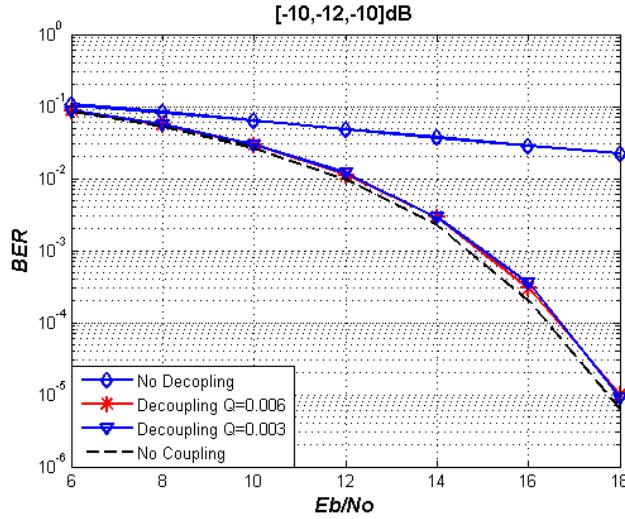


Figure 78: Performance of IFSA based decoupling module under moderate coupling

required for one correct IFSA execution is reduced to 1650 ($11 \times 1500/10$) which corresponds to reduction of processing load of more than 90% when compared to the original SA approach oriented on precision.

The reduction of permitted number of iterations to 1500 is verified in Figure 78 where the performance of IFSA decoupling module under moderate coupling conditions is presented. The first curve on the image presents the behaviour of system without the decoupling module under coupling of $[-10, -12, -10]$ dB. The lowest curve shows the theoretical system performance with only AWGN. Two curves in the middle present two decoupling module instances based on IFSA approach stopped after 1500 iterations. Actually, the coefficients used for generating this two curves correspond to the seventh and tenth IFSA execution depicted in Figure 77. Regardless of the numerical difference of cost functions, the performance of both decoupling modules is very similar. Since the search is abandoned when cost function reaches sufficient Q value defined with $Q_{suff} = 0.005$ the performance of decoupling module does not follow the theoretical transmission curve as close as is the SA approach.

The presented IFSA decoupling solution presents robust performance even though it is based on only 1500 iterations. When compared to SA method, the number of required iterations is considerably reduced while the loss of solution precision is generally small which confirms the adequate selection of adjusted IFSA search parameters. This backs the use of IFSA method in the decoupling module as an alternative to SA approach.

⇒Conclusion

The proposed Improved Fast SA search method is included in decoupling module as the fast version of the original SA search. It follows the same search logic as the original SA algorithm with the difference in the annealing schedule which is much steeper, the initial search space is wider and the probability of uphill movement is higher at the beginning. This way the search is carried out more thoroughly in the initial stage, allowing stepper search progress sooner than in the original SA approach. Furthermore, since the search space is reduced according

<i>Inherited SA Values</i>	<i>New IFS A Functions</i>
$Res_{symb} = 0.00001$	$T_L = \frac{T_\alpha}{\ln(L+1)}$ $T_\alpha = 0.69, L = 1, 2, 3\dots$
$Res_{coeff} = 0.00001$	$p(\Delta Q) = \frac{T_L}{(\Delta Q^2 + k_c T_L^2)^{(D+1)/2}}$
$Coeff_0(i) = \begin{cases} 1, & i = 1 \\ 0, & i = 2, \dots, 20 \end{cases}$	$k_c = \sqrt{\frac{1}{T_1^2} \left(\frac{(D+1)/2}{\sqrt{p_{max}}} \sqrt{\frac{T_1}{p_{max}}} - (3\sigma_{\Delta Q})^2 \right)}$
$Coeff_{new}(i) = \xi_i R_0 + Coef_{act}(i)$ $i = 1..20; \xi_i \in \langle -1, 1 \rangle$	$R_0(L) = \delta R_0(L - 1)$ $R_0(0) = r, L = 1, 2, \dots$
$Q = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_{1i} - \hat{x}_{1i})^2}$	$r = 0.001$ $\delta = 0.99$
$p_{max} = 50\%$	$N_{VCPL} = 10$
	$Q_{suff} = 0.005$

Table 7: Theoretic set of IFS A search parameters before empirical analysis

to the exponential function, the maximum allowed displacement is reduced with temperature, and as the consequence the search concentrates more on the small space around the global minimum area.

The parameters of IFSAs algorithm are adjusted to nonlinear coupling problem through empirical analysis based on different set of simulations. The adopted set of IFSAs search parameters is presented in Table 7. The table consists of two columns with inherited SA values in the left column, and new IFSAs functions and parameters values based on empirical analysis in the right one.

The performance of IFSAs decoupling module shows good decoupling results, with slightly lower BER level when compared to the original SA search. Nevertheless, at the same time the required number of iterations is reduced by more than 85% which justifies the precision loss. The IFSAs presents stable solution behaviour under all coupling conditions. The approach requires less processing power than SA, and as such might be used in mobile phones inside the decoupling module.

6.3 Guided Multi-level (GML) search

In previous chapters the problem of nonlinear coupling is resolved with decoupling module based on two different Simulated Annealing approaches that are founded on heuristic methods. In order to avoid the probability factor and, as a consequence search indecision, a deterministic solution in the form of a simple iterative algorithm denominated Guided Multi-level (GML) search is elaborated. The idea and the development of this, third decoupling approach, present the author major contribution to the surface approximation methods. The search name includes “multi-level” since it is divided in levels (similar to temperature levels of SA approach), while “guided” was inspired with military guided missile and in this case describes the speed and precision at the same time. The advantages of the proposed GML search include easy control of speed and precision, reduced processing requirements, simple implementation and avoidance of statistical dependence of the solution. Eventually, the approach ensures deterministic solution to the decoupling problem with excellent surface precision and low processing demands.

6.3.1 Describing GML search approach

Much like the first two decoupling module algorithms, the GML uses the same quadratic mean square function defined in eq.(85) as:

$$Q = \sqrt{\frac{1}{N} \sum_{k=1}^N (x_{1k} - \hat{x}_{1k})^2}$$

where N is the number of symbols used for calibration, x_{1k} is the sent k^{th} calibration symbol on first antenna, and \hat{x}_{1k} its corresponding value calculated with obtained decoupling coefficients. The number of coefficients and their model are chosen according to the inverse approximate surface defined in eq(81):

$$\hat{x}_{1k} = \text{Max}(|y_{1k}|, |y_{2k}|) \left\{ \sum_{i=1}^{i \leq N_1} a_i (y_{1k}^N)^i + \sum_{j=1}^{j \leq N_2} b_j (y_{2k}^N)^j + k_1 y_{1k}^N y_{2k}^N + k_2 \right\}$$

where y_{1k}^N and y_{2k}^N are the k^{th} normalized received symbols on the corresponding antennas, N is the number of calibration signals, $\text{Max}(|y_{1k}|, |y_{2k}|)$ is the maximum received absolute value level and N_1 and N_2 represent the degrees of freedom of variables y_1 and y_2 respectively. The surface model based on 20 degrees of freedom ($N_1 = 9, N_2 = 9, k = 2$) has proven as good surface approximation model and will be used here as well during the selection of all GML search parameters. The initial surface approximation is based on the set of coefficients $Coeff_0$ defined in eq.(90) and eq.(91) as:

$$Coeff_0 : \begin{cases} a = [1, 0, 0, 0, 0, 0, 0, 0, 0] \\ b = [0, 0, 0, 0, 0, 0, 0, 0, 0] \\ k = [0, 0] \end{cases} \Rightarrow Coeff_0 = \underbrace{[1, 0, 0, 0, \dots, 0, 0, 0]}_{19 \text{ zeros}}$$

where first nine elements correspond to coefficients $a_i, i = 1..9$, second nine elements correspond to $b_i, i = 1..9$, and the last two correspond to k_1 and k_2 . As already mentioned this configuration corresponds to ideal transmission without any distortions. Other similarities with the first two decoupling approaches include symbol resolution which are set to $Res_{symb} = 0.00001$ and maximum resolution of decoupling coefficients which is also set to $Res_{coef} = 0.00001$.

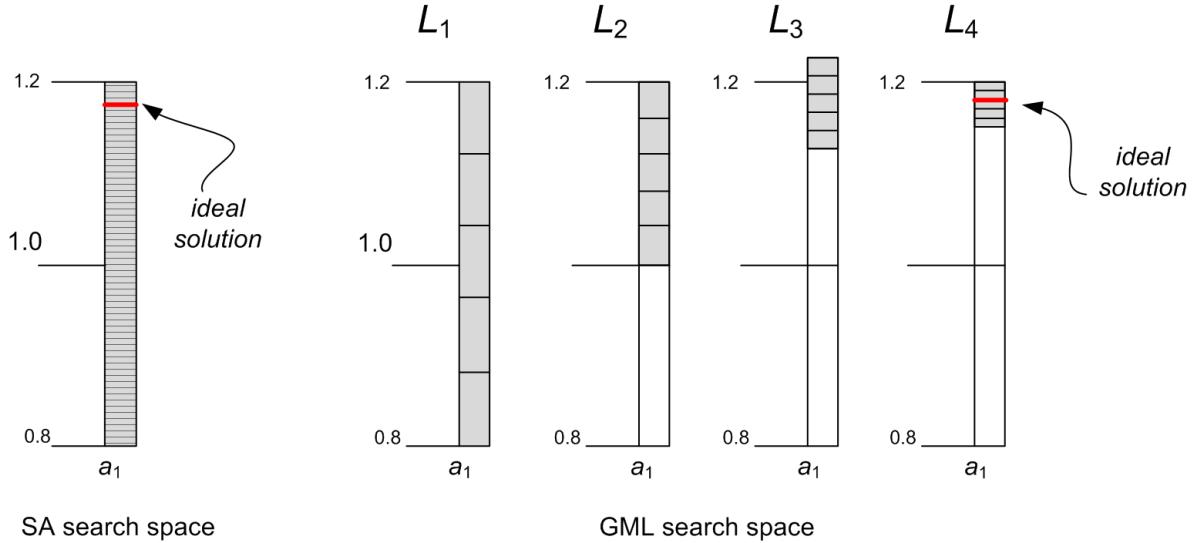


Figure 79: Construction of search space for SA and GML approach

Apart from these identical parameters characteristic for nonlinear decoupling problem some similarities between GML and SA based methods do exist but eventually the differences are greater. Initialization of both algorithms consists of the construction of the space where the search is carried out. As already pointed out, if M samples are chosen for each of N design variables, the number of possible configurations to be explored is M^N . In our case, with $N = 20$ degrees of freedom, using only 10 samples per variable builds the search space of 10^{20} possible configurations. The exhaustive exploration of this domain is practically impossible yet the SA method is based on this hostile domain. The GML search presents completely different approach to the problem. Unlike SA where the space resolution is static, in GML it changes with every search level as the solution domain is reduced. However, since the number of samples per variable is kept constant, by reducing the search domain the resolution of the corresponding space is increased. This is seen more clearly in Figure 79, where the search space of one coefficient is constructed for SA approach and in 4 consecutive search levels for GML approach. Since the coefficient corresponds to a_1 the space is built around numerical value of 1. The SA search space from the beginning exhibits dense solution space and since all solutions carry the same probability the ideal one can not be located easily as not all configurations can be tested. On the other hand, each search space of GML approach has exactly six samples and eventually all six of them can be visited. This way, the whole search space is covered, the probability factor is cast aside, and the search itself is accelerated.

The coefficient search area is constructed at the beginning of each level around the numerical value of the corresponding coefficient taken from the configuration showing the lowest cost function at the moment. Thus the search space of each coefficient is built independently and consequently each space can cover different numerical values. The space width is described through specific GML parameter entitled as W , which defines the space inside the $\langle Coef_{min}(i) - W, Coef_{min}(i) + W \rangle$ with i corresponding to coefficient index. The space is than evenly divided by parameter $NVCPL$ which, like in SA methods, defines the number of visited candidates per level. Thus, $NVCPL$ corresponds to the number of times the calculation of cost function is carried out at one temperature level for one coefficient. During the analysis of search space, all sampling points are checked one by one and at each iteration, the obtained actual cost function Q_{act} is compared to the temporary cost function minimum Q_{min} . If the new state possesses less energy ($Q_{act} < Q_{min}$), the solution is accepted as new Q_{min} value, and the corresponding configuration is stored in $Coef_{min}$. After all sampling points are visited, the search attention is transferred to next decoupling coefficient or, in case there are no coefficients available at the same level, new precision level is initiated.

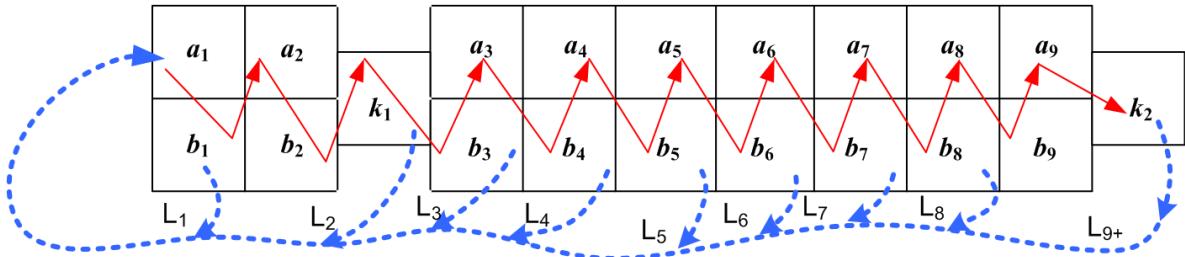


Figure 80: Zig-zag method used in GML search for obtaining the decoupling coefficients

In GML approach the search focuses on one coefficient at the time keeping the rest of them constant and assigning them the value corresponding to configuration stored in $Coef_{min}$. This way the correct search path is located faster and the progress is controlled easier. The order of coefficients is arranged in a zig-zag method depicted in Figure 80, starting with a_1 and finishing with k_2 . The most significant coefficient is assigned to a_1 since it undoubtly carries the highest energy, which means information. By putting k_2 at the tail of all other coefficients it becomes the least significant one. Eventually, all coefficients from left to right are lined up according to the power factor they correspond to. For this reason k_1 is located after a_2 and b_2 since all three of them are found alongside to square exponent. Moreover, the zig-zag structure is constructed from top to bottom in a way that it assigns higher importance to data obtained at the corresponding antenna. In case of reasonable coupling this makes sense

since this data carries more information about itself than it can be found in the coupled signal at other antenna. Apart from classifying the coefficients according to their importance into categories, the GML method proposal includes gradual introduction of coefficients categories into the surface model. According to this proposal the first level search denoted as L_1 in Figure 80, includes only decoupling coefficients a_1 and b_1 . The second one (L_2) introduces the square exponent and hence adjust the approximation surface using decoupling coefficients in the following order a_1, b_1, a_2, b_2, k_2 . This logic is continued up to the ninth level after which all 20 coefficients are introduced and all are analyzed one by one in subsequent search precision levels. The precision level is regarded as finished when all coefficients composing it are analyzed once.

The last parameter required for correct description of GML algorithm is the search space reduction factor denoted as β . This parameter, expressed numerically between $\langle 0, 1 \rangle$, defines the amount of preserved search space between the two consecutive precision levels. Hence, the width of the consecutive search spaces follows the exponential function according to:

$$W_{L+1} = \beta W_L \quad (113)$$

where L stand for the precision level. The width of search domain is constant for all coefficients at the same precision level. Thus, since decoupling coefficients are introduced gradually, their corresponding initial search space is narrower as their importance is lower. In fact when the search reaches tenth level the search spaces of two most important coefficients have been constructed and reduced nine times while the search spaces of a_9, b_9 and k_2 have only been constructed, and analyzed, ones. The GML algorithm is stopped when the surface approximation enhancement between the two consecutive temperature levels drops bellow 1%, or when user defined number of precision levels are analyzed.

The simplified functioning of GML search is shown in Figure 81 where the precision level is limited to only three level searches. At first level (L_1) the initial search space is set to $SD = 0.2$ and since $Coeff_{min}$ is equal to $Coeff_0$, the search space is built around $a_1 = 1$ and $b_1 = 0$. The search space reduction factor is set to $\beta = 0.5$ and the number of visited points per level is set to $NVCPL = 50$, meaning that initial search space exhibits the precision equal to 0.008 (twice the search space divided by the number of visited points). When the search is focused on one coefficient the rest of them is set to their corresponding numerical value stored in $Coeff_{min}$. Hence the in ital search configuration denoted as $Coeff_{min}(L_{11})$ is based on the

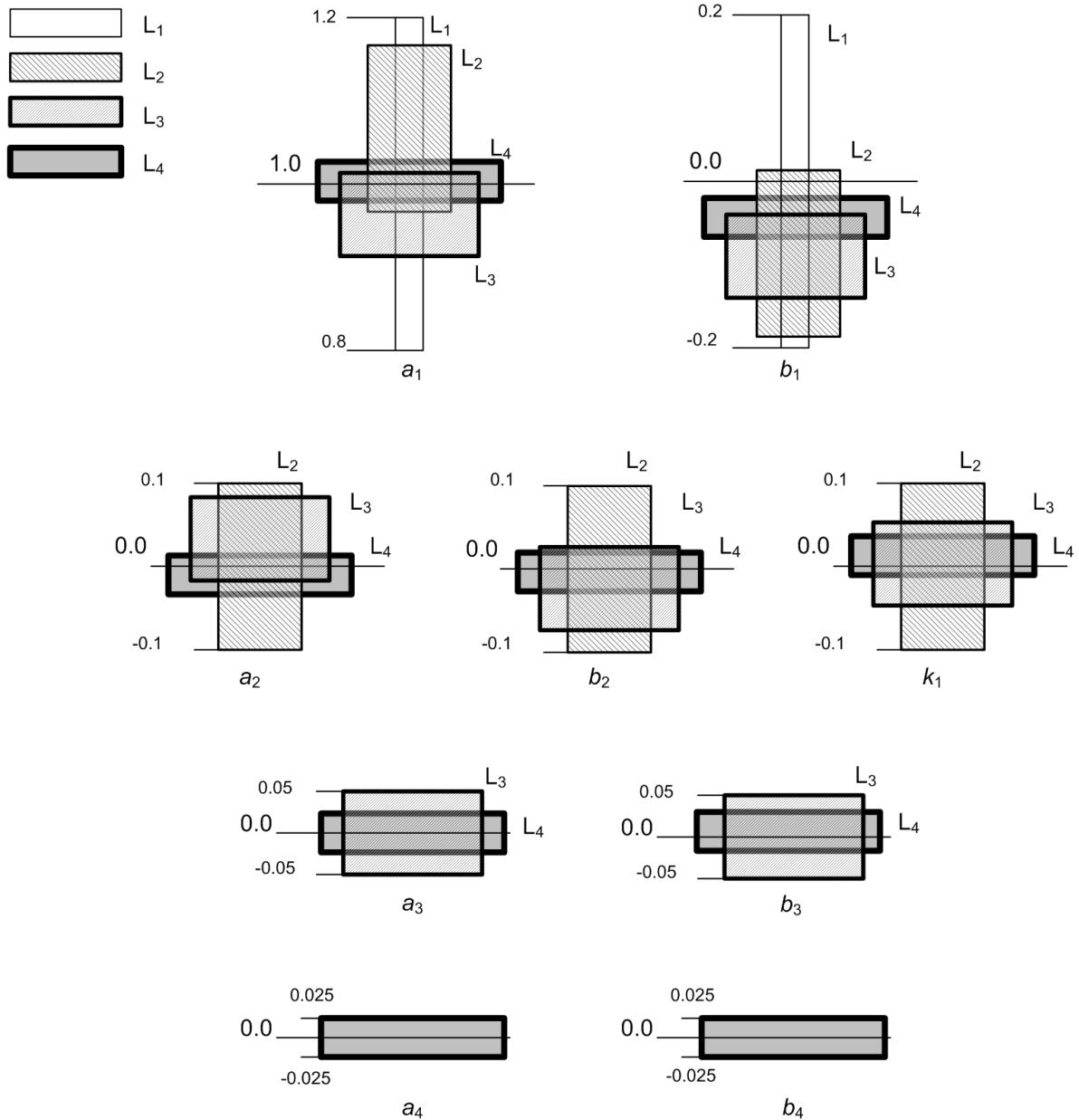


Figure 81: GML search example implemented on first four precision levels

following set of coefficients:

$$Coef_{min}(L_{11}) = \left[X, \underbrace{0, 0, 0, \dots, 0, 0}_{19 \text{ zeros}} \right]$$

where the indexing of L_{ij} corresponds to the i -th precision level, and j -th decoupling coefficient inside the $Coef$ structure. Thus, the coefficient marked with X in $Coef_{min}(L_{11})$ corresponds to the coefficient of interest which is changed for each cost function calculation and the rest of the coefficients is kept constant until N_{VCPL} number of candidates is visited. Hence, after 50 iterations the coefficient obtaining the minimum cost function denoted as $a_1^{L_1}$ is updated in $Coef_{min}$ configuration, and the search for the most adequate b_1 factor is carried out according to the updated configuration:

$$Coef_{min}(L_{12}) = \begin{bmatrix} a_1^{L_1}, X, \underbrace{0, 0, 0, 0, 0, \dots, 0, 0}_{18 \text{ zeros}} \end{bmatrix}$$

After the first level search is finished, the search space is reduced according to factor $\beta = 0.5$ to $SD = 0.1$. At this point three new decoupling coefficients a_2 , b_2 and k_1 are introduced and all together five search spaces with precision 0.004 are constructed. The search is than carried on the following configuration used for obtaining the $a_1^{L_2}$:

$$Coef_{min}(L_{21}) = \begin{bmatrix} X, b_1^{L_1}, \underbrace{0, 0, 0, 0, 0, \dots, 0, 0}_{18 \text{ zeros}} \end{bmatrix}$$

and the same pattern is repeated four more times. In general, the configuration corresponding the i -th precision level of eg. 6-th decoupling coefficient is described as:

$$Coef_{min}(L_{i6}) = \begin{bmatrix} a_1^{L_{i-1}}, b_1^{L_{i-1}}, a_2^{L_{i-1}}, b_2^{L_{i-1}}, k_1^{L_{i-1}}, X, b_3^{L_{i-1}}, \dots, b_9^{L_{i-1}}, k_2^{L_{i-1}} \end{bmatrix} \quad (114)$$

After the second precision level is finished, the search space is again reduced. Consequently, coefficient precision is increased up to 0.002 and the surface is approximated with seven coefficients at the end of this level. In the presented example the number of search levels is limited to four, which sets the final precision to 0.001 and number of coefficients to nine.

The low processing demands of GML approach are not evident at first sight as neither system precision nor the required number of iterations show significant alterations. However, by focusing the search on one coefficient at the time, the rest of them is kept constant. This fact can be exploited with simple mathematical optimization that substantially reduces the processing

load. The optimization consists in partial calculation of the surface model at the beginning of each search, immediately after the construction of search space. The corresponding part of surface is denoted as \hat{C}_{1k} since it is based on constant set of coefficients. It is calculated according to the same inverse approximate surface model defined in eq(81), with the corresponding coefficient of interest set to zero. Hence, if the search is carried out on, for example, decoupling coefficient a_3 , the constant part matches:

$$\hat{C}_{1k} = \text{Max}(|y_{1k}|, |y_{2k}|) \left\{ \begin{array}{l} N_1 \\ \sum_{i=1}^{N_1} a_i (y_{1k}^N)^i + \sum_{j=1}^{N_2} b_j (y_{2k}^N)^j + k_1 y_{1k}^N y_{2k}^N + k_2 \\ i \neq 3 \end{array} \right\} \quad (115)$$

The actual surface of the m -th visited candidate is then defined as:

$$\hat{x}_{1km} = \hat{C}_{1k} + a_{3m} \text{Max}(|y_{1k}|, |y_{2k}|) (y_{1k}^N)^3 \quad (116)$$

The processing load of \hat{C}_{1k} calculation approximately corresponds to calculation of one typical approximation surface built according to the proposed surface model and denominated as P_M . If the surface is described with M degrees of freedom, and the calculation of any coefficient requires the same amount of processing load then the surface based on m coefficients requires m/M part of the entire surface calculation. Hence the processing load of one GML precision level search, denoted as P_L corresponds to:

$$P_L = P_M m_L \left[\frac{1}{M} N_{VCPL} + \frac{m_{L-1}}{M} \right]; \quad m_0 = 1 \quad (117)$$

where N_{VCPL} is the number of visited points, m_L the number of used coefficients at the corresponding level, M the number of coefficients of the final surface model, and P_M the processing load of one calculation of surface with maximum number of degrees of freedom. The processing load corresponding to the entire GML search is equal to the sum of processing loads of all implemented precision levels:

$$P_{GML} = \sum_L P_L \quad (118)$$

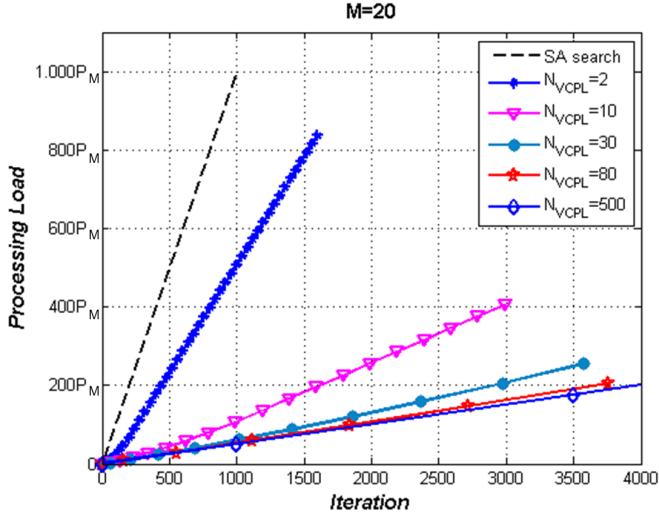


Figure 82: The processing load of GML search with different N_{VCPL} number

If the same number of visited points per temperature level is used with SA or IFSA approach, the processing load is expressed as:

$$P_{SA} = P_{IFSA} = P_M N_{VCPL} L \quad (119)$$

where number of temperature levels is L . It should be pointed out that the number of temperature levels in SA does not correspond to the number of precision levels in GML search as, for example, 1700 iterations correspond to 6 precision levels in GML, and 34 temperature levels of SA search.

The difference between the processing load can be seen more clearly in Figure 82, where the processing load expressed in P_M is depicted as a function of number of iterations. The figure shows the processing load of the original SA approach together with GML search based on different N_{VCPL} number applied to surface with 20 degrees of freedom. The processing load of SA search carries no optimization hence, the number of iterations corresponds to the number of complete surface calculations. The GML approach that exhibits the lowest processing load reduction is obtained with $N_{VCPL} = 2$, which corresponds to the smallest number of visited candidates per level that are required for correct functioning of GML approach. However, even in this case the processing load decrease of approximately 50% is achieved as 1000 iterations are transformed into 500 full surface calculations. Further on, as N_{VCPL} number is increased the processing load requirements drop considerably as for $N_{VCPL} = 80$ almost 3800 iterations

are transformed in approximately 200 full surface calculations, that is equal to 200 iterations in SA approach. Hence the processing load in this GML approach reduces to almost 5% the processing load of original SA approach. If N_{VCPL} is increased to 500, the processing load continues to fall but even though this enhancement could be justified mathematically, it makes no sense inside the GML approach as in this case the introduction of all 20 coefficients would require 49500 iterations which would put unnecessary weight on GML search. The adequate N_{VCPL} number, together with space reduction factor β , and the initial search space width W , will be established in the following section.

6.3.2 Adopting GML search to nonlinear decoupling

In order to accurately describe the GML search three parameters have to be defined: initial search space, number of visited points per level and the reduction factor. Since GML approach exhibits low processing requirements all parameters are selected only with respect to the obtained solution precision ignoring the number of required iterations. The parameters are determined with simulation using 64QAM Gray modulated signal in AWGN channel under strong and moderate nonlinear coupling conditions. Like in previous simulations, the coupling parameters on both antennas are set to same values and are expressed as $[C_{11}, C_{12}, C_{13}]$ in dB. According to surface fit equation defined in eq.(81) the degrees of freedom of coupling cancellation module are set to 20 with $(N_1, N_2) = (9, 9)$. The calibration is conducted with SNR value set to 100dB and based on 2000 calibration signals with the presented curves corresponding to the average of ten GML search executions. The number of iterations corresponds to the number of times the Q function is calculated. When determining the search parameters the starting search domain is determined first, followed by space reduction factor and leaving the number of visited candidates per level to be determined last.

$\Rightarrow W$

The initial search domain is selected as the first factor to be established since the other two factors are directly related to this one. When determining the initial search space it should be taken into account that the initial decoupling configuration corresponds to ideal signal transmission without any distortion. Thereby, the coefficients are initialized favouring the small coupling distortions which requires the search domain wide enough to offer the solution to strong nonlinear distortions. However, it should also prevent the search from drifting

pointlessly around the space when light coupling is experienced. If the initial search domain is selected adequately, the search should produce the results of the same precision regardless of N_{VCPL} parameter with only difference located in the number of required iterations. Since the search is abandoned when the enhancement in the cost function is smaller than 1% between two consecutive search precision levels, the reduction space factor β should be selected based on the previously determined initial search space. If the space is reduced slowly, the GML search can produce precise results only if the initial search domain is small and located close enough to the solution. However, if large reduction space reduction is implemented on this initial search domain, GML can easily get stacked in some local minimum. On the other hand, if the initial search space is wide, the small reduction factor can produce enhancements smaller than 1% which stop the search. Hence this space requires faster space reduction.

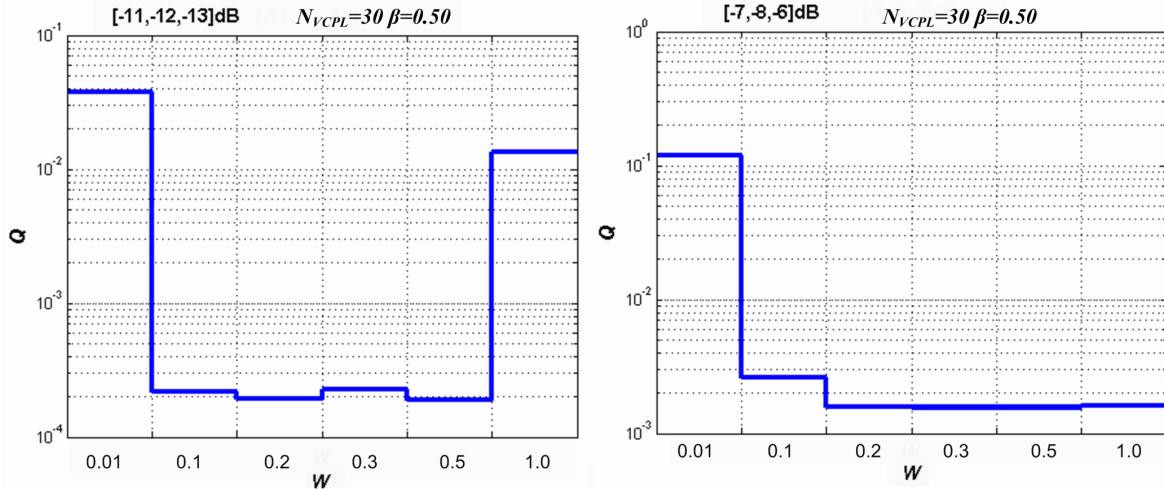


Figure 83: The precision of cost function obtained with different W factor

The simulations in Figure 83 depict the obtained precision level of the approximated surface based on different initial search area. The search is performed with $N_{VCPL} = 30$, and $\beta = 0.5$ under moderate and strong coupling conditions. The figure shows that the simulation with small initial search space of $W = 0.01$ exhibits low search ability for both distortions levels. This is explained with the location of the ideal surface configuration which is situated outside the initial search space, and can hardly be reached. Under strong coupling conditions this occurs even for $W = 0.1$ but with minor consequences. Contrary to small search space, the wide initial space showed excellent results under strong coupling, but disappointed under moderate coupling. This appeared again due to the location of the ideal surface configuration which is far from the initial search point under strong coupling and hence wide search area

favours its location. However, as the initial search space gets wider more precision levels have to be implemented in order to reach the desired solution. Hence, as the search path is increasing the probability of encountering local minimum is also increased. This is precisely the case which occurs under moderate coupling, and since this is not just based on one search execution but on the average of 10 search instances, reaching local minimum can not be described with low probability. Based on this two figures of cost function precision the initial search space should be selected in the interval $\langle 0.2, 0.5 \rangle$ as those parameters obtained lowest cost function results under both coupling distortions.

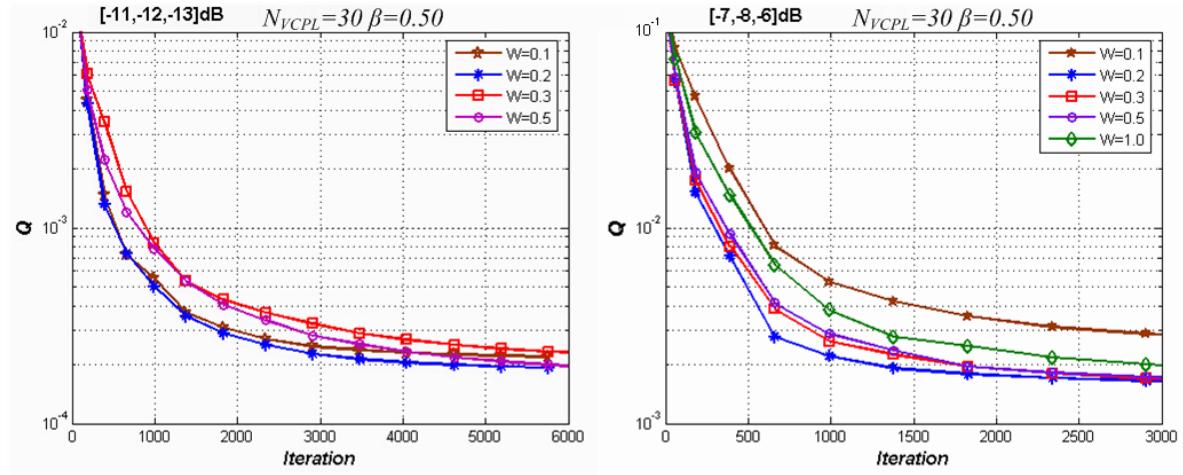


Figure 84: Cost function in dependence of number of iterations for different W factor

Figure 84 presents the dependence of the cost function with regard to number of iterations under different values of factor W . The results correspond to the same simulations as presented in previous figure. However, the final precision of cost function is not the same in both figures as in Figure 83 the number of iterations is not limited, while in Figure 84 not the whole search process is depicted. This is made in order to easily detect the function with steepest descent factor. In other words, the curve obtains lowest cost function value under lowest cost, which corresponds to the smallest number of iterations. This is achieved with $W = 0.1$ and $W = 0.2$ under moderate coupling but even though the curves are very similar $W = 0.2$ reaches better final precision. Similar situation appears under strong coupling where the best results are obtained with $W = 0.2$. In this case smaller search space of $W = 0.1$ is discarded due to its poor precision performance. Thus, the adequate selection of the initial search parameter is obtained by setting the corresponding GML parameter to $W = 0.2$.

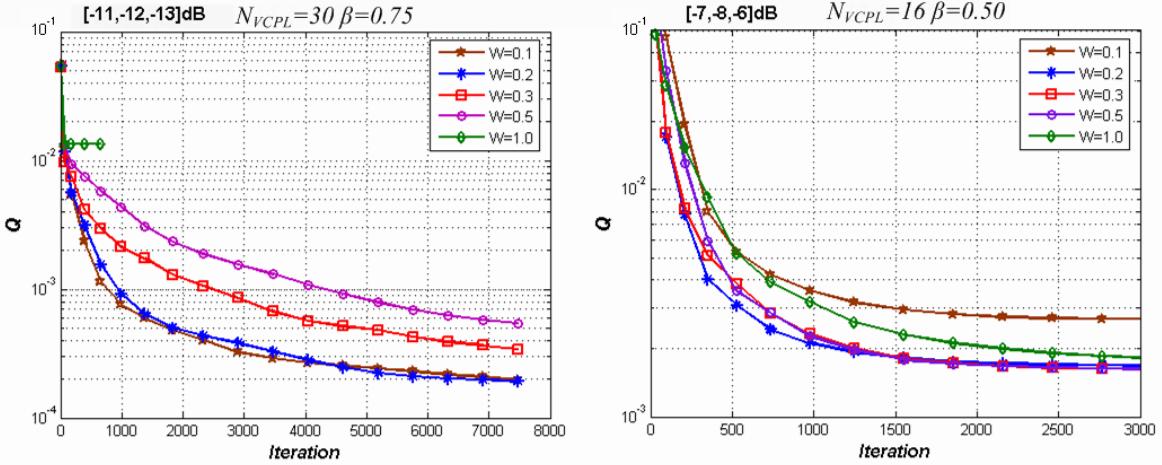


Figure 85: Q in dependence of number of iterations for different GML search parameters

The appropriate selection of the initial search space is confirmed with two more sets of simulations presented in Figure 85. The simulations on the right correspond to GML searches implemented with different space reduction factor under the same moderate coupling distortion as in previous simulations set. The speed of search space reduction is here decreased as β changes from 0.5 to 0.75. As a consequence the search converges slowly but the curves are positioned in the same order as in previous simulation. Large search space defined with $W = 1.0$ which is shown in this figure, is still stacked in local minimum but in this case the search lasts a few precision levels more. The GML based on $W = 0.3$ and $W = 0.4$ is here not so close to $W = 0.2$ which is actually almost identical as $W = 0.1$ and exhibits the best results. Hence, that initial search space of $W = 0.2$ exhibits excellent results under different space reduction factors and this verifies its selection.

The influence of the number of visited candidates per level on the initial search space factor can be obtained comparing the left images on Figures 84 and 85 which correspond to the same GML search, under the same, strong coupling distortion, implemented with different N_{VCPL} factor. Eventually, the curve distribution is the same with only difference found in the obtained cost function value after the first 500 iterations which is lower for $N_{VCPL} = 16$ than for $N_{VCPL} = 30$. However, the final precision is very much alike with $W=0.1$ showing the worst and $W = 0.2$ the the best performance results. Thus in this case the initial search space would also respond to $W = 0.2$ and this level will be used in following simulations.

$\Rightarrow \beta$

The reduction of search space is the key factor for obtaining the robust GML search performance. If β is close to one the search space is reduced slowly and the search progress is also slowed down. According to theory this approach should also produce excellent surface approximation configurations based on large number of iterations. However, since the search spaces of two consecutive levels are very alike the candidates of the new precision level are situated close to already tested configurations of the previous level. As GML algorithm is suspended when the precision enhancement between two consecutive search levels is lower than 1% , the search based on slow space reduction rarely reaches the global minimum. This is seen in Figure 86, where the behaviour of the GML search based on different space reduction factors under strong coupling is presented. The image on the left corresponds to the obtained precision regardless of the number of iterations, while on the right the dependence of cost function of the number of iterations, limited to 3000, is depicted. Both figures clearly show that the space reduction factor larger than 0.9, leaves the search in some local minimum as the presented Q function is lower than the one obtained with other β factors.

If the search space is reduced too fast the search space is not explored adequately, and the paths leading towards the optimal solution are lost in reduction phase. This is confirmed in the first two values of β which offer the highest value of Q function as a solution. Consequently, the adequate search space reduction factor is located in the interval $\langle 0.4, 0.6 \rangle$. These levels

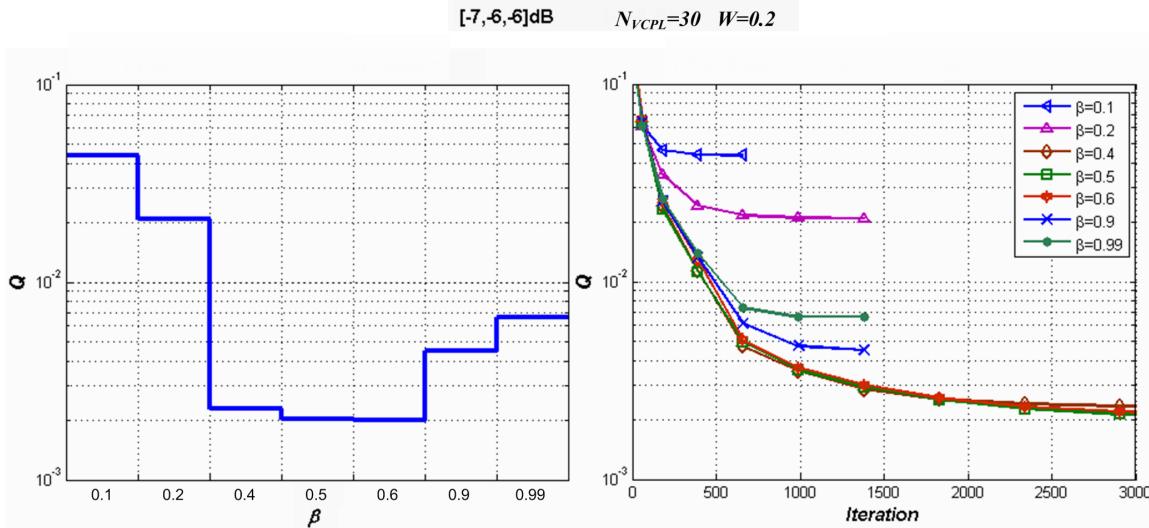


Figure 86: The behaviour of GML search with different space reduction factors

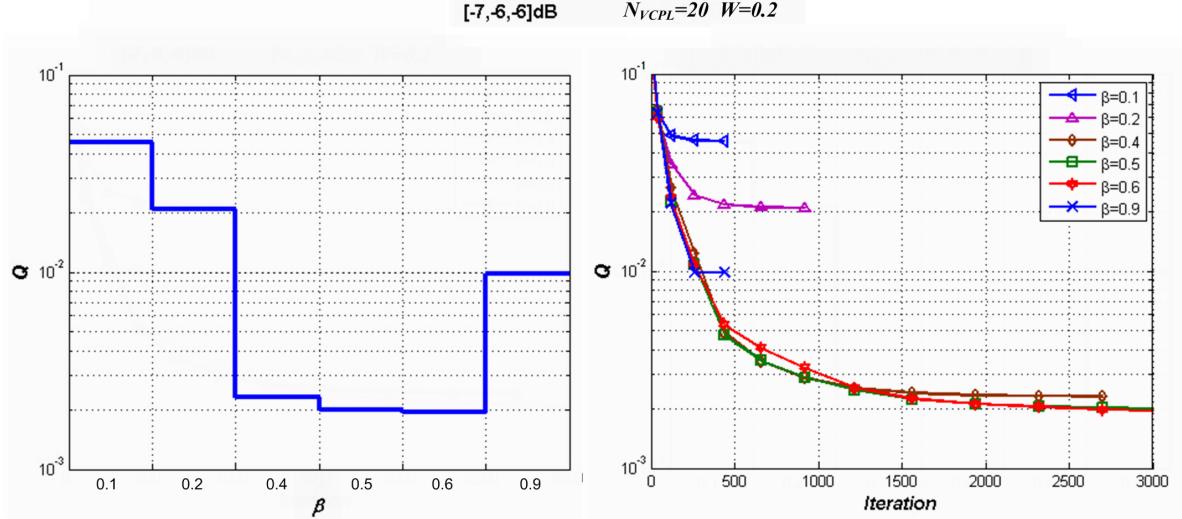


Figure 87: The behaviour of GML search with different space reduction factors

produce similar precision of the cost function and the behaviour of cost function draws almost identical curves with respect to the number of iterations as seen in the figure on the right.

The simulations in Figure 86 are based on initial search space set to $W = 0.2$ and number of visited candidates per level set to $N_{VCP}L = 30$. If the $N_{VCP}L$ factor is changed the behaviour of the GML search under the same distortion levels, is slightly changed. This is seen in Figure 87, where the search with lower $N_{VCP}L$ factor is implemented under strong coupling conditions. The precision of cost function with $N_{VCP}L = 20$ shows exactly the same behaviour as in previous simulations obtained with $N_{VCP}L = 30$. Hence, reducing the search domain with factor lower than 0.2, which corresponds to preserving only 20% of the previous search space, produces higher Q values regardless of number of visited candidates per level. The same conclusion can be obtained for slow space reduction higher than 0.9. The best Q results are again obtained with β between 0.4 and 0.6. However with lower $N_{VCP}L$ number the three presented curves of $\beta = 0.4$, $\beta = 0.5$ and $\beta = 0.6$ are here not so close. In fact in the range of 500 to 1000 iterations the curve corresponding to $\beta = 0.6$ exhibits inferior results than the other two factors. On the other hand, when the number of iterations is larger than 1500, the curve obtaining the worst results is the one corresponding to $\beta = 0.4$. Thus, based on these two set of simulations the search space reduction factor tuned to $\beta = 0.5$ adapts the best to nonlinear coupling cancellation problem regardless of the number of visited candidates per level.

$\Rightarrow N_{VCPL}$

Number of visited candidates per precision level defines the number of configurations which are tested for one decoupling coefficient at the corresponding precision level. Theoretically, with large number of visited candidates the algorithm is more robust to the local minimum pitfalls. However, this robust property is achieved with increased processing load but the final precision of the approximated surface is not enhanced. On the other hand, if the number of visited candidates per level is small, the search will progress at high speed but as the search path only has a few options at each precision level it will offer local minimum as the final solution. These two N_{VCPL} characteristics are easily seen in Figure 88 where the behaviour of GML search based on different N_{VCPL} factor is presented under strong coupling distortion.

The image presented on the left of Figure 88, shows the obtained Q function based on different N_{VCPL} factors and without any iteration restrictions. Apparently, only the search based on $N_{VCPL} = 2$ exhibits poor performance under this coupling distortion, while all GML searches based on any larger number of candidates achieve significantly better and similar precision levels. The search convergence speed is seen on the image on the right where $N_{VCPL} = 4$ clearly reaches the best results as the same cost function level is achieved with only 700 iterations. The iteration cost is increased to 2700 for $N_{VCPL} = 16$, to 5100 for $N_{VCPL} = 30$, 7700 for $N_{VCPL} = 50$ and, finally, 11500 for $N_{VCPL} = 100$.

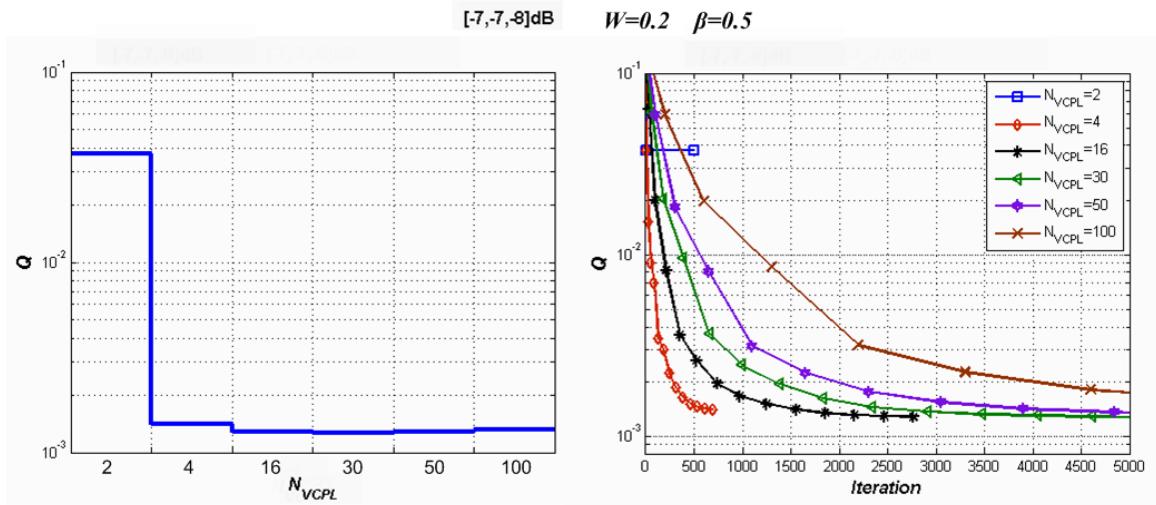


Figure 88: The behaviour of GML search based on different N_{VCPL} factor under strong coupling

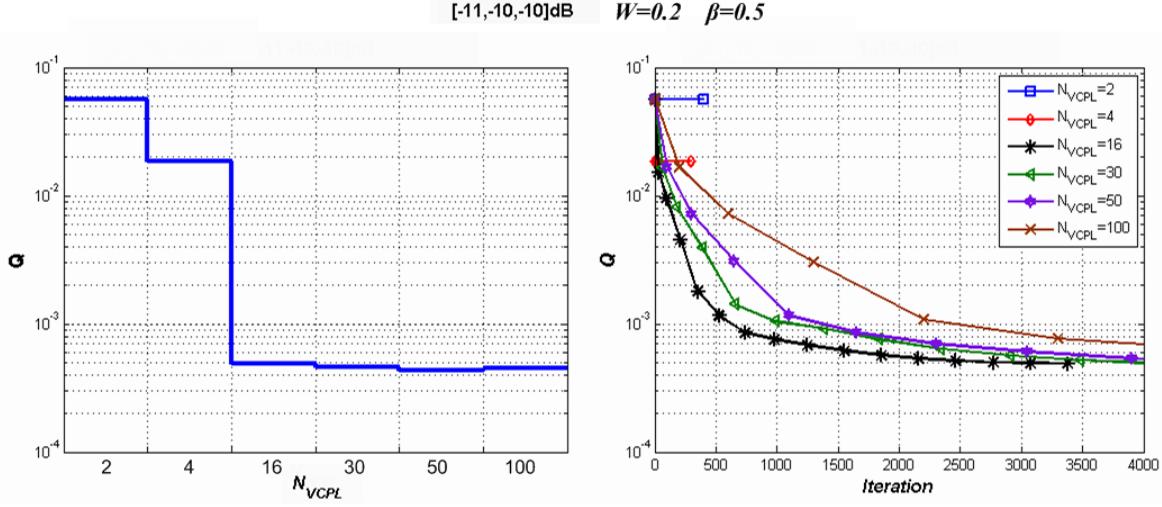


Figure 89: The behaviour of GML search based on different N_{VCPL} factor under moderate coupling

If the GML search with the same set of parameters is implemented on the moderate coupling of $[-11, -10, -10]$ dB the results are changed considerably. This is seen in Figure 89, where the GML search based on $N_{VCPL} = 4$ is stuck in local minimum and the best performance corresponds to $N_{VCPL} = 16$. According to these images dividing the search space into only four candidates produces excellent results in some conditions but as it offers no robustness it is easily stuck in local minimum and offers poor performance in other conditions. Consequently, the number of visited candidates per level should be in the interval $\langle 4, 16 \rangle$, offering robustness and requiring low number of iterations.

In order to select the appropriate number of visited candidates per level, two additional sets of simulations are performed as presented in Figure 90. The image on the left correspond to GML search under strong coupling of $[-8, -6, -7]$ dB while the image on the right is obtained under moderate coupling of $[-12, -10, -13]$ dB. The purpose of these simulations is to confirm the robustness of GML search based on higher N_{VCPL} factor and to select its lowest value which successfully avoids all local minimum traps. According to presented strong coupling conditions, the best results are obtained with $N_{VCPL} = 10$ as this option exhibits the same angle of descent as the curves obtained with $N_{VCPL} = 4$ and $N_{VCPL} = 16$, but the decline lasts longer and consequently lower Q function is achieved. The situation with respect to presented moderate coupling is not this obvious as the steepest descent is obtained with $N_{VCPL} = 4$ which ends up in the local minimum. This curve is followed by $N_{VCPL} = 8$ which

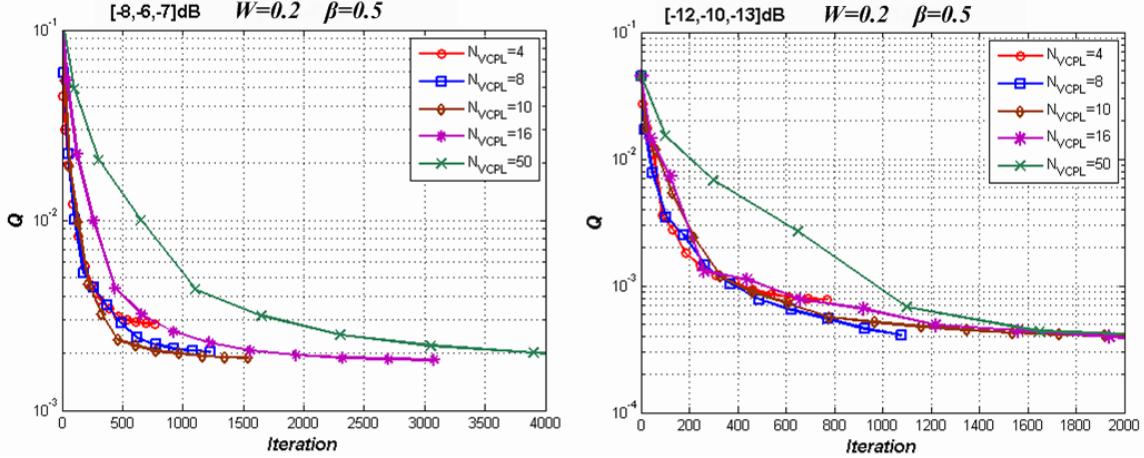


Figure 90: The behaviour of GML search based on different N_{VCPL} factors

shows excellent results for less than 1000 iteration. The third curve is the one obtained with $N_{VCPL} = 10$ which need almost double number of iterations to reach the level obtained with $N_{VCPL} = 8$. Eventually, the selection of adequate number of visited candidates per level is reduced to 8 or 10 configurations per decoupling coefficient. Since larger N_{VCPL} number offers higher robustness the GML search will be based on $N_{VCPL} = 10$.

⇒GML Search implementation example

In order to test the behaviour of the GML search based on parameters adjusted to nonlinear decoupling problem, the corresponding search is executed 10 times on moderate coupling of $[-10, -12, -12]$ dB. The search is done with calibration size set to 2000 symbols. Maximum number of iterations is not limited as the search is abandoned if two consecutive precision levels don't produce cost function enhancement of more than 1%.

The behaviour of GML approach is depicted in Figure 91, where Q value is presented as a function of the number of iterations for 10 instances of GML search under same operating conditions. Eventhough the curves can not be clearly distinguished there are exactly ten of them present in the image. The fact that the curves are grouped so tightly confirms the robustness of the GML search which is even more emphasized when number of iterations and obtained precision are taken into account. In this example, all ten search executions are stopped after 2300 iterations and all of them are stopped in the same cost function area of around 3×10^{-4} . In other words, all ten GML based searches used the same number of

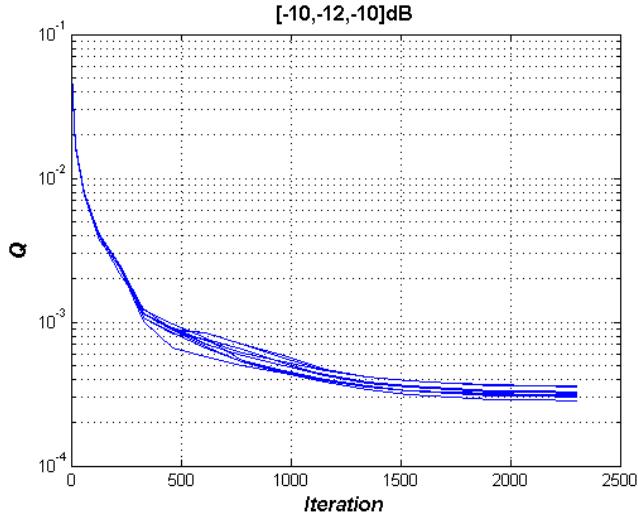


Figure 91: Ten instances of adjusted GML search under moderate coupling

precision levels to reach the same precision of the approximated surface which exhibits the search consistency and high robustness. Furthermore, it required the processing load equal to approximately tenth part of the approach based on SA or IFSA algorithm with the same amount of iterations which makes this method more adequate for practical use under limited mobile device conditions.

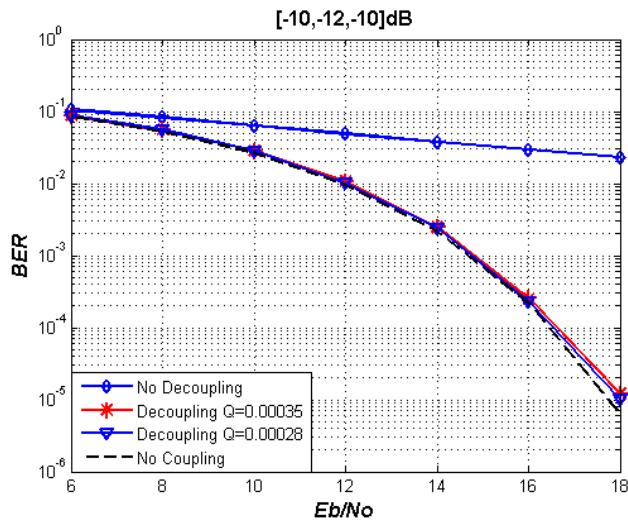


Figure 92: Performance of GML based decoupling module under moderate coupling

The performance of the decoupling module shows excellent decoupling results as shown in

Figure 92 where two decoupling surfaces exhibiting highest and lowest Q functions after 2300 iterations are implemented. Two decoupling modules are almost perfectly matched and are both very close to the theoretical curve without coupling. This confirms the correct functioning of the proposed GML search method and justifies its use in decoupling module.

⇒Conclusion

The proposed GML search algorithm is developed for decoupling approximation surface problem as the need for a simple, low processing, and fast algorithm emerged. The GML approach is adopted for the presented decoupling problem but in fact it can be used in any approximation problem dealing with many degrees of freedom. The most important characteristics of GML approach which clearly distinct it from SA and IFSAs methods is that it classifies the coefficients by weight factor into categories. The categories follow the surface approximation model which is constructed as a sum of two independent higher order polynomials, assigning one category to one order. Since the decoupling surface is restrained polynomials of ninth order, GML approach consists of 10 categories with the tenth one being the constant factor with lowest weight. The importance of all other categories is inversely proportional to their exponential order.

Another contribution of this method lies in the gradual introduction of the the corresponding categories into the surface model and focusing the search on one coefficient at the time. This approach enables the reduction of average processing load required for one cost function calculation, it simplifies the search process and consequently increases the convergence speed. The behaviour of proposed GML algorithm presented excellent decoupling efficiency obtained with number of iterations similar to IFSAs approach. However, when processing demands of these two methods are compared the GML approach is clear winner requiring only one tenth of IFSAs precessing demands. Based only on processing requirements this method is favourite for practical real time decoupling problems. Eventually, after GML search is adopted to the RF front-end decoupling problem, the algorithm is included in decoupling module as fast, precise, and low demanding search, capable of reaching precision limits set by SA methods.

7 Noiseless Calibration

Nonlinear systems are typically non-typical. Thus, it is expected that some algorithms are better suited for some systems than for others. In previous chapters three over-dimensioned decoupling module algorithms have been analyzed and adopted to general RF nonlinear coupling problem. The analysis offered a general insight of the capabilities of each approach including its decoupling precision, convergence speed and processing load.

Once the behaviour of the proposed search algorithms is analyzed, some specific properties of RF front-end decoupling problem can be included in order to simplify the system. The aim is to decrease the number of decoupling coefficients keeping the system BER performance as close as possible to the search based on 20 degrees of freedom. In the first section the simulations are carried out based on 2000 calibration samples. Afterwards, the calibration size is changed in order to see the reaction of the system based on less survey points. At the end, adequate calibration size with respect to decoupling precision, processing load and system robustness is proposed.

7.1 Decoupling Degrees of Freedom

In previous chapter the parameters of the decoupling module search algorithms are obtained based on the wide range of coupling conditions applied on the 64QAM signals with all system parameters over-dimensioned in order to develop precise decoupling approach without additional constraints in the analysis. Based on these premises, the SA search is implemented eventually without strict processing constraints, the number of symbols used for calibration purposes is set to $N = 2000$ and the decoupling surface is constructed with 20 degrees of freedom. Based on the same calibration size and decoupling surface form, it has been shown that IFSA and GML methods achieve similar surface approximation quality but with much lower processing demands. The next parameter which has to be adopted to real transmission system requirements is the number of decoupling coefficients which compose the decoupling approximation surface.

The direct coupling function of a real transmission system is defined as in eq.(27) as :

$$\begin{aligned} f_1(x_2, Coup_{21}) &= c_{11}x_2 + c_{12}x_2^2 + c_{13}x_2^3 \\ f_2(x_1, Coup_{12}) &= c_{21}x_1 + c_{22}x_1^2 + c_{23}x_1^3 \end{aligned} \quad (120)$$

with c_{ij} corresponding to the coupling of the j^{th} order experienced on the i^{th} coupling module input. As already analyzed, if the transmission system is exposed to this coupling function the decoupling surface that have excellent coupling cancellation is constructed of 9^{th} degree polynomial according to eq. (81) as:

$$\hat{x}_{1k} = \text{Max}(|y_{1k}|, |y_{2k}|) \left\{ \sum_{i=1}^{i \leq N_1} a_i (y_{1k}^N)^i + \sum_{j=1}^{j \leq N_2} b_j (y_{2k}^N)^j + k_1 y_{1k}^N y_{2k}^N + k_2 \right\} \quad (121)$$

where y_{1k}^N and y_{2k}^N are the k -th normalized received symbols on the corresponding antennas, $\text{Max}(|y_{1k}|, |y_{2k}|)$ is the maximum received absolute value level, N is the number of calibration signals and variables a_i, b_j, k_1, k_2 are the decoupling coefficients. The variables denoted as N_1 and N_2 are restricted to maximum value of 9, and represent degrees of freedom of variables y_1 and y_2 respectively. The set of parameters (N_1, N_2) will be changed in following simulations in order to analyze the required complexity of the decoupling surface approximation. The simulations are based on the SA approach as it offers highest precision and robustness. The calibration size is set to 2000, and the curves correspond to average level of ten independent executions.

The performance of SA algorithm implemented on decoupling surface constructed with different degrees of freedom under moderate coupling of $[-10, -5, -10]$ dB is seen in Figure 93. The image on the left corresponds to the Q function obtained with (N_1, N_2) set of decoupling coefficients. The total number of decoupling coefficients in simulations is $N_1 + N_2 + 2$, since k_1 and k_2 are used in all analysis. If the sufficient cost function value is regarded as $Q_{suff} = 0.005$, than beside the full load SA approach, the one based on coefficients set $(8, 7)$ should also obtain excellent results. However, as the number of decoupling coefficients is decreased the surface precision deteriorates and cost function drops to approximately $Q = 0.0014$ for $N_1 = 1$ and $N_2 = 3$.

According to cost function level only two decoupling surface forms $((9, 9)$ and $(8, 7))$ are expected to offer excellent decoupling solution. However, the corresponding BER performance

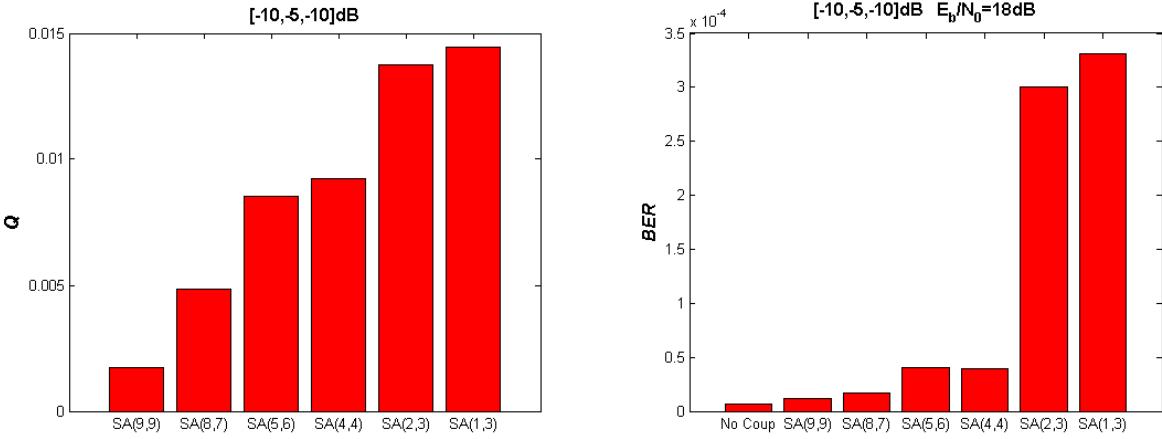


Figure 93: The performance of SA approach based on various decoupling surface forms

of the implemented decoupling surfaces depicted on the right image, together with the theoretical BER limit without coupling, shows that even four out of six surfaces might be used in this coupling case. The system BER performance corresponds to noise level of $E_b/N_o = 18$ dB. As expected BER of (9,9) and (8,7) show excellent characteristics with system performance close to the theoretical limit. The system based on next two decoupling surfaces experience slight deterioration but can still be regarded as excellent. Nevertheless, the last two surfaces based on six and seven decoupling coefficients exhibit severe precision drop and can not be recommended for precise decoupling solution. This is confirmed in Figure 94, where

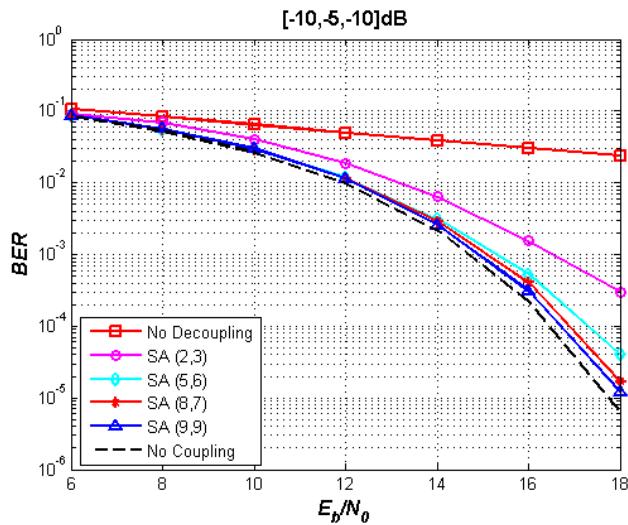


Figure 94: BER performance of different decoupling surface forms

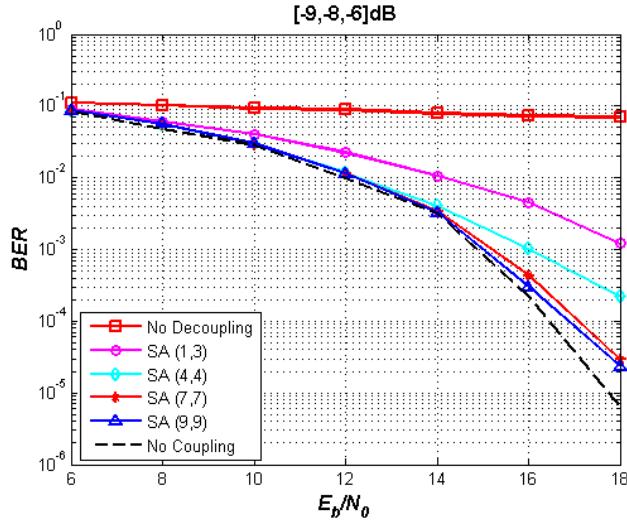


Figure 95: BER performance of different decoupling surface forms

BER performance under different E_b/N_0 noise levels is shown. In order to easily distinguish the curves, only four surface forms are depicted, together with theoretical curve and system performance without decoupling module. The presented results reaffirm the full complexity surface as the one adapting the best to the decoupling problem, but also prove that excellent coupling cancellation can be accomplished with fewer degrees of freedom. Apparently, under moderate coupling of $[-10, -5, -10]$ dB excellent coupling cancellation can be achieved starting with $N_1 = 4$ and $N_2 = 4$, which means all together 10 degrees of freedom. Eventually, the additional degrees of freedom contribute to final decoupling precision but their introduction in this case can be dependent on the required system decoupling precision and the available processing power.

If stronger cubic coupling, like $[-9, -8, -6]$ dB appears, the behaviour of BER curves is not identical to the one obtained under $[-10, -5, -10]$ dB coupling. As depicted in Figure 95, in this case the system based on ten decoupling coefficients denoted as $(N_1, N_2) = (4, 4)$ is not sufficient for excellent coupling cancellation, since decoupling surfaces constructed with degrees of freedom smaller than 16 are not located close to the theoretical limit case without coupling. This means that the precision of simpler decoupling surface depends directly on the suffered coupling distortion. Constructing the decoupling surface with smaller degrees of freedom favours the speed and processing requirements, but at the same time jeopardizes the system robustness as less complex approximated surface can not be adopted to all coupling conditions with the same efficiency. Actually, even very complex surface with 20 coefficients can guarantee

excellent decoupling performance in every strong coupling environment, and it would not be logical to expect much simpler surface to cover the whole coupling distortion area with the same precision. This conclusion is also reached in [Ara07a] where SA approach is implemented with different surface forms. The corresponding coupling distortions in this article use different coupling strength measurements taking into account only numerical coefficient ratios, hence from the numerical point of view, coupling of amplitudes $[-20, -30, -40]$ dB in [Ara07a] is equal to approximately $[-10, -9, -7]$ dB coupling strength calculated in this thesis.

Eventually, smaller number of decoupling coefficients should be used only when direct coupling functions are known and the decoupling precision of proposed surface form can be calculated in advance through simulations. Since analysis in this thesis is based on generic case of strong and moderate coupling distortions, the simplification of decoupling surface based on the simple neglecting of higher exponents is rejected in order to preserve the system robustness. However, up to now, the transmission system was observed as a simple transmitter-channel-coupling-receiver system without any specific conditions taken into account, but some system specific characteristics might relax the required number of surface degrees of freedom. For example, concentrating on the receiver based on I/Q downconverter, the square coupling distortion does not have influence on the system, as was analyzed and demonstrated mathematically in chapter 4.3.

If direct coupling function is defined as in eq.(27) and the coupling square parameter, which has no influence on the received signal, is neglected, the corresponding mathematical problem is defined as:

$$\begin{aligned} y_1 &= x_1 + c_{11}x_2 + c_{13}x_2^3 \\ y_2 &= x_2 + c_{21}x_1 + c_{23}x_1^3 \end{aligned} \tag{122}$$

Following the same logic, if the square distortion can be ignored in transmission there is no need to intend its cancellation with the decoupling surface approximation. Hence, all parameters related to the square exponential can also be neglected and qualified as not important. This leads to decoupling surface constructed of 9th degree polynomial, with all coefficients alongside even exponents set to zero. In other words, according to eq. (81), the decoupling surface is defined as:

$$\begin{aligned} \hat{x}_{1k} &= \text{Max}(|y_{1k}|, |y_{2k}|) \left\{ \sum_{i=1}^{i \leq N_1} a_i (y_{1k}^N)^i + \sum_{j=1}^{j \leq N_2} b_j (y_{2k}^N)^j + k_1 y_{1k}^N y_{2k}^N + k_2 \right\} \\ a_2, a_4, a_6, a_8, b_2, b_4, b_6, b_8, k_1 &= 0 \end{aligned} \tag{123}$$

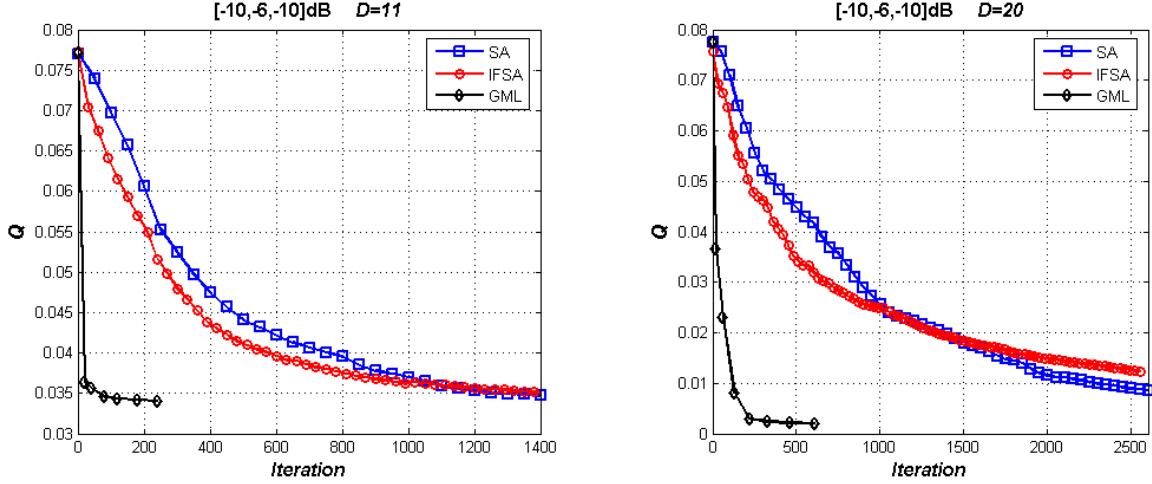


Figure 96: Convergence speed of SA, IFSA and GML with different degrees of freedom

where y_{1k}^N and y_{2k}^N are the k -th normalized received symbols on the corresponding antennas, $\text{Max}(|y_{1k}|, |y_{2k}|)$ is the maximum received absolute value level, and N is the number of calibration signals. Using the same notation as in previous examples, the variables a_i, b_j, k_1, k_2 are the decoupling coefficients. Maximum values of N_1 and N_2 which represent degrees of freedom of variables y_1 and y_2 respectively are restricted to 9. Hence, in this specific case when all even exponents are ignored, the degrees of freedom are reduced to 11 which are to be determined inside the surface approximation search.

In order to evaluate the proposed system simplification, and the corresponding reduction of degrees of freedom, the decoupling module based on 20 and 11 decoupling coefficients is tested under different coupling conditions. The configuration of simulated coupling environment is preserved as in all previous simulations described with three factors expressed in dB. The influence of square coupling on the surface form, and its corresponding decoupling performance, can easily be noticed if the system is submitted to coupling with strong square distortion. This is seen in Figure 96, where cost functions obtained with different degrees of freedom, under $[-10, -6, -10]$ dB coupling, are shown. The presented curves correspond to averaged results of ten search executions. The image on the left exhibits decoupling surface constructed with $D = 11$ decoupling coefficients, while the image on the right is obtained using all $D = 20$ degrees of freedom. The search is conducted for all three decoupling module approaches, using the parameters adopted to the decoupling problem in previous chapter.

As expected, regardless of the number of degrees of freedom, the form of the curves corre-

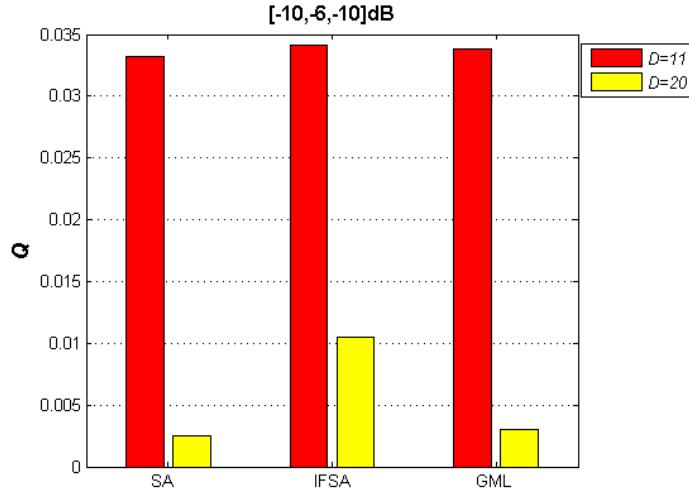


Figure 97: Precision of the three search approaches with different degrees of freedom

sponding to the same search approach are very alike. The SA approach is the slowest, while GML advances most rapidly in both images. Further on, the IFSA approach is steeper than SA in the beginning only to be driven into saturation much before the SA method. However, from the general form, the curves are quite different. In general the search progress achieved with $D = 11$ coefficients is almost twice as steep as in the module with $D = 20$ coefficients. This is seen with GML approach whose search is abandoned after 230 iterations with $D = 11$, and after approximately 600 iteration using $D = 20$ coefficients. Similarly, the IFSA is drown into saturation after around 1400 iterations on the left image, and around 2600 iteration on the right one. This is logical since all three searches process easier smaller number of unknown coefficients, and consequently route the search in correct direction without too much wandering around.

Eventhough the search progress is much faster when degrees of freedom are reduced to $D = 11$, the obtained cost function is rather poor when compared to the one exhibited with $D = 20$. This is depicted in Figure 97, where Q values of the same search approaches obtained with different degrees of freedom are shown side by side. In general, the surface precision obtained with $D = 20$ is almost ten times better for SA, approximately three times better for IFSA and eight times for GML. This appears due to the strong square coupling distortions which are not cancelled in the decoupling surface constructed with $D = 11$. Another interesting property experienced with $D = 11$ in this example is the similarity of the obtained Q values of all three approaches as all of them are located around $Q = 0.034$. This would indicate that the quality

of surface approximation is almost the same in all three methods. Also, it means that the SA search can not obtain significant improvement, when compared to IFSA, even with additional number of iterations.

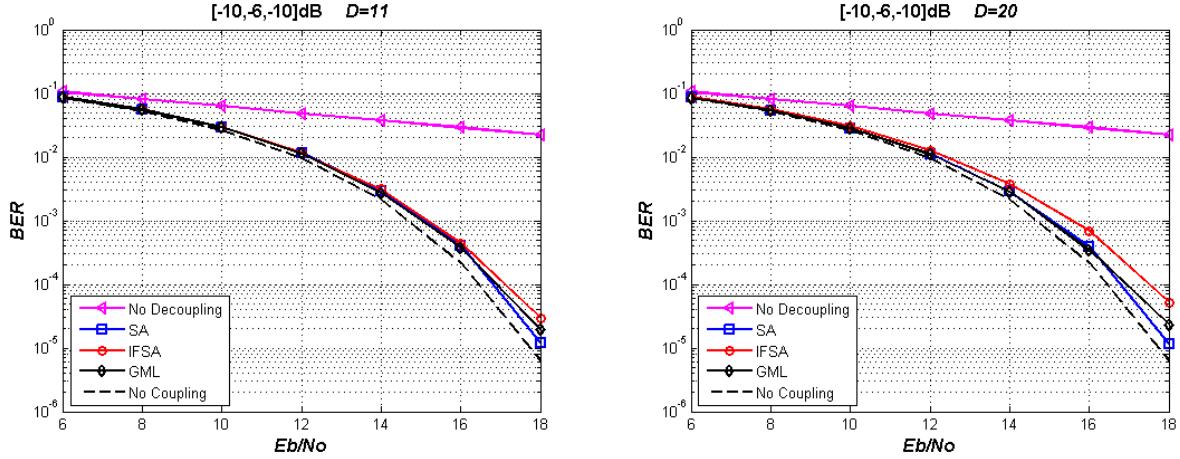


Figure 98: SA, IFSA and GML under $[-10, -6, -10]$ dB with different degrees of freedom

The effectivity of the decoupling solution is observed through system BER performance which is shown in Figure 98 for both sets of degrees of freedom. Alltogether five BER curves are shown in each image, with the lowest achievement corresponding to the system without decoupling module, and the best to the theoretical, ideal case, where no coupling is present. Eventually, the three decoupling search methods are located between these two curves in both images. According to the simulations in Figure 98, regardless of the cost function differences the performance of GML approach is preserved if even exponents are neglected. Furthermore, the search based on IFSA approach even produces better results with lower levels of freedom. This is probably due to the better initial search which can be carried out in more detail over smaller search space. Due to this reason the cost function obtained with IFSA using $D = 11$ coefficients is in the same range as the ones obtained with SA or GML approach, which is not the case with $D = 20$ coefficients.

Apparently, when moderate square distortion is decoupled based on the SA search implemented on $D = 11$ coefficients system performance is preserved when compared to the original SA search with 20 coefficients. This is confirmed in Figure 99 where SA, IFSA and GML performances under $[-12, -9, -13]$ dB coupling are shown. The simulated environment is again based on strong square coupling distortion, but this time the overall coupling is moderate.

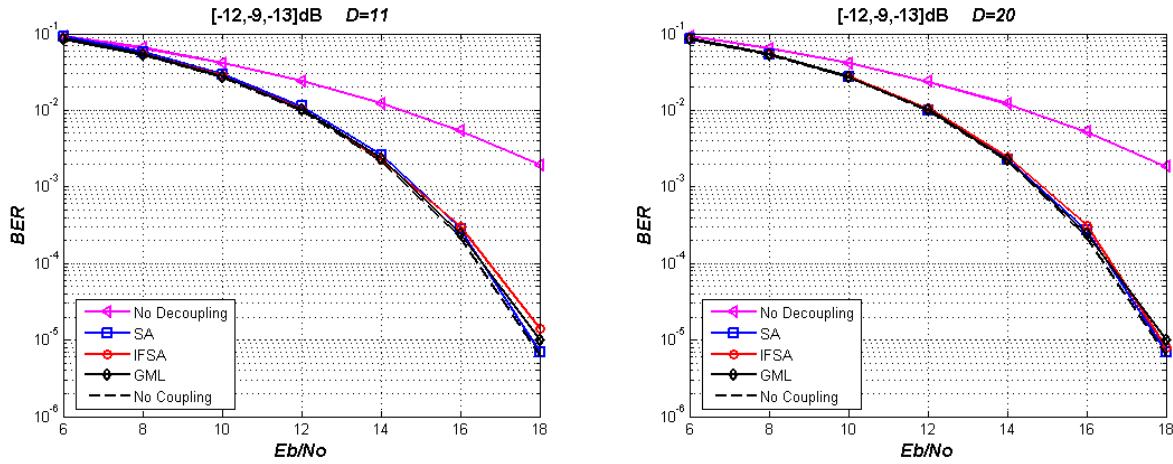


Figure 99: SA, IFSA and GML under $[-12, -9, -13]$ dB with different degrees of freedom

As can be seen in the two presented images, all three search approaches exhibit excellent decoupling results for both, $D = 20$ and $D = 11$ coefficients. In this case, only IFSA approach exhibits small BER deterioration, but this decrease is noticed only at $\text{BER} = 10^{-5}$ and does not present major concern.

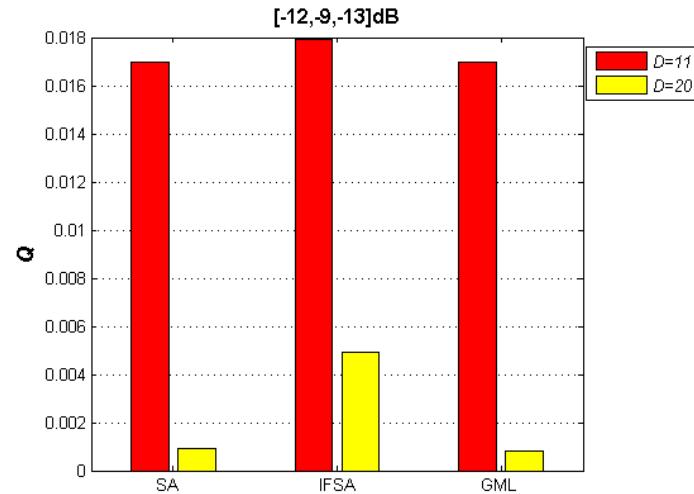


Figure 100: Precision of the three search approaches with different degrees of freedom

Under moderate coupling conditions the behaviour of the search algorithms show great similarities between the surface constructed with $D = 20$ and $D = 11$ coefficients. The search precision presented in Figure 100, shows that generally the obtained Q function is much larger

for lower degrees of freedom. Still, in both cases the SA and GML approach produced nearby results, while IFSAs approximation obtains higher cost function values.

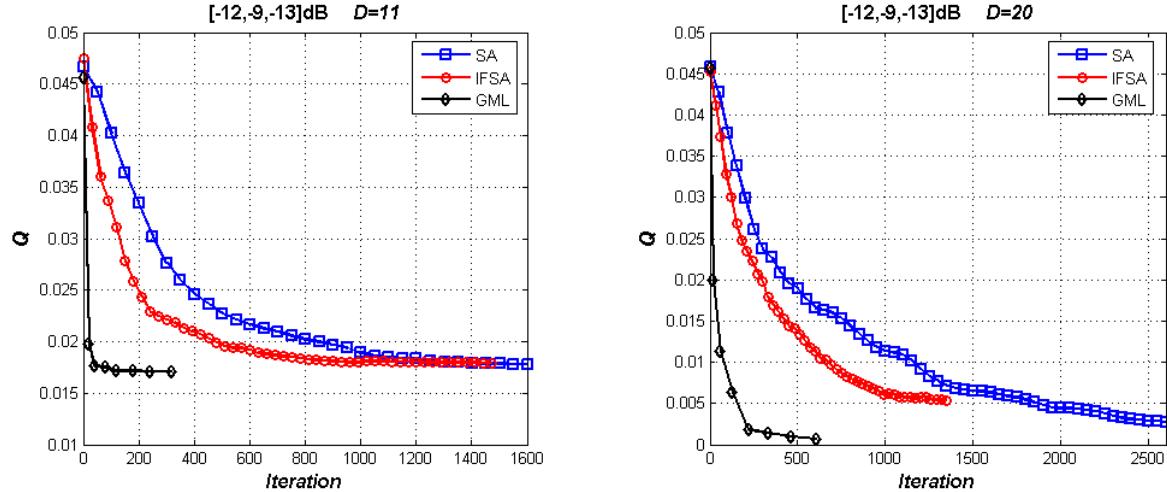


Figure 101: Convergence speed of SA, IFSAs and GML with different degrees of freedom

The resemblance between two different surface approaches is also seen in Figure 101 where the convergence speed of the SA, IFSAs and GML methods is presented. As can be noticed the curve layout follows the same formation with SA being the slowest and GML the fastest search approach. Also, as expected the convergence speed is higher with lower number of coefficients as all three methods enter the saturation phase much faster than with 20 levels of freedom. However, as for the case with $D = 11$ coefficients the saturation appears at higher Q values this indicates inferior decoupling precision. Still, as shown in Figure 99 this is not the case when moderate coupling conditions are analyzed.

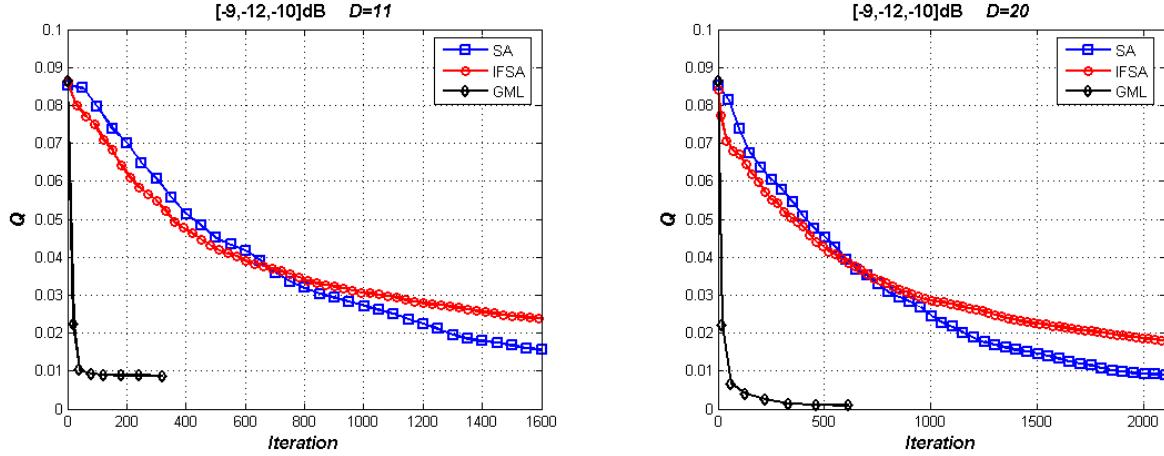


Figure 102: Convergence speed of SA, IFS, and GML with different degrees of freedom

If the three search methods are implemented for decoupling the system where square distortion parameter is not the most prominent one, the differences in BER performance between the complex 20 coefficients surface and the one with 11 are even less notable. In Figure 102, Q is depicted as a function of number of iterations for the two approximation surfaces implemented under moderate coupling. The curve formation is practically the same in both images. Hence the images can be distinguished only when number of iterations and search saturation level are taken into account.

The same conclusion can be reached for strong coupling whose convergence speed behaviour is presented in Figure 103 for $[-8, -7, -8]$ dB coupling. In general, when the approximation

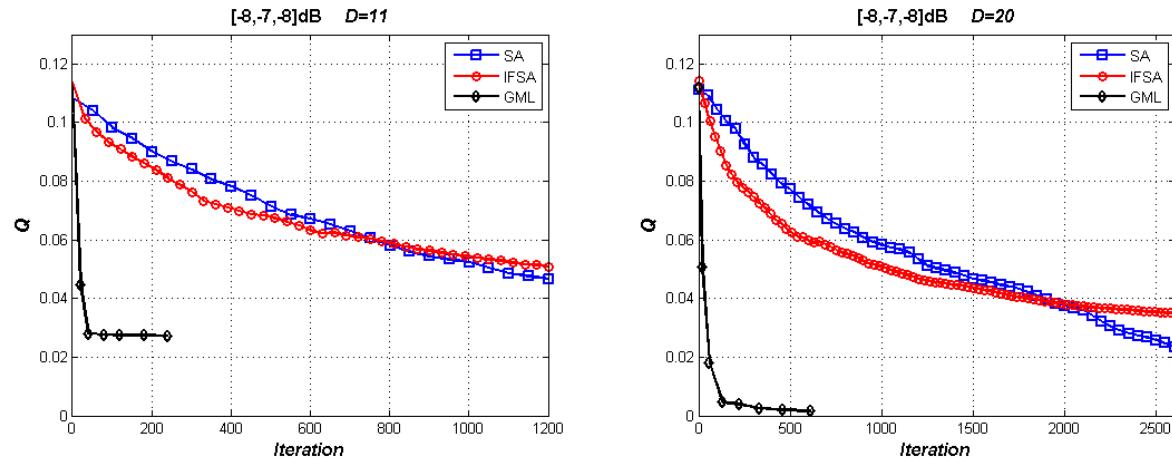


Figure 103: Convergence speed of SA, IFS, and GML with different degrees of freedom

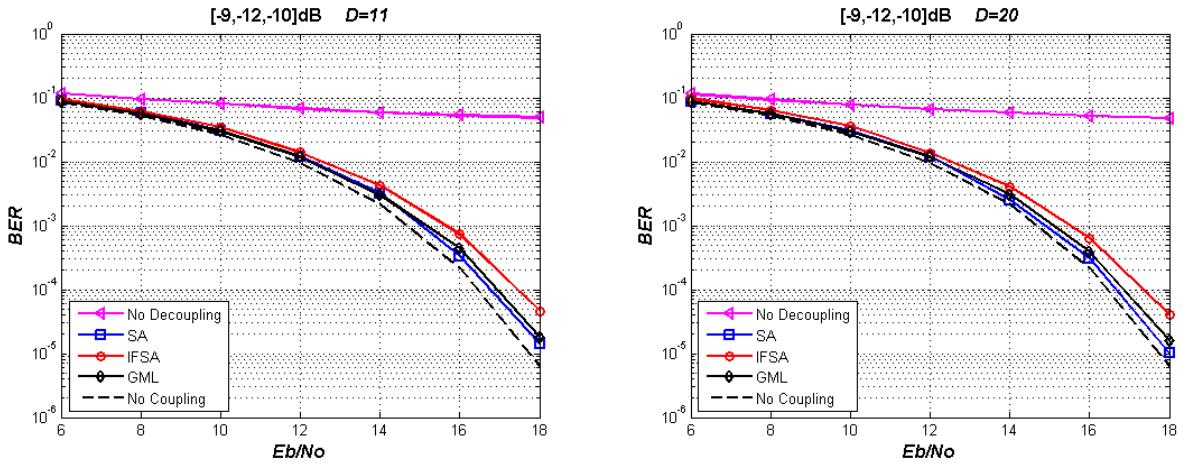


Figure 104: SA, IFSA and GML under [-9,-12,-10]dB with different degrees of freedom

surface is constructed with 11 degrees of freedom, without any even exponent, the search progress is approximately twice as fast in comparison with surface based on 20 degrees of freedom. As is shown in previous system simulations, again this difference in surface precision is not directly related to the BER performance in this case, since decoupling efficiency is completely preserved for IFSA and GML approach. This is confirmed in Figures 104 and 105, where BER behaviour is depicted for system suffering moderate and strong coupling, respectively.

In both figures, the images on the left appear as the copy of the ones on the right, which backs the theory that the even surface exponents can be neglected as square coupling influence

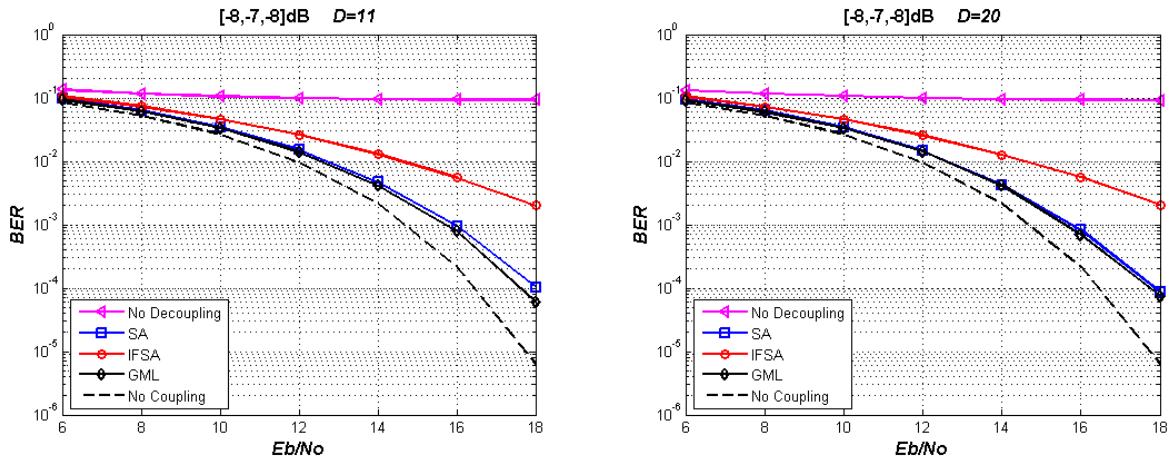


Figure 105: SA, IFSA and GML under [-8,-7,-8]dB with different degrees of freedom

is cancelled inside LPF of the I/Q downconverter. Using this approach the search progress is accelerated, the processing requirements are decreased, and the robustness of decoupling module is preserved. The drawback is that this decoupling surface is adopted to receiver based on I/Q downconverter which includes LPF in both signal branches. Actually, it is the LPF which performs square coupling cancellation and allows us to concentrate on linear and cubic coupling distortions. Due to this reason, the proposed simplification can not be implemented in receivers based on another receiver technology. Since the majority of today's receivers is based on I/Q downconverter approach, the rest of the simulations will be based on decoupling surface constructed with 11 degrees of freedom. In this case, the obtained processing requirements and decoupling precision will correspond to realistic RF Front-end conditions and can be used in mobile terminal planification phase.

7.2 Calibration size

Eventhough calibration is not performed often it carries the greatest part of decoupling module processing load requirements. Basically, once decoupling surface is established, its implementation consists of several multiplication and a few additions per each received sample, which is easily achieved under real time constraints even on a standard mobile terminal of second generation. However, the search for the decoupling set of coefficients is by far the most demanding part of decoupling module. Further on, inside this part the majority of processing time is spent on cost function calculations since this calculus is used as a measurement of the surface approximation quality. Naturally, this function has to be calculated for all visited surface configurations, and the number of Q calculations eventually corresponds to the number of iterations used throughout the thesis. According to the cost function definition, the processing load of cost function calculation is directly dependent on the corresponding calibration size as seen in the eq. (85):

$$Q = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_{1i} - \hat{x}_{1i})^2}$$

where N is the number of calibration samples.

Eventually, with N calibration samples, the Q function implementation consists of one square root calculation, one division, N additions, N subtractions, and N square calculations. If

the processing load of square root calculation is approximated with 2 square calculations, and additions and subtractions require the same processing power, Q function requirements are brought down to: $2N$ additions and $(N + 2)$ square calculation. Thus, the dependence is of linear nature, and if calibration size is reduced to half, the processing requirements on the receiver side can be relaxed by the same amount.

When analyzing the required calibration size the aim is to select the size which not only can offer excellent decoupling conditions but which can produce it constantly. This thought corresponds to the small calibration size which naturally can not retrieve the entire coupling information. In this case Q function gives us the indication of the surface precision, which is basically very low, since the search is conducted correctly but in this case the cost function value does not carry any information of the corresponding decoupling performance. Eventually, small cost function is not necessarily transformed into excellent decoupling behaviour hence the required calibration size will not be analyzed based on the produced cost function values.

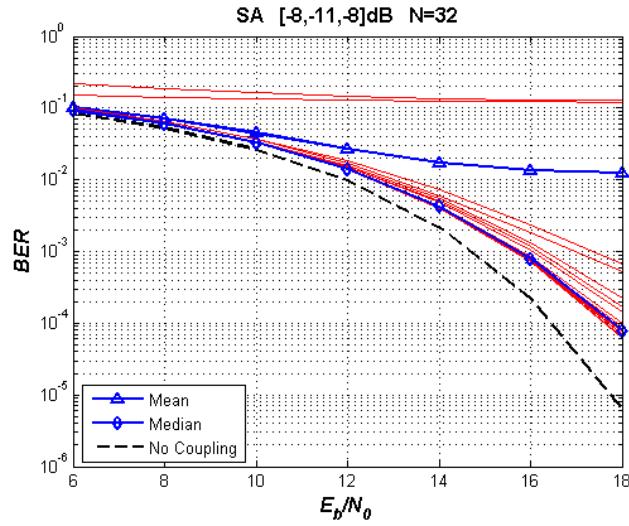


Figure 106: The performance of SA based decoupling with calibration size $N = 32$

In order to select the correct calibration size the attention should be centered on the system mean or median performance. This approach is depicted in Figure 106 where the two bold curves correspond to mean and median calculations, dotted line is the theoretical case without coupling and twenty simple lines correspond to the twenty independent decoupling module executions under the same conditions. The search is based on calibration of size $N = 32$, and SA approach implementing the set of search parameters established in previous chapter. Based

on the presented results, this calibration size often produces good decoupling conditions, but its implementation is not robust as in 10% of the cases the produced decoupling coefficients offer no performance enhancement. Still the correct measurement appears to be the one based on the median calculations as this curve is located inside the area with highest curve population. The mean curve in this case doesn't offer the correct information of one typical decoupling performance and will not be used for the decoupling quality measurement. However, it can serve as an indication of the system decoupling persistency, as when mean and median curves are close one to another the performance fluctuations are low and the overall system can be regarded as robust.

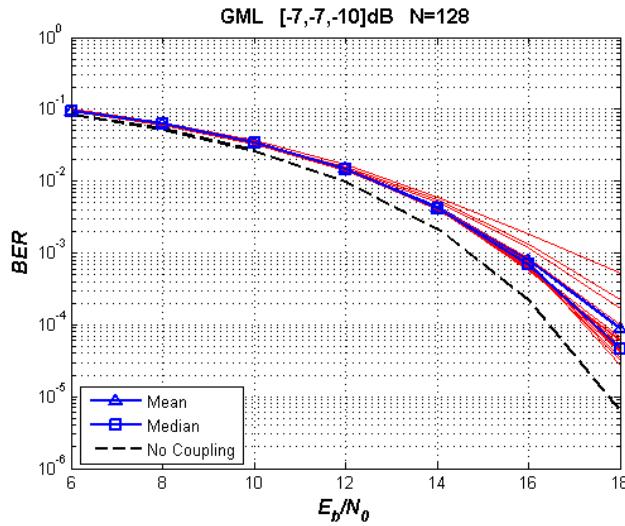


Figure 107: The performance of GML based decoupling with calibration size $N = 128$

If the calibration size takes higher value, the system robustness is increased as the curves corresponding to the independent executions under the same conditions are grouped closely together. This is seen in Figure 107, where the search based on the GML approach is applied on the coupling surface constructed out of $N = 128$ samples. In this case the mean curve does not enter the saturation phase at $E_b/N_0 = 14$ dB as is the case with $N = 32$, and it actually follows the median curve behaviour. Eventually, the median curve is located close to the bottom area of all decoupling executions and as such this curve will be used in the following simulations for the selection of the appropriate calibration size. The corresponding size will be set with respect to the processing requirements, surface precision and system robustness. The simulations are performed on 64QAM Gray signals, in AWGN channel, using all three decoupling module searches in order to analyze the behaviour of all approaches on the

calibration size. The presented BER curves correspond to median value of ten independent executions.

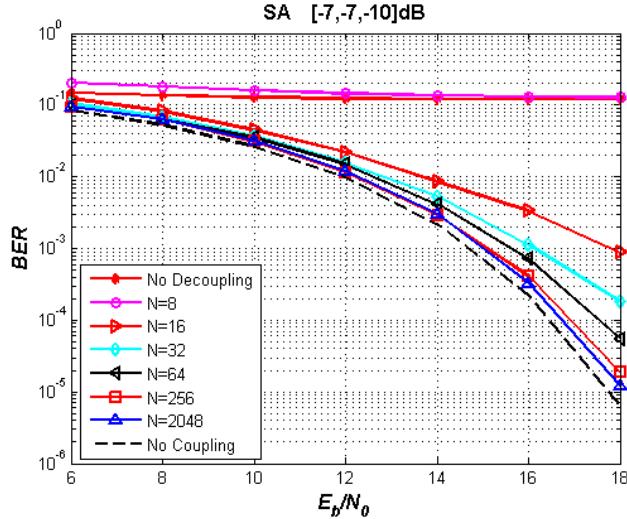


Figure 108: The performance of SA based decoupling with different calibration size implemented

The first set of simulations depicted in Figure 108 presents the results obtained with SA search approach under strong coupling of $[-7, -7, -10]$ dB. The SA search is implemented using search parameters adopted for decoupling problem in previous chapter. The systems without antenna coupling, with coupling, and with coupling and its different cancellations, are shown. As can be appreciated in the image, the reference surface based on only $N = 8$ samples doesn't offer any performance improvement. Moreover, the performance is even deteriorated at E_b/N_0 values lower than 14dB. However, with the introduction of additional surface reference points, the performance steadily improves. Naturally, the system performance closeness to the theoretical curve obtained without coupling is achieved with the highest number of calibration samples. In this case, the maximum number of calibration samples is set to $N = 2048$. The difference between this curve and the one obtained with 256 coupling samples is practically zero when E_b/N_0 is lower than 16dB. For larger E_b/N_0 values the performance deterioration can be detected but as the processing load of $N = 256$ is eight times lower than the one of $N = 2048$, the obtained precision justifies the precision lost. Eventually, based on the presented BER performance the calibration size required for the good coupling cancellation based on SA approach should be at least of $N = 256$ samples.

The behaviour of GML approach implemented on slightly lower coupling of $[-8, -11, -8]$ dB

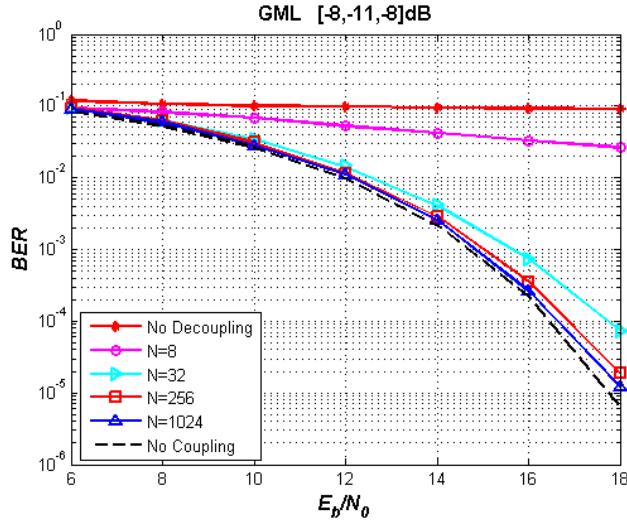


Figure 109: The performance of GML based decoupling with different calibration size

is shown in Figure 109. In this case the search based on $N = 8$ reference points does produce BER improvement, but the overall decoupling performance is too low even to be regarded as good decoupling. As expected two coupling size exhibiting excellent decoupling performance are again located above $N = 256$ samples. Eventhough the performance difference between $N = 256$ and $N = 1024$ is noticed, in this case, from $E_b/N_0 = 14$ dB the surface with 256 samples can be used inside the decoupling module.

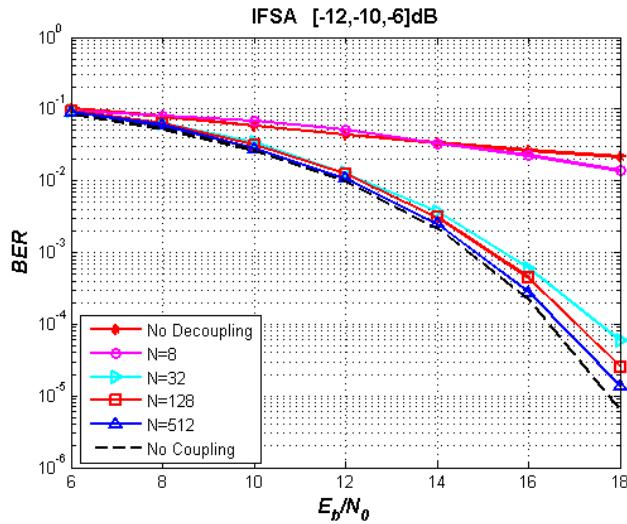


Figure 110: The performance of IFSA based decoupling with different calibration size

In the case of IFSA search in Figure 110, implemented under moderate coupling of [-12,-10,-6]dB system performance in dependence of calibration size follows the same logic as in previous two cases with the difference that in this case even with $N = 128$ excellent coupling cancellation can be achieved. This is logical since the coupling distortion is smaller and consequently the reference surface is smoother with less particularities. Eventually, based on this image, the IFSA approach can be used in moderate coupling conditions with the reference surface constructed out of $N = 128$ calibration samples.

The presented results confirm that SA, IFSA and GML decoupling achievements, depend on the calibration size which is in the accordance with the theoretical surface construction background. All three approaches present the same type of calibration size dependency as neither one of them exhibited significantly lower requirements or experienced untypical calibration size sensitivity. According to the presented results the tolerated calibration size for the moderate coupling starts with $N = 128$ samples, while this number is increased to $N = 256$ for the strong coupling. The complete coupling distortion cancellation is achieved with at least $N = 512$ calibration symbols in both cases. Since decoupling module used in practice is based on IFSA or GML approach, the calibration size can be selected without large pressure of processing load requirements. Taking into account that the generalized system, like ours, should be designed for the worst case scenario, the overall calibration size is set to $N = 512$ as in [Ara07c]. This calibration size is chosen in order to preserve system precision and robust behaviour demonstrated in previous simulations throughout the thesis based on 2000 calibration samples.

Conclusion

In this chapter the decoupling module consisting of three search algorithms is analyzed in noiseless calibration environment. This noiseless environment forms part of the proposed industrial calibration which, as already mentioned, can be implemented inside the reverberation chamber. The analysis included the behaviour of the decoupling module with respect to the number of the surface degrees of freedom and number of calibration samples used for reference surface construction.

It has been shown that simple reduction of decoupling coefficients is immediately perceived at the corresponding cost function level. This confirms that even though twenty levels of freedom are used for reference coupling surface construction that this number is not too high as

mathematically speaking each level of freedom contributed its share of coupling surface approximation. Since cost function levels are transferred into the system BER performance it has been demonstrated that by directly neglecting higher exponentials the decoupling achievements deteriorate. However, taking into account the LPF located in each branch of RF front-end based on I/Q downconverter the number of decoupling coefficients can be efficiently decreased. The idea is based on the fact that square coupling produces only high frequency components which are rejected in LPF and do not affect the received signal. Since LPF automatically cancels the square component there is no need to use the decoupling surface for the same task. Hence, it has been demonstrated that surface requirements can be relaxed, and the same precision can be achieved with twenty as with eleven decoupling coefficients. The decoupling surface in this case is constructed without even exponents. This simplified decoupling surface form is afterwards used for the calibration size analysis.

When adopting the search algorithms to decoupling problem the calibration size was set to 2000 samples, in order to guarantee the sufficient coupling information in the receiver. However, the most demanding part of the decoupling module, that is the cost function, directly depends on the number of calibration samples. Naturally, as the number of calibration points is decreased the processing load of the decoupling module also decreases. Hence the implemented calibration size in practical case should not be over-dimensioned if not necessary. The executed simulations for all three search approaches presented similar behaviour without any special calibration size requirements as all three of them exhibited the same calibration size sensitivity. It has been shown that as fewer points are used for describing the coupling, less information is available in the decoupling module, and as consequence decoupling exhibits performance degradation. It has also been demonstrated that the system robustness is enhanced as calibration size is increased. Since robustness is one of the most important properties of decoupling module, the calibration size is set to 512 samples favouring precision to processing load. This is done taking into account the low processing requirements of GML approach which is seen as the key search algorithm for the practical decoupling module implementation. Eventually, the enhancement of the received signal using the decoupling module depends on the level of coupling distortion and the precision requirements. If the precision requirements are high, the signals that suffered larger distortions need more calibration symbols to satisfy this requirement. Therefore the adequate calibration size should be established for each transmission system independently.

8 Calibration with noise

Up to this point the decoupling module has been analyzed and tested in noiseless conditions and the obtained decoupling coefficients are then used for system transmission in channel with AWGN noise. This approach allowed the development of three approximation surface algorithms in noise friendly environment with the emphasis on precision, convergence speed, robustness and processing load requirements. However, since the proposed decoupling module is physically located inside the mobile terminal, low noise conditions will rarely occur in practical use. Eventually the calibration process will have to be implemented in this environment also, and this case is presented in this chapter.

The presented calibration process corresponds to second and third calibration level according to the calibration types explained in chapter 5.2. In the second calibration level the proper terminal calibration path is used, while third level corresponds to real transmission system between mobile terminal and base station. The difference in the two calibration is in the noise power which is smaller at second level, and real time constraints which appear in third level calibration. The performed analysis is made in E_b/N_0 range from 6dB to 18dB which covers both calibration levels.

Eventhough each calibration level can be executed one after another using the solution of the previous level as a fitting surface starting point, in this thesis the analysis will be made only on independently generated solutions in order to locate the decoupling limits of each search approach. Due to this reason in the last section the three decoupling module search algorithms are compared with respect to the obtained precision, processing load, convergence speed and solution consistency. Each search algorithm is based on the corresponding search parameters adopted to the decoupling problem, while implemented surface form and calibration size are established in this chapter prior to comparison section.

8.1 Adaptation of surface form and calibration size

In previous chapter it has been shown that decoupling efficiency is directly dependent on the number of exponents used for the decoupling surface construction. Further on, it was demonstrated that if receiver's RF front-end is based on I/Q downconverter, its LPFs, located

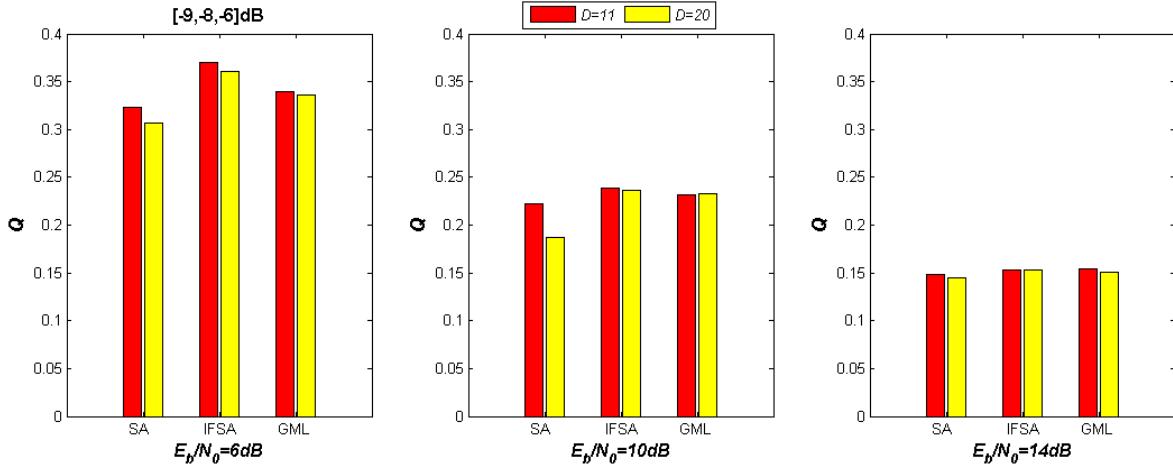


Figure 111: Behaviour of two decoupling surface forms under AWGN noise

in each signal branch, cancel the square coupling distortion. Hence decoupling surface which neglects even exponents can be used as a simplified surface form in noiseless environment preserving the decoupling precision. The same analysis is performed here for the calibration implemented in AWGN noise environment. The noise level implemented in the calibration stage corresponds to the noise level experienced at the transmission level where decoupling coefficients are finally employed. The simulations are conducted based on over-dimensioned calibration size, set to $N = 4000$ samples, in order to reduce its influence on the decoupling results. The presented results of the cost function correspond to average of ten independent search executions.

The obtained Q function implemented under strong coupling conditions of $[-9, -8, -6]\text{ dB}$, is shown in Figure 111 for three E_b/N_0 levels of 6 dB , 10 dB and 14 dB and for all three search methods. Based on the presented results two conclusions can be reached. First, the approximated surface precision deteriorates as noise level is increased, and second that the surface precision is only slightly decreased as fewer levels of freedom are used. In contrast to the noiseless conditions where Q function of surface with $D = 20$ coefficients exhibited up to ten times smaller Q values, when AWGN noise is added the increased number of coefficients can only achieve small precision enhancement. This is actually due to the noise distortions which prevent better approximation precision. Eventually, in this case, it is observed that surface approximation is more sensitive to noise increment than on the reduction of decoupling coefficients. This is seen in Q function drop from around $Q_{14\text{dB}} \approx 0.33$ to $Q_{6\text{dB}} \approx 0.15$ when E_b/N_0 is changed from 14dB to 6dB while by reducing levels of freedom to almost half, the cost

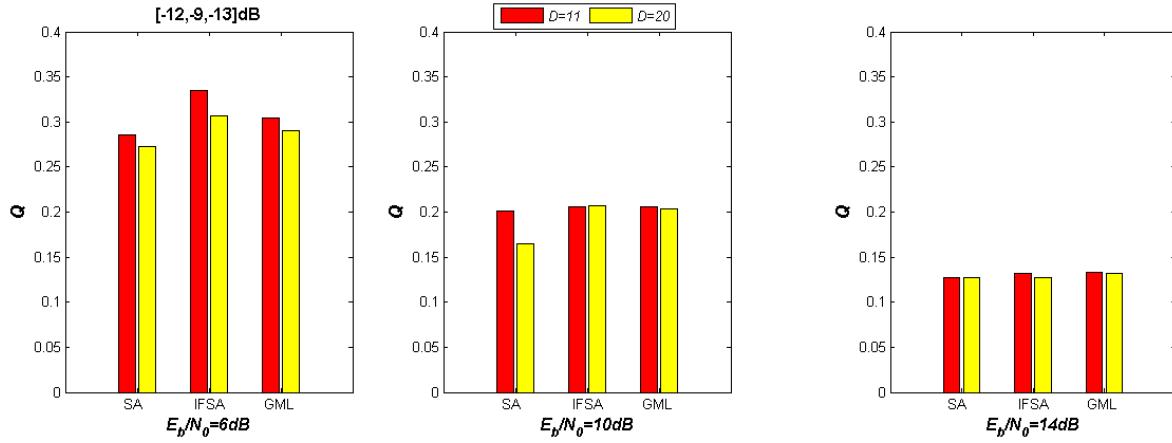


Figure 112: Behaviour of two decoupling surface forms under AWGN noise

function is decreased from $Q_{D20} = 0.22$ to $Q_{D11} = 0.18$ for SA approach under $E_b/N_0 = 10\text{dB}$.

The same conclusion can be reached when moderate coupling with AWGN noise distortion is analyzed. This case is depicted in Figure 112 for coupling strength of $[-12, -9, -13]\text{dB}$. The obtained precision of all three search approaches is again located in the same area for both decoupling surface forms with SA offering slightly better precision. Since SA approach defines the surface precision limit, and GML and IFSA obtain the solutions close to SA, this images indicate that these two approaches can be used with same efficiency under AWGN noise regardless of their reduced number of iterations. This indication is analyzed in the following images where the performance of the corresponding search approaches with two different surface forms is analyzed.

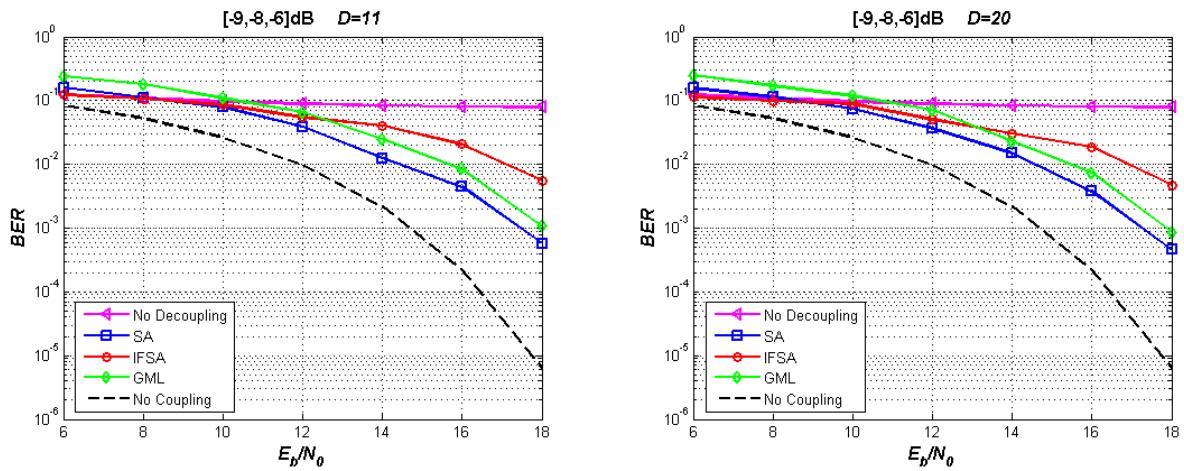


Figure 113: SA, IFSA and GML under $[-9, -8, -6]\text{dB}$ with based on different surface forms

In the Figure 113, decoupling efficiency of the coefficients obtained in previous search example, and implemented in decoupling module under corresponding E_b/N_0 noise conditions is presented. The curves correspond to the median value of ten simulations. The left image corresponds to surface with eleven, and the right image to the one constructed with twenty coefficients. If the two images are compared, decoupling results are very much alike, which backs surface simplifications based on the elimination of even decoupling coefficients. According to the presented results on both figures, the SA approach shows perceptible efficiency gain when compared to GML and even larger signal enhancement when compared to IFSA method. This is either due to the fact that the SA search is more robust to the AWGN noise or to the fact that small cost function differences are transformed into significant signal enhancement difference at transmission level.

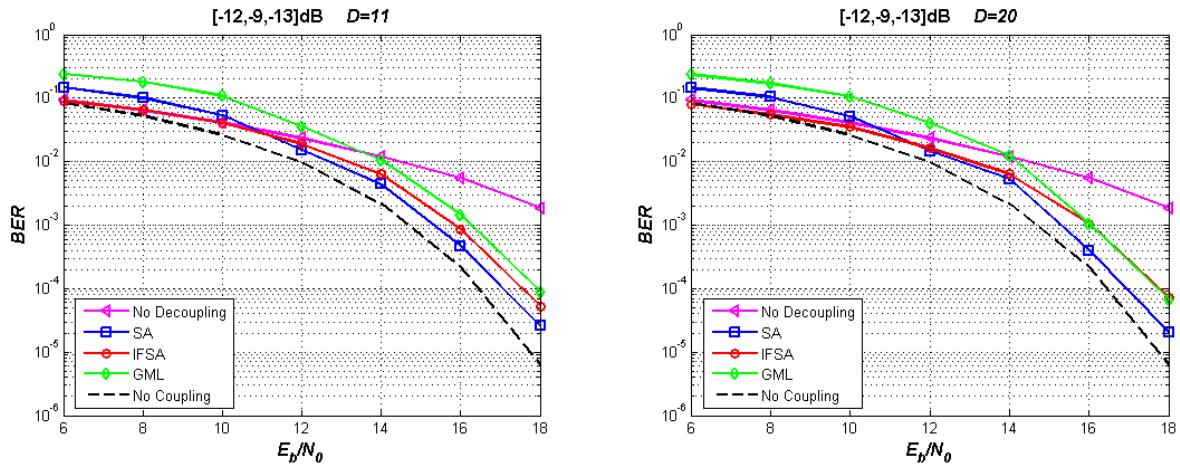


Figure 114: SA, IFSA and GML under $[-12, -9, -13]\text{dB}$ with based on different surface forms

If the decoupling module is implemented on moderate coupling the exhibited results lead towards the same conclusion. This is seen in Figure 114, where the influence of even decoupling coefficients is neglected on the left image. Similar to previous strong coupling case, regardless of this freedom reduction, and the resulting surface approximation degradation, the decoupling efficiency is preserved. Hence this kind of surface form, built with eleven degrees of freedom without even exponents, can be used in calibration process when receiver is based on RF Front-end with I/Q downconverter. This approach is highly recommended for practical decoupling situations since processing load is decreased considerably while decoupling results and robustness are kept almost unchanged.

All previous simulations with AWGN noise are made with reference coupling surface built

with 4000 samples. This system parameter is over-dimensioned with the aim of discarding the influence of calibration size on the decoupling efficiency. It was set to 4000 having in mind that coupling information stored in reference surface is altered under strong noise distortions. Thus, the size of reference surface larger than the one used in noiseless conditions is required to achieve consistent decoupling results. However, since processing load is paid expensively in real system the over-dimensioned parameters are generally avoided unless they are highly necessary. In order to determine the acceptable calibration size factor the decoupling efficiency should be evaluated in terms of precision, processing requirements and robustness as was the case for noiseless calibration.

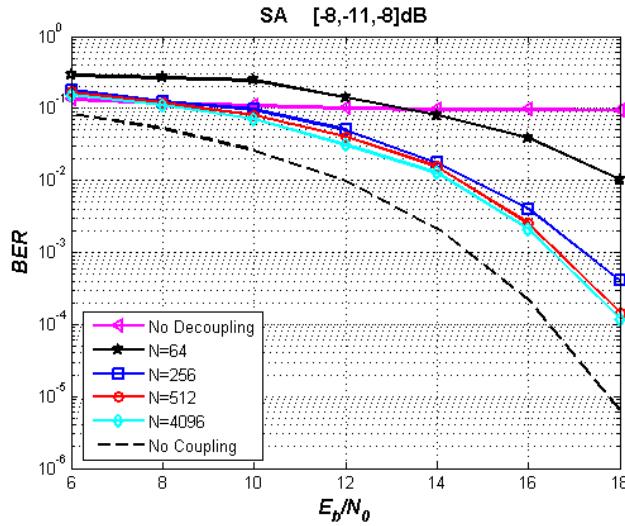


Figure 115: SA approach with different calibration size under AWGN

The set of simulations depicted in Figure 115 shows the results obtained with SA algorithm implemented on decoupling surface, with eleven degrees of freedom, under strong coupling of $[-8, -11, -8]$ dB. The image consists of two characteristic curves corresponding to the theoretical system limit without coupling marked with dotted line, and the flat one corresponding to the system with coupling but without implemented decoupling module. All other curves represent the achievements of the particular calibration size correspond to the median of ten independent simulation executions. As expected with the increase of calibration size factor, the decoupling module is more effective and the performance curve is located closer to the theoretical limit case without coupling. However, even though the best precision is achieved with $N = 4096$, the results obtained with $N = 512$, which is eight times smaller, are very similar. Hence the increase in processing load is not justified. On the other hand, if calibra-

tion size is lowered to $N = 256$, the significant performance difference is noticed only when smaller noise levels are present. Under $N = 64$, which is the smallest calibration size presented in the figure, the decoupling module produces signal enhancement only at levels higher than $E_b/N_0 = 14\text{dB}$, as stronger AWGN noise destroys the received signal and misleads the search. As a consequence the decoupling surface is constructed on faulty bases and the introduction of decoupling module lead to signal deterioration in this case. This is experienced for $N = 64$ at E_b/N_0 smaller than 13dB , and at $E_b/N_0 < 8\text{dB}$ for all other calibration sizes. According to the presented results, SA search can not offer excellent decoupling results under strong coupling, regardless of the calibration size factor. However, acceptable decoupling is obtained with calibration size above 512 samples.

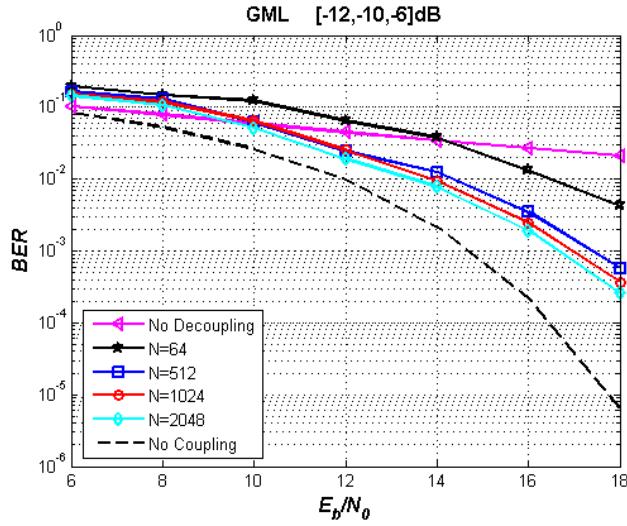


Figure 116: GML approach with different calibration size under AWGN

The second set of simulations presented in Figure 116, corresponds to the GML approach implemented with different calibration size factors under moderate coupling conditions of $[-12, -10, -6]\text{dB}$. As it is perceived in the image, much like in previous case, the reference surface based on only $N = 64$ samples experience poor decoupling performance and even deteriorate the received signal under $E_b/N_0 < 14\text{dB}$. Nevertheless, in this case the performance obtained with $N = 512$ is not so close to the best performance as in the previous case. This indicates that calibration size of $N = 512$, can not be regarded as robust. This conclusion stands, since the performance curve of $N = 512$ is not always located close to the best calibration curve. This performance difference can be appreciated in the last two figures, especially at $E_b/N_0 > 14\text{dB}$. Apparently, in this case, the adequate size of calibration samples

is achieved with $N = 1024$.

Conclusion

The presented simulation indicate that the introduction of AWGN noise into the calibration process influences the decoupling results considerably. Apparently, the presence of noise restrains the decoupling performance as only a part of the coupling information, carried inside calibration reference surface, reaches correctly the receiver. This implies that reference surface carries less coupling information as the noise is amplified. For this reason the calibration size of $N = 512$, recommended for noiseless scenario is not sufficient for AWGN noise environment. In contrast, the surface based on eleven degrees of freedom without even exponents, experienced the same behaviour in noiseless conditions as under AWGN noise. The obtained decoupling results presented the same performance levels as the ones obtained with twenty degrees of freedom, which ratifies the adequate surface simplification. Hence, this surface model can be implemented in decoupling module under AWGN noise conditions. Eventually, according to the shown decoupling results the calibration size factor should be increased to at least $N = 1024$ samples as this size offers excellent performance for moderate coupling in AWGN environment.

8.2 Algorithm Comparison

The algorithm comparison in under AWGN noise is done for strong and moderate coupling conditions. The AWGN noise in calibration path corresponds to the noise that experiences a signal in transmission path. The calibration size consists of 1024 samples, and the decoupling surface is based on simplified surface constructed with eleven degrees of freedom. The results presented in the figures correspond to the median of ten independent executions.

The convergence speed, obtained under strong coupling of $[-8, -8, -5]$ dB, of all three search methods is shown in Figure 117 for the first 1200 iterations. The figure on the left side corresponds to environment with strong noise, with $E_b/N_0 = 6$ dB, while the one on the right corresponds to low noise environment of $E_b/N_0 = 14$ dB. The difference in noise conditions can also be appreciated in the initial value of cost function which is lowered from $Q_0 = 0.43$ to

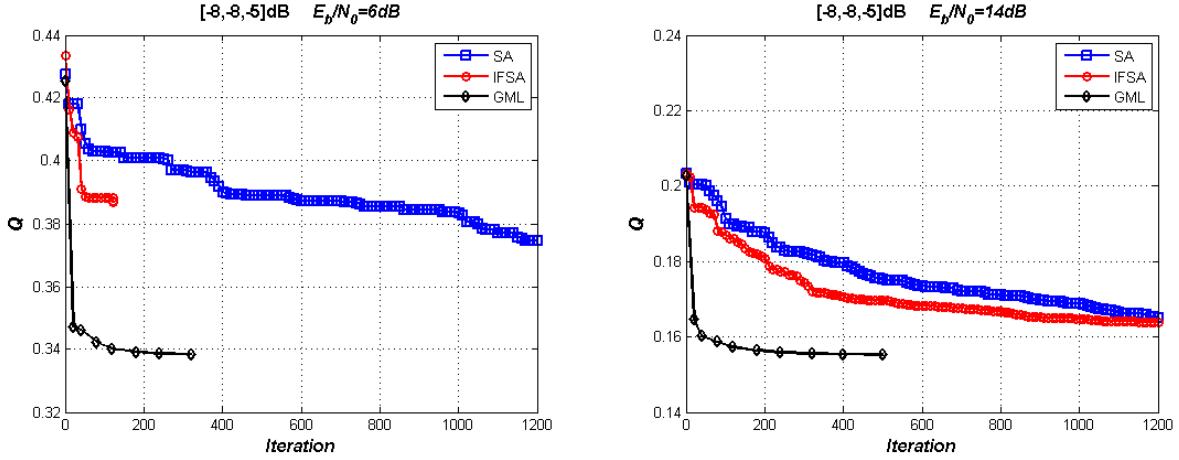


Figure 117: The convergence speed under strong coupling and AWGN noise

$Q_0 = 0.21$ when E_b/N_0 is changed from 6dB to 14dB. The obtained cost function results show that the GML approach exhibits fastest convergence regardless of the level of noise that is present in the system, while, as expected, SA method experiences slowest convergence. Since the search progress is depicted for only first 1200 iterations, the final precision of SA approach can not be evaluated. However, the final approximation precision is shown in Figure 118, where the obtained cost function is presented for E_b/N_0 of 6, 10 and 14dB. In this case SA method offers highest precision with respect to the cost function regardless of the AWGN noise level. It is followed by GML approach while IFSA shows lowest precision. The figure

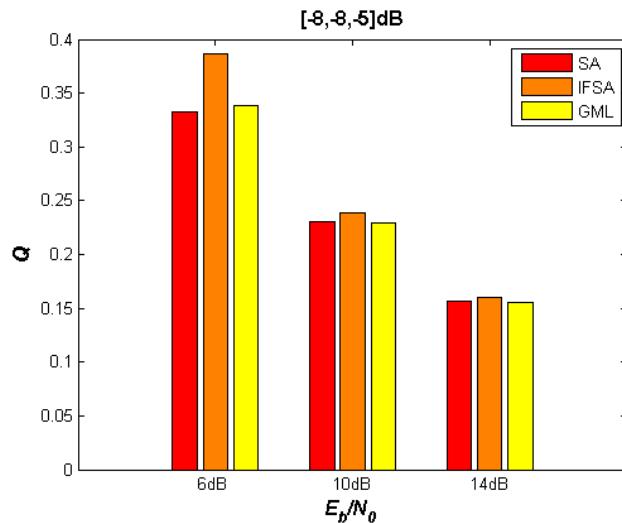


Figure 118: The obtained cost function under strong coupling and AWGN noise

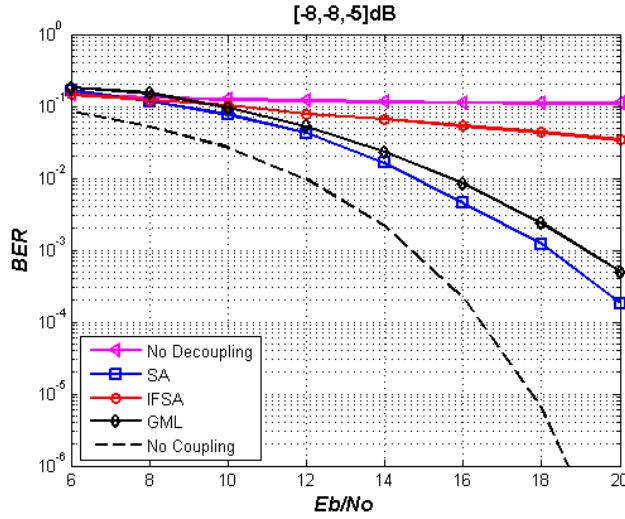


Figure 119: The performance of search methods under strong coupling and AWGN noise

also shows decreasing tendency of cost function value as the AWGN noise is reduced. This confirms the correct functioning of all three approaches under AWGN noise environment.

The BER performance of the the decoupling module in strong coupling conditions and under AWGN noise is seen in Figure 119. The best performance corresponds to the SA approach which is closely followed by GML method. The IFSA approach shows poor decoupling performance as its decoupling curve experience slow signal enhancement with noise reduction. Nevertheless, for E_b/N_0 lower than 10dB all three approaches show similar performance. In this area signal improvement is low and in case of E_b/N_0 lower than 8dB the introduction of decoupling module even deteriorates the received signal. This is due to the combination of strong coupling and strong AWGN noise which make impossible correct reference surface construction. As a result the decoupling surface is of poor quality and is constructed on erroneous bases so it eventually destroys the received signal instead of enhancing it.

In case of moderate coupling the behaviour of the decoupling methods is similar to the one experienced under the strong coupling. The convergence speed of all three approaches at $[-12, -13, -10]$ dB coupling and AWGN noise is shown in Figure 120. The number of iteration is reduced to 600 in order to clearly seen the search progress of GML and IFSA approach. As expected, the GML approach continues to show the fastest convergence while the SA approach is still the slowest one.

However, when the final precision is analyzed, the IFSA approach experienced lowest Q value

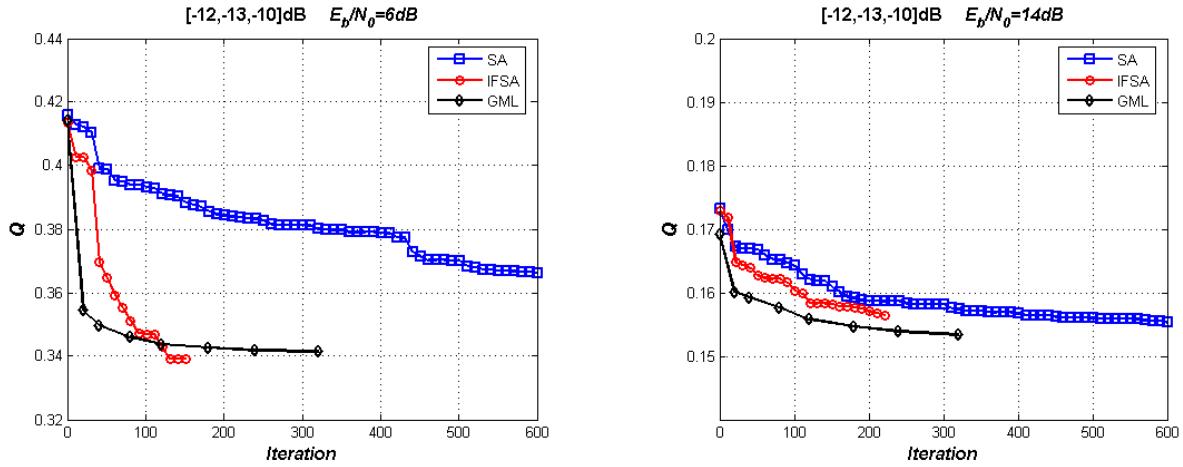


Figure 120: The convergence speed under moderate coupling and AWGN noise

for E_b/N_0 lower than 10dB. This is seen in Figure 121, where the obtained cost function is presented for E_b/N_0 equal to 6, 10 and 14dB. Apparently, in this coupling case the IFSA method exhibits the best precision under stronger AWGN noise levels like $E_b/N_0 = 6$ or 10dB and it is only at higher E_b/N_0 values when SA reclaims its high precision property. This is explained with wide initial search area which permits the IFSA approach to avoid the local minimums which correspond to AWGN noise, while SA loses too much time in uphill movement and can not offer the adequate solution with available number of iterations.

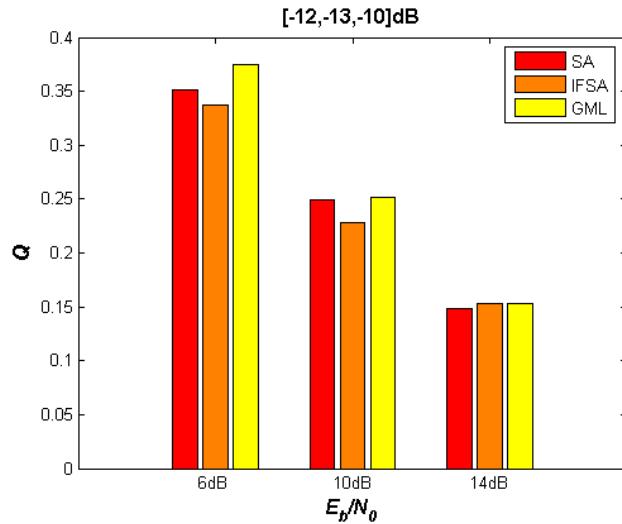


Figure 121: The obtained cost function under moderate coupling and AWGN noise

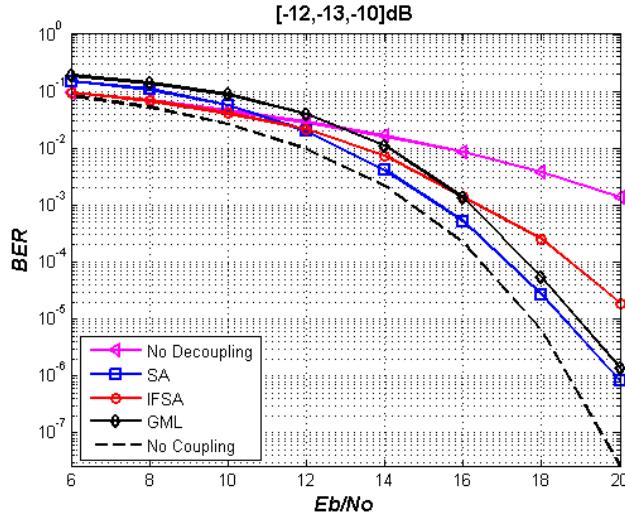


Figure 122: The performance of search methods under moderate coupling and AWGN noise

The precision of IFSA approach under strong AWGN noise is also confirmed in Figure 122, where the BER performances of three decoupling methods are presented. The IFSA method shows the best performance for E_b/N_0 lower than 12dB, but the performance curve doesn't follow the theoretical curve and apart from this level the SA decoupling shows the best signal enhancement. Since IFSA method again shows slow decoupling enhancement progress, it is not recommended for E_b/N_0 larger than 16dB. This E_b/N_0 level corresponds to the case when GML approach starts to show better performance. Hence the practical implementation would include IFSA approach at high AWGN noise levels and GML at lower noise levels.

Conclusion

When AWGN noise is introduced in the calibration path the reference surface consisting of calibration samples is altered and the coupling information is reduced. As AWGN noise level is increased, more coupling information is lost, and, as a consequence, decoupling surface approximation can not achieve excellent decoupling. If strong AWGN noise is combined with strong coupling case the environment is so hostile that eventually decoupling even deteriorates more the received signal. This is because the search is based on erroneous surface information. However, the analysis showed that significant signal enhancement can be achieved in practical case under moderate coupling when IFSA is implemented at strong noise and GML at low noise levels. The required number of calibration samples for this environment is set to 1024

samples. Also, it has been demonstrated that the simplified surface based on eleven coefficients can be used with almost the same efficiency as the one constructed with twenty.

9 Conclusions

When analyzing the multiple antenna RF Front-end distortions the coupling phenomena is usually not considered. However, as most of the modern communication systems incorporate multiple antennas transceivers inside a small space, the RF coupling is highly probable, and its investigation is required. In this thesis, the impact of RF front-end nonlinear coupling on the system performance is analyzed and decoupling module, located after A/D converter, is proposed for coupling reduction. The following sections include the most significant contributions indicated chapter by chapter.

Chapter 1

The introductory chapter is dedicated to the review of the mobile terminal background and eventual coupling friendly environment evolving inside the present product tendencies. The detected tendencies include product miniaturization, increased processing power and the introduction of multiple antennas and many other hardware components not related directly to communication purposes. At the same time product low costs requirement are permanently present which means that the inside of mobile terminal is becoming over-populated with low-cost electronics component. Since low-cost components tend to dissipate energy, a rather hostile environment is easily created. Further on, digital components are interlaced with a lot of data metallic wires and different analog elements, which create magnetic fields, increase operating temperature and finally favour the appearance of different signal distortions. Eventually, this kind of environment stimulates the appearance of coupling as these coupling distortions occur mainly inside the places where no energy dissipation protection is present (due to the lack of space), or where low-cost electronic elements are used in order to reduce the product costs. However, not all mobile phone tendencies are negative to signal processing. Actually, taking advantage of the increased mobile phones processing power is proposed in the first chapter of the thesis, for achieving software coupling cancellation. Even though coupling appears both in transmission and reception paths in mobile terminal, this thesis focuses only on the cancellation at the reception. Actually, coupling appears in all MIMO transmission, but this particular case is chosen since it is the most demanding from the point of view of processing power and battery restrictions. When coupling is present at transmission in mobile terminal, more sophisticated and computationally complex algorithms can be used at base station for its cancellation. Thus, the objective of this thesis is to develop the cheap, yet effective

way to overcome, or at least reduce, the coupling effects using the resources available at the mobile phone.

Chapter 2

In the second chapter the RF front-end architecture and various distortions appearing in single antenna architecture are analyzed. At present, the employed RF front-end is still mainly based on low-IF or zero-IF architecture with analog signal reception, digitalization at IF stage and signal processing based on DSP technology. However, as DSP processors are becoming faster, many of the radio system functions are implemented in software and A/D conversion is moving closer to the antenna, leading toward the software defined radio system (SWR). Eventually, it is shown that the future ideal SWR terminal still includes analog elements like amplifiers or local oscillator which can not be digitally replaced. This indicates that the same kind of signal distortion will also appear in future RF Front end technologies and that the results obtained in the thesis can be applied with the same efficiency. The analyzed distortions include DC offset, flicker noise, second order intermodulation and I/Q imbalance imperfections. All of them appear to present the nonlinear behaviour, which the majority of authors tries to bypass in order to present the system as simple as possible. It is interesting to notice that all presented distortions can be resolved with both, analog and digital approach. Still, the cancellation methods are converging towards digital signal processing as processors are getting more powerfull. These distortions are important since this kind of signal imperfections experienced at one antenna can eventually occur as part of coupling distortion and can be observed on the other antenna. Inside RF front-end with multiple antennas all these imperfections are intensified as the environment gets more hostile.

Chapter 3

When building a system model, the emphasis is usually made on simplification and reducing the behavior to linear dependency. However, it has been shown in many scientific areas that nonlinear foundations adapt better to real life situations and offer higher accuracy solutions. The nonlinear modelling is avoided largely due to complex inverse functions which requires advanced mathematical methods for obtaining the solution. Nevertheless, since the majority of the imperfections presented in Chapter 2 exhibit strong nonlinear behaviour the accurate

coupling model can not be based on linear bases. Thus, nonlinear coupling model consisting of polynomials limited to third order is proposed in the thesis. Since coupling actually corresponds to energy swap between two physical media, the idea was to describe it like a simple *black box* with its input and output values. This approach makes the model independent of the actual coupling source. The developed model is able to adapt easily to different distortion scenarios, being able to support the appearance of various distortions at the same time together with the nonlinear coupling phenomena. Since it is not limited to multi-antenna systems, the model is of general importance and the obtained decoupling methods can be implemented in other scientific areas which also face coupling distortion. Chapter 3 also proposes the means for relating the coupling occurring at physical level with the corresponding numerical case in mathematical model. Eventually, the coupling model is defined with a set of six coupling parameters expressed in [dB] describing the direct coupling strength of linear, square and cube distortions at two receivers antennas. The general coupling model is than adopted to the RF front-end with I/Q downconverter. Afterwards, the influence of linear, square and cube coupling on the output of I/Q downconverter is analyzed. It is demonstrated that the square coupling generates only high frequency components, which are later rejected in LPF located in each signal branch of I/Q receiver. Thus, eventhough square coupling appears inside RF front-end, its influence at the output is zero. However, I/Q architecture is not immune to the existence of linear and cube coupling whose cancellation require additional signal processing. This is seen in the last section of fourth chapter where the implemented simulation system is presented. The simulations presented show the necessity of coupling cancellation module, as otherwise system performance is easily drawn to saturation.

Chapter 4

In Chapter 4, the mathematical analysis of inverse coupling function is performed. It is demonstrated that the analytical solution is very complex, and that the generic form of unique solution can not be derived as its form changes depending on the coupling and the set of input signals. Furthermore, it is shown that the inverse solution to coupling problem might not exist at all, or it may offer more than one possible solution which is unacceptable. As a result Sturm functions are introduced for unique solution confirmation and afterwards Newton method is applied for the root search. However, as this approach requires important amount of processing power, surface fitting approximation in combination with calibration process is introduced as the inverse coupling solution. The shortcoming of the approximate methods, commonly

termed heuristics, is that no theoretical guarantee for finding the optimal, or even feasible solutions, is provided. On the other hand the solution is not ambiguous and once the inverse surface is well defined only a small portion of the algorithm has to be executed to obtain the desired solution. Thus, the pursuit for analytical solution is avoided and the attention centers in obtaining inverse coupling function with mathematical surface approximations. From the technological point of view the calibration can be implemented at three physical levels which back up the this decoupling approach. Eventually, with the introduction of surface approximation the complexity of nonlinear inverse coupling function is transformed into highly demanding search for the adequate decoupling surface form. The search is carried on in the decoupling module which is located between the outputs of A/D converter and any further digital signal processing. This location is chosen since it is least dependent on the transmission signal type as surface data consists directly on samples taken at the sampling rate of A/D converter. Inside the module, signal enhancement is achieved based on approximation surface constructed as a sum of two polynomials of up to ninth degree. This surface corresponds to the analytical inverse function which is also of the ninth degree. Since this kind of functions tend to be "lumpy" with several local minimums and maximums the approximation surface should be able to correctly adopt to any coupling case and hence higher order polynomials appear.

Chapter 5

The fifth chapter develops three approximation surface search algorithms and adopts them to generic coupling cancellation case. The algorithms include Simulated Annealing (SA), Improved Fast Simulated Annealing (IFSA) and Guided Multi-Level (GML) search. The SA method is proposed since it is widely used in applications dealing with many levels of freedom and enormous number of possible solutions. The name and inspiration come from the annealing in metallurgy, as this technique mimics the nature of the metal which is heated and then cooled slowly in order to reach its absolute minimum energy state. Basically, the SA approach is frequently used since it can statistically guarantee finding an optimal solution, but as no estimate on the required number of iterations is available, it is also well known for its slow convergence. This means that the search can get unacceptably exhaustive and time consuming which calls for a trade off between the guaranteed precision and speed. For this reason the original SA code based on physical annealing process suffered minor changes as convergence speed is improved. However, since the precision range is almost kept unchanged, the algorithm enhancements could not be radical, and consequently the required processing time is far beyond

the real time constraints. Nevertheless, when developing the SA algorithm the sight was never set on speed, but principally on the approximation precision as this approach eventually provides the lower limit of the decoupling precision module. Furthermore, it is also used for selecting the adequate type of cost function, choosing the adequate starting search point, and, what is most important, for testing the surface model selection through system BER performance.

The second decoupling module algorithm, denoted IFSA, is introduced in order to significantly speed up the search progress. It is developed by the author specifically for nonlinear decoupling problem and is intended to be used in practical implementations. Inside IFSA, the heuristic methods are preserved, and the search progress is accelerated on the cost of the approximation surface precision. The corresponding processing load is decreased to less than 10%, when compared to the first SA approach, but this is still too demanding to be used under real time constraints of the present mobile terminal devices. Another drawback of IFSA is that it also depends on the probability factor, and since processing load is eventually restrained to some limit, this makes impossible the precision guarantee. Hence, if the obtained solution is below the certain precision limit, the search has to be executed again.

Finally, in order to avoid the probability factor and its consequent indecision, a deterministic solution in the form of a simple iterative algorithm denominated GML search is developed by the author. This approach enables the reduction of average processing load required for one cost function calculation, it simplifies the search process and consequently increases the convergence speed. It is adopted for the presented decoupling problem but in fact it can be used in any approximation problem dealing with many degrees of freedom. The advantages of the proposed GML search include easy control of speed and precision, reduced processing requirements, simple implementation and avoidance of statistical dependence of the solution. The basic idea is to construct the decoupling surface on solid bases starting with only two, and then gradually introducing more levels of freedom. Eventually, the approach ensures deterministic solution to the decoupling problem with excellent surface precision, very fast convergence speed and low processing demands.

Throughout this chapter all three proposed coupling cancellation algorithms are analyzed and adopted to nonlinear coupling cancellation problems. The analysis is done in noiseless calibration scenario with large calibration size in order to reduce their influence on the decoupling results. This scenario in fact corresponds to the industrial calibration stage (reverberation chamber). The obtained results confirm the SA approach as the most precise under all cou-

pling scenarios. The SA approach is followed by GML whose precision is very similar, while IFSA exhibits notable precision deterioration. When comparing the convergence speed GML and IFSA require significantly smaller number of iterations, but the corresponding processing requirement are considerably lowered in GML case. Actually, GML algorithm requires 10% of the IFSA processing requirement, while IFSA corresponds to 10% of processing power of the original SA approach. All three methods are adopted to the decoupling problem taking into account the robustness factor, hence all three experience solid robustness in noiseless calibration conditions.

Chapter 6

In chapter six the noiseless analysis is performed in more depth centering on further system optimization. First, the simplification of surface form based on the reduction of higher exponents is analyzed. The obtained results showed that the decoupling performance is eventually affected with this simplification and the grade of deterioration depend on the specific coupling case. Eventually, in general good decoupling results can be achieved by limiting the surface polynomials to fifth order, that is, limiting the number of degrees of freedom to twelve. However, if the search algorithms take advantage of the specific behaviour of RF front end based on I/Q downconverter the number of degrees of freedom can be reduced almost without performance loss. This is confirmed for all three approaches. The reduction is based on the fact that square coupling produces only high frequency components which are rejected in two LPF of I/Q downconverter. Hence, this kind of coupling does not affect the received signal and the approximated surface requirements can be relaxed as square cancellation can be neglected. The result is surface constructed with eleven decoupling coefficients without square and any higher even exponent. Afterwards, the search based on this reduced surface is used for determining the adequate number of calibration samples required for reference surface construction. The focus of this analysis is set on the system robustness which gets stronger as the number of calibration samples is increased. All three search algorithms are tested, and all of them presented the same level of sensitivity regarding this system parameter. Based on the obtained performance results the adequate calibration size of 512 samples is proposed.

Chapter 7

The last chapter describes the decoupling module optimization when the calibration is done under AWGN noise conditions. This case corresponds to the calibration implemented inside the mobile terminal or between the terminal and base station. In contrast to the noiseless case where reference surface and decoupling coefficients are constructed in perfect conditions, here the reference surface is constructed under the conditions equal to the ones of the real received signal. That is, if the transmission system is experiencing the AWGN noise resulting in E_b/N_0 value of 10dB, the implemented decoupling coefficients are obtained under the same level and not under, for example, E_b/N_0 of 100dB. Again, all three search methods were tested with simplified surface form, and confirmed its correct functioning. The analysis of required calibration size, implemented on the surface with eleven degrees of freedom was carried out, and, according to the obtained results, the proposed number of calibration samples is 1024. Afterwards, the comparison of the three search methods under AWGN noise is conducted. The presented results showed the same behaviour with respect to required number of iterations and processing load as experienced in noiseless conditions. However, BER performance has suffered some changes as the best results do not correspond to SA approach over the entire E_b/N_0 range. In fact, IFSA method has proven to be the only one able to produce signal enhancement under strong noise conditions, but the overall performance can not be described as excellent since it exhibited poor decoupling performance for E_b/N_0 lower than 10dB. Since GML algorithm follows the behaviour of SA approach with small difference, but under low processing demands, the coupling under AWGN noise eventually can be combated with combination of IFSA and GML approach.

Future research

The main contributions of this thesis are nonlinear analysis of RF front end coupling, and the development of simple GML method for the approximation surface search. Consequently, possible future research can be routed in two separate investigation courses. The first direction would be the further development of GML approach and the analysis of its efficiency in other coupling-type problems like audio coupling in 3.1 systems, or coupling appearing in vdd or gnd networks. It would be interesting to see the behaviour of GML algorithm in general complex problem environment when many degrees of freedom are present but limited processing power is available.

The second investigation line would follow the course of RF front end coupling analysis. The coupling model developed in this thesis, assumes coupling to be time invariant, and frequency independent. However, these assumptions are done only in order to simplify the overall system and future research can include the investigation of their influence on coupling environment. Future work can also include the introduction of more than two antennas in the system and the consecutive complex coupling environment. Nevertheless, the most important research should be focused on measurements obtained in real life systems. Since this thesis is based on theory and it would be interesting to adopt these measurements to proposed coupling model and to compare the obtained theoretical results with practical decoupling efficiency.

Summary of scientific contributions

Some of the problems encountered in RF front end nonlinear coupling cancellation analysis are presented in the following publications [Ara07b, Ara07a, Ara07c, Rao06b, Ara02a, Ara02b]. While the former three contributions are directly related to the problems addressed in this thesis, the last three are more general, focused on analysis and design of wireless systems, that helped the author to gain the necessary background and experience. A number of research projects in which the author participated made possible the formation in area of multiple antennas systems. Some of the projects, those related to multiple antenna systems, are listed below:

- MATRICE (IST-2001-32620) Multicarrier CDMA Transmission Techniques for Integrated Broadband Cellular Systems; funded by 5th Framework European Commission IST Programme; duration, from: January 2002 to January 2005
- 4MORE (IST-2002-07039) MC-CDMA Multiple Antenna System-on-Chip for Radio Enhancements; funded by: 6th Framework European Commission IST Programme; duration, from January 2004 to June 2006
- Wireless Hybrid Enhanced Mobile Radio Estimators (WHERE); funded by 7th Framework European Commission IST Programme; duration, from January 2008 to June 2010
- Arquitecturas para radiocomunicaciones digitales de banda ancha; funded by: Spanish Ministry of Science and Technology; duration from 2001 to 2003.

- Transceptores integrados multielemento y monocircuito para radiocomunicación digital de banda ancha; funded by: Spanish Ministry of Science and Technology; duration from January 2004 to December 2006
- Transceptores integrados multielemento y monocircuito para radiocomunicación multimedia de banda ancha y muy ancha; funded by Spanish Ministry of Education and Culture; duration from January 2007 to December 2009

References

- [Aar97] E. Aarts and J. Lenstra. *Local Search in Combinatorial Optimization*. Wiley, 1997.
- [Abe01] M. Abe, N. Sasho, R. Morelos-Zaragoza and S. Haruyama. An RF Architecture for Software Defined Recivers. *Microwave Workshop & Exhibition 2002, Yokohama*, Dec. 2001.
- [Abi95] A. A. Abidi. Direct Conversion radio transcievrs for digital communications. *IEEE Journal of Solid-State Circuits*, vol. 30:pp. 1399–1410, Dec. 1995.
- [Abi96] A. A. Abidi. *Low-power radio-frequency ICs for portable communications*. Norwood, MA: Artech House, 1996.
- [Aga05] J. Agar. *Constant Touch: A Global History of the Mobile Phone*. New York: Totem, 2005.
- [Agn06] F. Agnelli and et al. Wireless Multi-Standard Terminals: System Analysis and Design of a Reconfigurable RF Front-end. *IEEE Circuits and Systems Magazine*, pp. 38–59, 2006.
- [Ara02a] I. Arambasic, D. Begusic and J. C. Quiros. Analisis de Interconexión de Redes GSM/GPRS y WLAN. *XVII Simposium Nacional de la Unión Cientifica Internacional de Radio, URSI 2002*, pp. 77–78, Sept. 2002.
- [Ara02b] I. Arambasic, I. Raos, F. J. C. Quiros and D. Begusic. GSM/GPRS-WLAN Interworking Solutions. *10th International Conference on Software, Telecommunications & Computer Networks, SOFTCOM2002*, pp. 531–535, Oct. 2002.
- [Ara07a] I. Arambasic, F. J. C. Quiros and I. Raos. Efficient RF Front-End Non-Linear Multi Antenna Coupling Cancellation Techniques. *PIMRC 2007. IEEE 18th International Symposium on Personal, Indoor and Mobile Radio Communications, 2007*, pp. 1 – 5, Sept. 2007.
- [Ara07b] I. Arambasic, F. J. C. Quiros and I. Raos. Improvement of multiple antennas diversity systems through receiver nonlinear coupling cancellation. *2007. International Waveform Diversity and Design Conference*, pp. 102 – 106, June 2007.
- [Ara07c] I. Arambasic, F. J. C. Quiros and I. Raos. Receiver’s Multi-Antenna Coupling Cancellation Based on Guided Multi-level Search. *WISP 2007. IEEE International Symposium on Intelligent Signal Processing, 2007*, pp. 1 – 6, Oct. 2007.

- [Bar04] C. Barquinero, T. Fernandéz, J. Casajús, I. Raos, J. Rossi, A. Arkhipov and R. Raulefs. RF front end design of a MIMO transceiver. Tech. rep., ACORDE, UPM, FT, DLR, July 2004.
- [Bau00] E. E. Bautista, B. Bastani and J. Heck. A high IIP2 downconversion mixer using dynamic matching. *IEEE Journal of Solid-State Circuits*, vol. 35(12):pp. 1934 – 1941, Dec. 2000.
- [Ber97] J. E. Berruto and G. Colombo. Protocol and Architectural issues for Software Radio. *SW Radio Workshop, Brussels, Belgium*, May 1997.
- [Bos99] V. Bose, M. Ismert, M. Welborn and J. Guttag. Virtual radios. *IEEE Journal on Selected Areas of Communications*, vol. 17:pp. 591–602, April 1999.
- [Bud07] P. Budde. 2007 Global Mobile Communications - Statistics, Trends & Forecasts. Tech. rep., 2007.
- [Bur00] E. Buracchini. The Software Radio Cocept. *IEEE Communications Magazine*, vol. 38:pp. 138–143, Sept. 2000.
- [Cav93] J. K. Cavers and M. W. Liao. Adaptive compensation for imbalance and offset losses in direct conversion transceivers. *IEEE Transactions on Vehicular Technology*, vol. 42:pp. 581 – 588, Nov. 1993.
- [CH00] C. Chih-Hung, M. J. Deen and M. M. C. Yuhua. Extraction of the channel thermal noise in MOSFETs. *Proceedings of the 2000 International Conference on Microelectronic Test Structures, 2000. ICMTS 2000.*, pp. 42 – 47, 2000.
- [Cha94] J. Chang, A. A. Abidi and C. R. Viswanathan. Flicker noise in CMOS transistors from subthreshold to strong inversion at various temperatures. *IEEE Transactions on Electron Devices*, vol. 41(11):pp. 1965 – 1971, Nov 1994.
- [Cha02] J. Chapin and A. Shah. Software radio technology and challenges. *Proceedings of the Thirty-Eighth Annual International*, October 2002.
- [Che82] V. Cherny. A thermodynamical approach to the travelling salesman problem: An efficient simulation algorithm. Report, Comenius University, Bratislava, Czechoslovakia, 1982.

- [Che96] J. Y. C. Cheah. Introduction to wireless communications applications and circuit design. *RF and Microwave Circuit Design for Wireless Communications*, pp. 17–41, 1996.
- [Chi01] D. Chien. *Digital Radio System on a Chip - A System Approach*. Norwell, MA Kluwer Academic Publishers, 2001.
- [Chu99] S. Chuprun, C. Bergstrom and D. Torrieri. Adaptive Spectrum Exploitation Using Emerging Software Defined Radios. *IEE Radio and Wireless Conference RAWCON 99*, pp. 113–116, 1999.
- [Cof01] D. Coffing and E. Main. Effects of offsets on bipolar integrated circuit mixer even-order distortion terms. *IEEE Transactions on Microwave Theory and Techniques*, vol. 49(1):pp. 23 – 30, Jan 2001.
- [Cro98] J. Crols and M. S. J. Steyaert. Low-IF topologies for high-performance analog front ends of fully integrated receivers. *IEEE Trans. Circuits Syst. II*, vol. 45:pp. 269–282, March 1998.
- [Dar00] H. Darabi and A. A. Abidi. Noise in RF-CMOS mixers: a simple physical model. *IEEE Journal of Solid-State Circuits*, vol. 35(1):pp. 15 – 25, Jan 2000.
- [Den83] J. E. Dennis and R. B. Schnabel. *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*. SIAM, 1983.
- [Dra92] G. Drago, A. Manella, M. Nervi, M. Repetto and G. Secondo. A Combined Strategy for Optimization in Non Linear Magnetic Problems Using Simulated Annealing and Search Techniques. *IEEE Transactions on Magnetics*, vol. 28(2):pp. 1541–1544, March 1992.
- [Edw64] H. M. Edwards. A Generalized Sturm Theorem. *The Annals of Mathematics*, vol. 80(1):pp. 22–57, Jul 1964.
- [Fer04] T. Fernández, C. Barquinero, J. Casajús, J. A. Beracoechea and I. Raos. Measurements and compensation techniques for multiantenna RF front end version 1. Tech. rep., Acorde, UPM, Dec. 2004.
- [Fin04] S. Finnerty and S. Sen. Simulated annealing based classification. *Sixth International Conference on Tools with Artificial Intelligence, 1994. Proceedings.*, (Nov):pp. 824 – 827, 2004.

- [Fos98] G. J. Foschini and M. J. Gans. On Limits of Wireless Communications in a Fading Environment when Using Multiple Antennas. *Wireless Personal Communications*, vol. 6(3):pp. 311–335, Mar. 1998.
- [Fre94] M. Freking. *Digital Signal Processing in Communication Systems*. New York: Chapman & Hall, 1994.
- [Gab97] T. J. Gabara and W. C. Fischer. Capacitive coupling and quantized feedback applied to conventional CMOS technology. *IEEE Journal of Solid-State Circuits*, vol. 32:pp. 419 – 427, March 1997.
- [Gar07] L. G. García. *Contribution to the characterization and evaluation of multiple antenna systems for communications: prototyping, propagation and antenna perspective*. phd thesis, Universidad Politécnica de Madrid, Escuela Técnica Superior de Ingenieros de Teledomunicación, Dec. 2007.
- [Gil05] G.-T. Gil, I.-H. Sohn, J.-K. Park and Y. Lee. Joint ML estimation of carrier frequency, channel, I/Q mismatch, and DC offset in communication receivers. *IEEE Transactions on Vehicular Technology*, vol. 54:pp. 338 – 349, Jan 2005.
- [Gla98] J. Glas. Digital I/Q imbalance compensation in a low-IF receiver. *IEEE Global Telecommunications Conference, 1998. GLOBECOM 98. The Bridge to Global Integration.*, vol. 3:pp. 1461 – 1466, Nov. 1998.
- [Gra95] P. P. Gray and R. G. Meyer. Future Directions in Silicon ICs for RF Personal Communications. *Custom Integrated Circuits Conference*, pp. 83–90, 1995.
- [Hay00] A. R. Hayes, Z. Ghassemlooy, N. L. Seed and R. McLaughlin. Baseline-wander effects on systems employing digital pulse-interval modulation. *IEE Proceedings-Optoelectronics*, vol. 147:pp. 295 – 300, Aug. 2000.
- [Hen99] T. Hentschel, M. Henker and G. Fettweis. The Digital Front-End of Software Radio Terminals. *IEEE Personal Communications*, pp. 40–46, August 1999.
- [Hot04] M. Hotti, J. Ryynanen, K. Kivekas and K. Halonen. An IIP2 calibration technique for direct conversion receivers. *Proceedings of the 2004 International Symposium on Circuits and Systems, ISCAS '04.*, vol. 4:pp. IV – 257–260, May 2004.
- [Hua04] X. Huang. On transmitter gain/phase imbalance compensation at receiver. *IEEE Commun. Lett.*, vol. 4:pp. 363–365, Nov 2004.

- [Hul93] C. D. Hull and R. G. Meyer. A systematic approach to the analysis of noise in mixers. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, vol. 40(12):pp. 909 – 919, Dec 1993.
- [Hwa04] M.-W. Hwang and et al. A high IIP2 direct-conversion mixer using an even-harmonic reduction technique for cellular CDMA/PCS/GPS applications. *Digest of Papers. 2004 IEEE Radio Frequency Integrated Circuits (RFIC) Symposium*, pp. 39 – 42, June 2004.
- [Ing89] L. Ingber. Very fast simulated re-annealing. Lester Ingber Papers 89vfsr, Lester Ingber, 1989. Available at <http://ideas.repec.org/p/lei/ingber/89vfsr.html>.
- [Ing92] L. Ingber and B. Rosen. Genetic algorithms and very fast simulated reannealing: a comparison. *Mathematical and Computer Modelling*, pp. 87–100, 1992.
- [Ing93] L. Ingber. Simulated Annealing: Practice versus theory. *Mathl. Comput. Modelling*, vol. 18(11):pp. 29–57, 1993.
- [Ing96] L. Ingber. Adaptive simulated annealing (ASA): Lessons learned. Lester Ingber Papers 96as, Lester Ingber, 1996. Available at <http://ideas.repec.org/p/lei/ingber/96as.html>.
- [ITU97] ITU. *Land Mobile (including Wireless Access) Volume 2: Principles and Approaches on Evolution to IMT-2000/FPLMTS*. ITU, 1997.
- [ITU03] ITU. *Migration to IMT-2000 Systems - Supplement 1 to the Handbook on Deployment of IMT-2000 Systems*. ITU, 2003.
- [Joh87] D. Johnson, C. Aragon, L. McGeoch and C. Schevon. Optimization by simulated annealing: An experimental evaluation (Parts 1 and 2). Tech. rep., AT&T Bell Laboratories, Murray Hill, NJ, 1987.
- [Joh89] D. S. Johnson, C. R. Aragon, L. McGeoch and C. Schevon. Optimization by simulated annealing: An experimental evaluation. *Operations Research*, pp. 865–892, 1989.
- [Ken00] P. B. Kenington and L. Astier. Power Consumption of A/D Converters for Software Radio Applications. *IEEE Transactions on vehicular Technology*, vol. 49(2):pp. 643–650, March 2000.

- [Kil04] P. S. Kildal and K. Rosengren. Correlation and capacity of MIMO systems and mutual coupling, radiation efficiency, and diversity gain of their antennas: simulations and measurements in a reverberation chamber. *IEEE Communications Magazine*, vol. 42(12):pp. 104 – 112, Dec 2004.
- [Kir83] S. Kirkpatrick, C. D. Galett and M. P. Vecchi. Optimisation by Simulated Annealing. *Science*, vol. 220(4598):pp. 621–680, May 1983.
- [Kiv01] K. Kivekas, A. Parssinen and K. A. I. Halonen. Characterization of IIP2 and DC-offsets in transconductance mixers. *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, vol. 48(11):pp. 1028 – 1038, Nov. 2001.
- [Kiv02] K. Kivekas, A. Parssinen, J. Ryynanen and J. Jussila. Calibration techniques of active BiCMOS mixers. *IEEE Journal of Solid-State Circuits*, vol. 37(6):pp. 766 – 769, June 2002.
- [Kol96] T. E. Kolding. Multi-standard mixed-signal transceivers for wireless communications - A research overview. Tech. rep., Department of Communication Technology, Aalborg University, Aalborg, Denmark, Tech. Rep. R-96-1005, Dec. 1996.
- [Kra97] B. Kraemer and et al. Advances in Semiconductor Technology Enabling Software Radio. *SW Radio Workshop, Brussels, Belgium*, May 1997.
- [Laa87] P. J. M. van Laarhoven and E. Aarts. *Simulated Annealing: Theory and Applications*. D. Reidel, 1987.
- [Lac95] R. L. Lackey and D. W. Upmal. Speakeasy: The Military Software Radio. *IEEE Commun. Mag.*, pp. 56–61, May 1995.
- [Lad01] M. Laddomada, F. Daneshgaran, M. Mondin and R. Hickling. A PC-based software receiver using a novel front-end technology. *IEEE Communications Magazine*, vol. 39(8):pp. 136–145, Aug. 2001.
- [Lau99] S. Laursen. Second order distortion in CMOS direct conversion receivers for GSM. *Proc. ESSCIRC 1999*, pp. 342–345, Sept 1999.
- [Lee98] T. H. Lee. *The Design of CMOS Radio-Frequency Integrated Circuits*. Cambridge University Press, 1998.
- [Lev04] P. Levinson. *Cellphone: The Story of the World's Most Mobile Medium and How It Has Transformed Everything!*. New York: Macmillan, 2004.

- [Li99] Y. Li and W. M. Snelgrove. A novel adaptive mismatch cancellation system for quadrature IF radio receivers. *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, vol. 46(6):pp. 789 – 801, june 1999.
- [Li02] Y. Li, J. Yao and D. Yao. An Efficient Composite Simulated Annealing Algorithm for Global Optimization. *IEEE 2002 International Conference on Communications, Circuits and Systems and West Sino Expositions*, vol. 21:pp. 1165– 1169, July 2002.
- [Li07] L.-H. Li, F.-L. Lin and H.-R. Chuang. Complete RF-System Analysis of Direct Conversion Receiver (DCR) for 802.11a WLAN OFDM System. *IEEE Transactions on Vehicular Technology*, vol. 56:pp. 1696 – 1703, July 2007.
- [Liu98a] C.-L. Liu. Impacts of I/Q imbalance on QPSK-OFDM-QAM detection. *IEEE Transactions on Consumer Electronics*, vol. 44(3):pp. 984 – 989, Aug. 1998.
- [Liu98b] W. Liu and et al. BSIM3v3.2 MOSFET model user's manual. Tech. rep., The Regents of the University of California, 1998.
- [Liu04] J. Liu, A. Bourdoux and et al. Impact of front-end effects on the performance of downlink OFDM-MIMO transmission. *Proc. IEEE Radio and Wireless Conf*, pp. 159–162, Sept 2004.
- [Liw00] S. Liwei, J. C. Jensen and L. E. Larson. A wide-bandwidth Si/SiGe HBT direct conversion sub-harmonic mixer/downconverter. *IEEE Journal of Solid-State Circuits*, vol. 35(9):pp. 1329 – 1337, Sept. 2000.
- [Lok02a] A. Loke and F. Ali. Direct conversion radio for digital mobile phones-design issues, status, and trends. *IEEE Transactions on Microwave Theory and Techniques*, vol. 50:pp. 2422 – 2435, Nov 2002.
- [Lok02b] A. Loke and F. Loke. Direct conversion radio for digital mobile phones-design issues, status, and trends. *IEEE Transactions on Microwave Theory and Techniques*, vol. 50:pp. 2422 – 2435, Nov. 2002.
- [Lu99] J. Lu, K. B. Letaief, J. C.-I. Chuang and M. L. Liou. M-PSK and M-QAM BER Computation Using Signal-Space Concepts. *IEEE Transactions on Communications*, vol. 47(2):pp. 181 – 184, February 1999.

- [Lu00] W. W. Lu. Compact multidimensional broadband wireless: The convergence of wireless mobile and access. *IEEE Communications Magazine*, vol. 38:pp. 119–123, Nov. 2000.
- [Man02] D. Manstretta and F. Svelto. Analysis and optimization of IIP2 in CMOS direct down-converters. *Proceedings of the IEEE 2002 Custom Integrated Circuits Conference*, pp. 243 – 246, May 2002.
- [Man03] D. Manstretta, M. Brandolini and F. Svelto. Second-order distortion mechanisms in CMOS downconverters. *IEEE Journal of Solid-State Circuits*, vol. 38(3):pp. 394–406, March 2003.
- [Mat00] B. Matinpour, S. Chakraborty, M. Hamai, C. Chun and J. Laskar. A Novel DC-offset Cancellation Technique for Even-Harmonic Direct Conversion Receivers. *2000 IEEE MTT-S Digest*, vol. 1:pp. 631–634, 2000.
- [Men97] P. Mendonca and L. Caloba. New simulated annealing algorithms. *Proceedings of 1997 IEEE International Symposium on Circuits and Systems, 1997. ISCAS '97.*, vol. 3:pp. 1668 – 1671, Jan 1997.
- [Met53] N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller and E. Teller. Equation of state calculations by fast computing machines. *J. Chem. Phys.*, vol. 21(6):pp. 1087–1092, 1953.
- [Mey86] R. G. Meyer. Intermodulation in high-frequency bipolar transistor integrated-circuit mixers. *IEEE Journal of Solid-State Circuits*, vol. 21(4):pp. 534 – 537, Aug. 1986.
- [Mil03] E. Milotti. *1/f noise: a pedagogical review*, dipartimento di fisica, università di udine and i.n.f.n. - sezione di trieste ed., 2003.
- [Mir00a] S. Mirabbasi and K. Martin. Classical and modern receiver architectures. *IEEE Communications Magazine*, vol. 38(11):pp. 132 – 139, Nov 2000.
- [Mir00b] S. Mirabbasi and K. Martin. Classical and modern receiver architectures. *IEEE Commun. Mag.*, vol. 38:pp. 132–139, Nov. 2000.
- [Mit95] J. Mitola. The Software Radio Architecture. *IEEE Communications Magazine*, vol. 33(5):pp. 26–38, May 1995.
- [Mit97] J. Mitola. Software Radio Technology Challenges and Opportunities. *SW Radio Workshop, Brussels, Belgium*, May 1997.

- [Nam00] W. Namgoong. Performance of a direct-conversion receiver with AC coupling. *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, vol. 47:pp. 1556 – 1559, Dec 2000.
- [Nam01] W. Namgoong and T. H. Meng. Direct-conversion RF receiver design. *IEEE Transactions on Communications*, vol. 49(3):pp. 518 – 529, March 2001.
- [Nik04] K. Nikitopoulos and A. Polydoros. Decision-directed compensation of phase noise and residual frequency offset in a space-Time OFDM receiver. *IEEE Communications Letters*, vol. 8:pp. 573 – 575, Sept. 2004.
- [Nol86] M. Noll. *Introduction to Telephones and Telephone Systems*. Artech House, 1986.
- [Ou99] J. Ou, J. Xiaodong, H. Chenming and P. Gray. Submicron CMOS thermal noise modeling from an RF perspective. *Symposium on VLSI Technology, 1999. Digest of Technical Papers.*, pp. 151 – 152, June 1999.
- [Pha05] A.-T. Phan and et al. A novel 1.5V DC offset cancellation CMOS down conversion mixer. *IEEE International Symposium on Circuits and Systems, ISCAS 2005*, vol. 4:pp. 3704 – 3707, May 2005.
- [Pol95] T. Pollet, M. V. Bladel and M. Moeneclaey. BER sensitivity of OFDM systems to carrier frequency offset and Wiener phase noise. *IEEE Trans. Commun.*, vol. 43(2/3/4):pp. 191–193, Feb./Mar./Apr. 1995.
- [Pre02] H. Pretl. RF integrated homodyne transceivers for 3G mobile radio. *Workshop IEEE International Microwave Symposium, Seattle 2002*, pp. –, 2002.
- [Pun01] K. P. Pun and et al. Correction of frequency dependent I/Q mismatches in quadrature receivers. *IEEE Electron. Lett.*, vol. 37(23):pp. 1415–1417, Nov. 2001.
- [Rad04] Radio Science Conference. *The Key Issues to Design Software Radio*. IEEE, 2004.
- [Rao04] R. M. Rao and B. Daneshrad. Analysis and correlation of I/Q mismatch for OFDM systems. *Proc. IST Mobile Communication Summit*, vol. 2:pp. 621–626, June 2004.
- [Rao06a] R. M. Rao and B. Daneshrad. Analog Impairments in MIMO-OFDM Systems. *IEEE Transactions on Wireless Communications*, vol. 5(12):pp. 3382 – 3387, Dec 2006.
- [Rao06b] I. Raos, S. Zazo and I. Arambasic. Slepian pulses for Multicarrier OQAM. *14th European Signal Processing Conference, EUSIPCO 2006*, Sept. 2006.

- [Raz97] B. Razavi. Design considerations for direct-conversion receivers. *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, vol. 44(6):pp. 428 – 435, June 1997.
- [Raz98a] B. Razavi. Architectures and circuits for RF CMOS receivers. *Proceedings of the IEEE 1998 Custom Integrated Circuits Conference*, pp. 393 – 400, May 1998.
- [Raz98b] B. Razavi. RF IC design challenges. *Proceedings Design Automation Conference, 1998.*, pp. 408 – 413, June 1998.
- [Raz99] B. Razavi. *Design of Analog CMOS Integrated Circuits*. Boston, MA: McGraw-Hill, 1999.
- [Ren96] T. Renyuan, Y. Shiyou, L. Yan, W. Geng and M. Tiemin. Combined strategy of improved simulated annealing and genetic algorithm for inverse problem. *IEEE Transactions on Magnetics*, vol. 32:pp. 1326 – 1329, May 1996.
- [Rob95] P. Robertson and S. Kaiser. Analysis of the effects of phase noise in orthogonal frequency division multiplex (OFDM) systems. *IEEE International Conference on Communications, 1995. ICC 95 Seattle, Gateway to Globalization*, vol. 3:pp. 1652 – 1657, 1995.
- [Ros92] B. Rosen. Function Optimization Based On Advanced Simulated Annealing. *Workshop on Physics and Computation, 1992*, pp. 289–293, Oct. 1992.
- [Rut71] J. Rutman. A relation between frequency instability and RF spectrum and its application to a thermal noise perturbated oscillator. *Proceedings of the IEEE*, vol. 59:pp. 99 – 101, Jan 1971.
- [Sac03] E. Sacchi, I. Bietti, S. Erba, L. Tee, P. Vilmercati and R. Castello. A 15 mW, 70 kHz 1/f corner direct conversion CMOS receiver. *Proceedings of the IEEE 2003 Custom Integrated Circuits Conference*, pp. 459 – 462, Sept 2003.
- [Sam92] S. Sampei and K. Feher. Adaptive DC-offset compensation algorithm for burst mode operated direct conversion receivers. *1992 IEEE 42nd Vehicular Technology Conference*, vol. 1:pp. 93 – 96, May 1992.
- [Sch93] M. Schiess and B. Johansson. Excess noise in heterodyne receivers. *IEEE Proceedings Optoelectronics*, pp. 217 – 220, June 1993.

- [Sch01a] A. Schuchert and R. Hasholzner. A novel I/Q imbalance compensation scheme for the reception of OFDM signals. *IEEE Trans. Consumer Electron*, vol. 47(3):pp. 313–318, Aug. 2001.
- [Sch01b] A. Schuchert, R. Hasholzner and P. Antoine. A novel IQ imbalance compensation scheme for the reception of OFDM signals. *IEEE Transactions on Consumer Electronics*, vol. 47(3):pp. 313 – 318, 2001.
- [Sch04a] T. Schenk, T. Xiao-Jiao, P. Smulders and E. Fledderus. Influence and suppression of phase noise in multi-antenna OFDM. *IEEE 60th Vehicular Technology Conference, 2004. VTC2004-Fall.*, vol. 2:pp. 1443 – 1447, Sept. 2004.
- [Sch04b] A. J. Scholten and et al. Compact modelling of noise for RF CMOS circuit design. *IEE Proceedings Circuits, Devices and Systems*, vol. 151:pp. 167 – 174, April 2004.
- [Sch07] H. Schmid. Aaargh! I just loooove flicker noise. *IEEE Circuits and Systems Magazine*, vol. 7(1):pp. 32 – 35, 2007.
- [Sev00] J. Sevenhuijsen, B. Verstraeten and S. Taraborrelli. Trends in silicon radio large scale integration: Zero IF receiver! Zero I&Q transmitter! Zero discrete passives! *IEEE Commun. Mag.*, vol. 38:pp. 142–147, Jan. 2000.
- [Sha48] C. Shannon. A Mathematical Theory of Communication. *Bell System Technical Journal*, vol. 27:pp. 379–423, 623–656, July, October 1948.
- [Sha03a] S. Shang, S. Mirabbasi and R. Saleh. A technique for DC-offset removal and carrier phase error compensation in integrated wireless receivers. *ISCAS '03. Proceedings of the 2003 International Symposium on Circuits and Systems, 2003.*, vol. 1:pp. 173–176, May 2003.
- [Sha03b] S. Shang, S. Mirabbasi and R. Saleh. A technique for DC-offset removal and carrier phase error compensation in integrated wireless receivers. *Proceedings of the 2003 International Symposium on Circuits and Systems, ISCAS '03.*, vol. 1:pp. 173–176, May 2003.
- [Shi00] D.-S. Shiu, G. J. Foschini, M. J. Gans and J. M. Kahn. Fading correlation and its effect on the capacity of multielement antenna systems. *IEEE Transactions on Communications*, vol. 48(3):pp. 502 – 513, March 2000.

- [Siv05] P. Sivonen, A. Vilander and A. Parssinen. Cancellation of second-order intermodulation distortion and enhancement of IIP2 in common-source and common-emitter RF transconductors. *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 52(2):pp. 305 – 317, Feb. 2005.
- [Svi05] R. Svitek and S. Raman. DC Offsets in Direct-Conversion Receivers: Characterization and Implications. *IEEE Microwave Magazine*, pp. 76–86, Sept 2005.
- [Szu87] H. H. Szu and R. L. Hartley. Nonconvex optimization by fast simulated annealing. *Proceedings of the IEEE*, vol. 75(11):pp. 1538 – 1540, Nov. 1987.
- [Tar05a] A. Tarighat, R. Bagheri and A. H. Sayed. Compensation schemes and performance analysis of IQ imbalances in OFDM receivers. *IEEE Trans. Signal Processing*, vol. 53(8):pp. 3257–3268, Aug. 2005.
- [Tar05b] A. Tarighat, R. Bagheri and A. H. Sayed. Compensation schemes and performance analysis of IQ imbalances in OFDM receivers. *IEEE Trans. Signal Processing*, vol. 53(8):pp. 3257–3268, Aug 2005.
- [Tar05c] A. Tarighat and A. H. Sayed. MIMO OFDM receivers for systems with IQ imbalances. *IEEE Trans. Signal Processing*, pp. 3583–3596, Sept 2005.
- [Ter00] M. T. Terrovitis and R. G. Meyer. Intermodulation distortion in current-commutating CMOS mixers. *IEEE Journal of Solid-State Circuits*, vol. 35(10):pp. 1461 – 1473, Oct 2000.
- [Tsi99] J. Tsividis. *Operation and Modeling of the MOS Transistors*. Second Edition, Boston: McGraw-Hill, 1999.
- [Val00a] M. Valkama and M. Renfors. Advanced DSP for I/Q imbalance compensation in a low-IF receiver. *2000 IEEE International Conference on Communications, ICC 2000*, vol. 2:pp. 768 – 772, June 2000.
- [Val00b] M. Valkama, M. Renfors and V. Koivunen. Blind Source Separation Based I/Q Imbalance Compensation. *The IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium 2000. AS-SPCC.*, pp. 310–314, Oct. 2000.
- [Val00c] M. Valkama, M. Renfors and V. Koivunen. On the performance of interference canceller based I/Q imbalance compensation. *Proceedings. 2000 IEEE Interna-*

- tional Conference on Acoustics, Speech, and Signal Processing, 2000. ICASSP '00.,* vol. 50:pp. 2885 – 2888 vol.5, June 2000.
- [Val01] M. Valkama, M. Renfors and V. Koivunen. Advanced methods for I/Q imbalance compensation in communication receivers. *IEEE Transactions on Signal Processing*, vol. 49(10):pp. 2335 – 2344, Oct 2001.
- [Val02] M. Valkama, K. Salminene and M. Renfors. Digital I/Q Imbalance Compensation in Low-IF Receivers: Principles and Practice. *14th International Conference on Digital Signal Processing, DSP 2002*, vol. 2:pp. 1179–1182, July 2002.
- [Vou93] K. Voudouris and J. Noras. Effects of amplitude, phase, and frequency imperfections on the performance of a direct conversion receiver (DCR) for personal communication systems. *IEEE Microwave and Guided Wave Letters*, vol. 3:pp. 313 – 315, Sept. 1993.
- [Wal99] J. P. Walser. *Integer Optimization by Local Search : A Domain-Independent Approach*. Lecture Notes in Artificial Intelligence 1637. Springer, 1999.
- [Wan03] H. Wang, M. Lin, Y. Li and H. Chen. A novel dynamic DC-offset canceller. *5th International Conference on ASIC, 2003. Proceedings*, vol. 1:pp. 639 – 642, Oct 2003.
- [Wep95] J. A. Wepman. Analog-to-digital converters and their applications in radio deceivers. *IEEE Commun. Mag.*, vol. 33:pp. 39–45, May 1995.
- [Win04] M. Windish and G. Fettweis. Adaptive I/Q imbalance compensation in low-IF transmitter architectures. *Proc. IEEE Vehicular Technology Conference*, vol. 30:pp. 2096–2100, Sept 2004.
- [Xu05] X. Xu, R. B. and K. Wu. Analysis and implementation of software defined radio receiver platform. *Microwave Conference Proceedings, 2005. APMC 2005. Asia-Pacific Conference Proceedings*, vol. 5, Dec. 2005.
- [Yam98] T. Yamaji, H. Tanimoto and H. Kokatsu. An I/Q active balanced harmonic mixer with IM2 cancelers and a 45° phase shifter. *IEEE Journal of Solid-State Circuits*, vol. 33(12):pp. 2240 – 2246, Dec 1998.
- [Yos98] H. Yoshida, H. Tsurumi and Y. Suzuki. DC offset canceller in a direct conversion receiver for QPSK signal reception. *The Ninth IEEE International Symposium on*

Personal, Indoor and Mobile Radio Communications, 1998., vol. 3:pp. 1314 – 1318,
Sept 1998.

- [Zer94] J. Zerubia, Z. Kato and M. Berthod. Multi-temperature annealing: a new approach for the energy-minimization of hierarchical Markov random field models. *Proceedings of the 12th IAPR International Conference on Pattern Recognition, 1994. Vol. 1 - Conference A: Computer Vision & Image Processing*, vol. 1:pp. 520–522, Oct. 1994.